Cheng-Few Lee
John C. Lee
Editors

Handbook of Financial Econometrics and Statistics

Springer Reference

Handbook of Financial Econometrics and Statistics

Cheng-Few Lee • John C. Lee
Editors

Handbook of Financial Econometrics and Statistics

With 281 Figures and 490 Tables

Springer Reference

Editors
Cheng-Few Lee
Department of Finance and Economics, Rutgers Business School
Rutgers, The State University of New Jersey
Piscataway, NJ, USA

and

Graduate Institute of Finance
National Chiao Tung University
Hsinchu, Taiwan

John C. Lee
Center for PBBEF Research
North Brunswick, NJ, USA

ISBN 978-1-4614-7749-5 ISBN 978-1-4614-7750-1 (eBook)
ISBN 978-1-4614-7751-8 (print and electronic bundle)
DOI 10.1007/978-1-4614-7750-1
Springer New York Heidelberg Dordrecht London

Library of Congress Control Number: 2014940762

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Preface

Financial econometrics and statistics have become very important tools for empirical research in both finance and accounting. Econometric methods are important tools for doing asset pricing, corporate finance, options and futures, and conducting financial accounting research. Important econometric methods used in this research include: single equation multiple regression, simultaneous regression, panel data analysis, time series analysis, spectral analysis, non-parametric analysis, semi-parametric analysis, GMM analysis, and other methods.

Portfolio theory and management research have used different statistical distributions, such as normal distribution, stable distribution, and log normal distribution. Options and futures research have used binomial distribution, log normal distribution, non-central chi square distribution, Poission distribution, and others. Auditing research has used sampling survey techniques to determine the sampling error and non-sampling error for auditing.

Based upon our years of experience working in the industry, teaching classes, conducting research, writing textbooks, and editing journals on the subject of financial econometrics and statistics, this handbook will review, discuss, and integrate theoretical, methodological, and practical issues of financial econometrics and statistics. There are 99 chapters in this handbook. Chapter 1 presents an introduction of financial econometrics and statistics and shows how readers can use this handbook. The following chapters, which have been contributed by accredited authors, can be classified by the following 14 topics.

i. Financial Accounting (Chapters 2, 9, 10, 61, 97)
ii. Mutual Funds (Chapters 3, 24, 25, 68, 88)
iii. Microstructure (Chapters 4, 44, 96, 99)
iv. Corporate Finance (Chapters 5, 21, 30, 38, 42, 46, 60, 63, 75, 79, 95)
v. Asset Pricing (Chapters 6, 15, 22, 28, 34, 36, 39, 45, 47, 50, 81, 85, 87, 93)
vi. Options (Chapters 7, 32, 37, 55, 65, 84, 86, 90, 98)
vii. Portfolio Analysis (Chapters 8, 26, 35, 53, 67, 73, 80, 83)
viii. Risk Management (Chapters 11, 13, 16, 17, 23, 27, 41, 51, 54, 72, 91, 92)
ix. International Finance (Chapters 12, 40, 43, 59, 69)
x. Event Study (Chapters 14)
xi. Methodology (Chapters 18, 19, 20, 29, 31, 33, 49, 52, 56, 57, 58, 62, 74, 76, 77, 78, 82, 89)

xii. Banking Management (Chapters 64)
xiii. Pension Funds (Chapters 66)
xiv. Futures and Index Futures (Chapters 48, 70, 71, 94)

In addition to this classification, based upon the keywords of chapter 2-99, we classify the information into a) finance and accounting topics and b) methodology topics. This information can be found in chapter 1 of this handbook.

In the preparation of this handbook, first, we would like to thank the member of advisory board and contributors of this handbook. In addition, we would like to make note that we appreciate the extensive help from the Editor Mr. Brian Foster, our research assistants Tzu Tai, Lianne Ng, and our secretary Ms. Miranda Mei-Lan Luo. Finally, we would like to thank the financial support from the Wintek Corporation and APEX International Financial Engineering Res. & Tech. Co. Ltd. that allowed us to write the edition of this book.

There are undoubtedly some errors in the finished product, both typo-graphical and conceptual. I would like to invite readers to send suggestions, comments, criticisms, and corrections to the author Professor Cheng F. Lee at the Department of Finance and Economics, Rutgers University at the email address lee@business.rutgers.edu.

December 2012

Cheng-Few Lee
John C. Lee

Advisory Board

About the Editors

Cheng-Few Lee is a Distinguished Professor of Finance at Rutgers Business School, Rutgers University and was chairperson of the Department of Finance from 1988–1995. He has also served on the faculty of the University of Illinois (IBE Professor of Finance) and the University of Georgia. He has maintained academic and consulting ties in Taiwan, Hong Kong, China and the United States for the past four decades. He has been a consultant to many prominent groups including, the American Insurance Group, the World Bank, the United Nations, The Marmon Group Inc., Wintek Corporation, and Polaris Financial Group.

Professor Lee founded the *Review of Quantitative Finance and Accounting* (RQFA) in 1990 and the *Review of Pacific Basin Financial Markets and Policies* (RPBFMP) in 1998, and serves as managing editor for both journals. He was also a co-editor of the *Financial Review* (1985–1991) and the *Quarterly Review of Economics and Finance* (1987–1989). In the past 39 years, Dr. Lee has written numerous textbooks ranging in subject matters from financial management to corporate finance, security analysis and portfolio management to financial analysis, planning and forecasting, and business statistics. In addition, he edited two popular books, *Encyclopedia of Finance* (with Alice C. Lee) and *Handbook of Quantitative Finance and Risk Management* (with Alice C. Lee and John Lee). Dr. Lee has also published more than 200 articles in more than 20 different journals in finance, accounting, economics, statistics, and management. Professor Lee was ranked the most published finance professor worldwide during the period 1953–2008.

Professor Lee was the intellectual force behind the creation of the new Masters of Quantitative Finance program at Rutgers University. This program began in 2001 and has been ranked as one of the top ten quantitative finance programs in the United States. These top ten programs are located at Carnegie Mellon University, Columbia University, Cornell University, New York University, Princeton University, Rutgers University, Stanford University, University of California at Berkley, University of Chicago, and University of Michigan.

John C. Lee is a Microsoft Certified Professional in Microsoft Visual Basic and Microsoft Excel VBA. He has a Bachelor and Masters degree in accounting from the University of Illinois at Urbana-Champaign.

John has worked over 20 years in both the business and technical fields as an accountant, auditor, systems analyst and as a business software developer. He is the author of the book on how to use MINITAB and Microsoft Excel to do statistical analysis which is a companion text to *Statistics of Business and Financial Economics*, 2nd and 3rd, of which he is one of the co-authors. In addition, he has also coauthored the textbooks *Financial Analysis*, *Planning and Forecasting*, 2ed (with Cheng F. Lee and Alice C. Lee), and *Security Analysis*, *Portfolio Management*, *and Financial Derivatives* (with Cheng F. Lee, Joseph Finnerty, Alice C. Lee, and Donald Wort). John has been a Senior Technology Officer at the Chase Manhattan Bank and Assistant Vice President at Merrill Lynch. Currently, he is the Director of the Center for PBBEF Research.

Contents

Volume 1

Volume 4

Contributors

Shigeyuki Abe Faculty of Policy Studies, Doshisha University, Kyoto, Japan

Raj Aggarwal University of Akron, Akron, OH, USA

Mercedes Alda García Facultad de Economía y Empresa, Departamento de Contabilidad y Finanzas, Universidad de Zaragoza, Zaragoza, Spain

Heitor Almeida University of Illinois at Urbana-Champaign, Champaign, IL, USA

James S. Ang Department of Finance, College of Business, Florida State University, Tallahassee, FL, USA

Augustine C. Arize Texas A & M University-Commerce, Commerce, TX, USA

Zhidong D. Bai KLAS MOE & School of Mathematics and Statistics, Northeast Normal University, Changchun, China

Department of Statistics and Applied Probability, National University of Singapore, Singapore, Singapore

Gurdip S. Bakshi Department of Finance, College of Business, University of Maryland, College Park, MD, USA

Turan G. Bali McDonough School of Business, Georgetown University, Washington, DC, USA

Alok Bhargava School of Public Policy, University of Maryland, College Park, MD, USA

Carl S. Bonham College of Business and Public Policy, University of Alaska Anchorage, Anchorage, AK, USA

Ivan E. Brick Department of Finance and Economics, Rutgers, The State University of New Jersey, Newark/New Brunswick, NJ, USA

Matthew D. Brigida Department of Finance, Clarion University of Pennsylvania, Clarion, PA, USA

Murillo Campello Cornell University, Ithaca, NY, USA

Charles Cao Department of Finance, Smeal College of Business, Penn State University, University Park, PA, USA

Oscar Carchano Department of Financial Economics, University of Valencia, Valencia, Spain

Jow-Ran Chang National Tsing Hua University, Hsinchu City, Taiwan

Shih-Kang Chao Ladislaus von Bortkiewicz Chair of Statistics, C.A.S.E. – Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin, Berlin, Berlin, Germany

An-Chi Chen KGI Securities Co. Ltd., Taipei, Taiwan

Hong-Yi Chen Department of Finance, National Central University, Taoyuan, Taiwan

Li-jiun Chen Department of Finance, Feng Chia University, Taichung City, Taiwan

Li-Shya Chen Department of Statistics, National Cheng-Chi University, Taipei City, Taiwan

Ren-Raw Chen Graduate School of Business Administration, Fordham University, New York, NY, USA

Rong Chen Department of Finance, Xiamen University, Xiamen, China

Sheng-Syan Chen National Central University, Zhongli City, Taiwan

Zhiwu Chen School of Management, Yale University, New Haven, USA

Anna Chernobai Department of Finance, M.J. Whitman School of Management, Syracuse University, Syracuse, NY, USA

Thomas C. Chiang Department of Finance, Drexel University, Philadelphia, PA, USA

Wan-Jiun Paul Chiou Department of Finance and Law College of Business Administration, Central Michigan University, Mount Pleasant, MI, USA

Chun-Yuan Chiu National Chiao–Tung University, Taiwan, Republic of China

Institute of Information Management, National Chiao Tung University, Taiwan, Republic of China

Hengchih Chou Department of Shipping and Transportation Management, National Taiwan Ocean University, Keelung, Taiwan

Ray Yeutien Chou Institute of Economics, Academia Sinica and National Chiao Tung University, Taipei, Taiwan

Win Lin Chou Department of Economics and Finance, City University of Hong Kong, Hong Kong, China

Huimin Chung Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan

Michael L. Clutter Warnell School of Forestry and Natural Resources, University of Georgia, Athens, GA, USA

Richard Cohen University of Hawaii Economic Research Organization and Economics, University of Hawaii at Manoa, Honolulu, HI, USA

Tian-Shyr Dai National Chiao-Tung University, Taiwan, Republic of China

Huong Dang University of Canterbury, Christchurch, New Zealand

Darinka Dentcheva Department of Mathematical Sciences, Stevens Institute of Technology, Hoboken, NJ, USA

Dean Diavatopoulos Finance, Villanova University, Villanova, PA, USA

Gang Nathan Dong Columbia University, New York, NY, USA

Diep Duong Department of Business and Economics, Utica College, Utica, NY, USA

Frank J. Fabozzi EDHEC Business School, EDHEC Risk Institute, Nice, France

Luis Ferruz Facultad de Economía y Empresa, Departamento de Contabilidad y Finanzas, Universidad de Zaragoza, Zaragoza, Spain

Wayne E. Ferson University of Southern California, Los Angeles, CA, USA

Cheng-der Fuh Graduate Institute of Statistics, National Central University, Zhongli City, Taiwan

Antonio F. Galvao University of Iowa, Iowa City, IA, USA

Gerard L. Gannon Deakin University, Burwood, VIC, Australia

Fernando Gómez-Bezares Universidad de Deusto, Bilbao, Spain

John W. Goodell College of Business Administration, University of Akron, Akron, OH, USA

Hongtao Guo Bertolon School of Business, Salem State University, Salem, MA, USA

Jia-Hau Guo Institution of Finance, College of Management, National Chiao Tung University, Hsinchu, Taiwan

Manak C. Gupta Temple University, Philadelphia, PA, USA

Ahmed Hachicha Department of Economic Development, Faculty of Economics and Management of Sfax, University of Sfax, Sfax, Tunisia

Fatma Hachicha Department of Finance, Faculty of Economics and Management of Sfax, Sfax, Tunisia

Chuan-Hsiang Han Department of Quantitative Finance, National Tsing Hua University, Hsinchu, Taiwan, Republic of China

Wolfgang Karl Härdle Ladislaus von Bortkiewicz Chair of Statistics, C.A.S.E. – Center for Applied Statistics and Economics, Humboldt–Universität zu Berlin, Berlin, Berlin, Germany

Lee Kong Chian School of Business, Singapore Management University, Singapore, Singapore

Richard Heaney Accounting and Finance, The University of Western Australia, Perth, Australia

Yuna Heo Rutgers Business School, Rutgers, The State University of New Jersey, Newark-New Brunswick, NJ, USA

Hemantha Herath Department of Accounting, Faculty of Business, Brock University, St. Catharines, ON, Canada

Thomas S. Y. Ho Thomas Ho Company Ltd, New York, NY, USA

Der-Tzon Hsieh Department of Economics, National Taiwan University, Taipei, Taiwan

Chin-Wen Hsin Yuan Ze University, Zhongli City, Taiwan

Chun-Pin Hsu Department of Accounting and Finance, York College, The City University of New York, Jamaica, NY, USA

Yu Chuan Hsu National Chi Nan University, Nantou, Taiwan

Jian Hua Baruch College (CUNY), New York, NY, USA

Bwo-Nung Huang National Chung-Cheng University, Minxueng Township, Chiayi County, Taiwan

Chin-Wen Huang Department of Finance, Western Connecticut State University, Danbury, CT, USA

Jing-Zhi Huang Smeal College of Business, Penn State University, University Park, PA, USA

Yongchang C. Hui School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, China

Ken Hung Division of International Banking & Finance Studies, Texas A&M International University, Laredo, TX, USA

Mao-Wei Hung College of Management, National Taiwan University, Taipei, Taiwan

John S. Jahera Jr. Department of Finance, College of Business, Auburn University, Auburn, AL, USA

Kiridaran Kanagaretnam Schulich School of Business, York University, Toronto, ON, Canada

Long Kang Department of Finance, Antai College of Economics and Management, Shanghai Jiao Tong University, Shanghai, China

The Options Clearing Corporation and Center for Applied Economics and Policy Research, Indiana University, Bloomington, IN, USA

Lie-Jane Kao Department of Finance and Banking, Kainan University, Taoyuan, ROC, Taiwan

Young Shin (Aaron) Kim College of Business, Stony Brook University, Stony Brook, NY, USA

Gary Kleinman Montclair State University, Montclair, NJ, USA

April Knill The Florida State University, Tallahassee, FL, USA

Alexander Kogan Rutgers Business School, Rutgers, The State University of New Jersey, Newark–New Brunswick, NJ, USA

Rutgers Center for Operations Research (RUTCOR), Piscataway, NJ, USA

Joshua Krausz Yeshiva University, New York, NY, USA

Wikil Kwak University of Nebraska at Omaha, Omaha, NE, USA

Tze Leung Lai Stanford University, Stanford, CA, USA

Miranda S. Lam Bertolon School of Business, Salem State University, Salem, MA, USA

Kenneth D. Lawrence New Jersey Institute of Technology, Newark, NJ, USA

Sheila M. Lawrence Rutgers, The State University of New Jersey, New Brunswick, NJ, USA

Alice C. Lee State Street Corp., USA

Cheng-Few Lee Department of Finance and Economics, Rutgers Business School, Rutgers, The State University of New Jersey, Piscataway, NJ, USA

Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan

Chia-Ju Lee College of Business, Chung Yuan University, Chungli, Taiwan

Chien-Hui Lee National Kaohsiung University of Applied Sciences, Kaohsiung, Taiwan

Geul Lee University of New South Wales, Sydney, Australia

Heeseok Lee Korea Advanced Institute of Science and Technology, Yuseong-gu, Daejeon, South Korea

Jang-Yi Lee Tunghai University, Taichung, Taiwan

Jieun Lee Economic Research Institute, Bank of Korea, Seoul, South Korea

John C. Lee Center for PBBEF Research, North Brunswick, NJ, USA

Kin-Wai Lee Division of Accounting, Nanyang Business School, Nanyang Technological University, Singapore, Singapore

Kuo-Hao Lee Department of Finance, College of Business, Bloomsburg University of Pennsylvania, Bloomsburg, PA, USA

Sang Bin Lee Hanyang University, Seong-Dong-Ku, Seoul, Korea

Miguel A. Lejeune George Washington University, Washington, DC, USA

Jiandong Li Chinese Academy of Finance and Development (CAFD) and Central University of Finance and Economics (CUFE), Beijing, China

Chan-Chien Lien Treasury Division, E.SUN Commercial Bank, Taipei, Taiwan

Donald Lien The University of Texas at San Antonio, San Antonio, TX, USA

Si Rou Lim National Chi Nan University, Nantou, Taiwan

Emily Lin St. John’s University, New Taipei City, Taiwan

Fang-Pang Lin National Center for High Performance Computing, Hsinchu, Taiwan

Hai Lin School of Economics and Finance, Victoria University of Wellington, Wellington, New Zealand

Hsuan-Chu Lin Graduate Institute of Finance and Banking, National Cheng-Kung University, Tainan, Taiwan

Nathan Liu Department of Finance, Feng Chia University, Taichung, Taiwan

Gerald J. Lobo C.T. Bauer College of Business, University of Houston, Houston, TX, USA

Anastasia Maggina Business Consultant/Research Scientist, Avlona, Attikis, Greece

Afif Masmoudi Department of Mathematics, Faculty of Sciences of Sfax, Sfax, Tunisia

Philippe Masset Ecole Hôtelière de Lausanne, Le-Chalet-à-Gobet, Lausanne 25, Switzerland

Robert Mathieu School of Business and Economics, Wilfrid Laurier University, Waterloo, ON, Canada

Bin Mei Warnell School of Forestry and Natural Resources, University of Georgia, Athens, GA, USA

Thomas Meinl Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany

Kristina L. Minnick Bentley University, Waltham, MA, USA

Heather Mitchell RMIT University, Melbourne, VIC, Australia

Bruce Mizrach Department of Economics, Rutgers, The State University of New Jersey, New Brunswick, NJ, USA

Kiseok Nam Yeshiva University, New York, NY, USA

Ali Nejadmalayeri Department of Finance, Oklahoma State University, Oklahoma, OK, USA

Shou Zhong Ng Hong Kong Monetary Authority, Hong Kong, China

Oded Palmon Department of Finance and Economics, Rutgers Business School – Newark and New Brunswick, Piscataway, NJ, USA

Dilip K. Patro RAD, Office of the Comptroller of the Currency, Washington, DC, USA

Robert L. Porter Department of Finance School of Business, Quinnipiac University, Hamden, CT, USA

Svetlozar T. Rachev Department of Applied Mathematics and Statistics, College of Business, Stony Brook University, SUNY, Stony Brook, NY, USA

FinAnalytica, Inc, New York, NY, USA

Vikash Ramiah School of Economics, Finance and Marketing, RMIT University, Melbourne, Australia

Raafat R. Roubi Department of Accounting, Faculty of Business, Brock University, St. Catharines, ON, Canada

Andrzej Ruszczynski Department of Management Science and Information Systems, Rutgers, The State University of New Jersey, Piscataway, NJ, USA

G.V. Satya Sekhar Department of Finance, GITAM Institute of Management, GITAM University, Visakhapatnam, Andhra Pradesh, India

Robert A. Schwartz Zicklin School of Business, Baruch College, CUNY, New York, NY, USA

Thomas V. Schwarz Stetson University, DeLand, FL, USA

Yuan-Chung Sheu National Chiao-Tung University, Hsinchu, Taiwan

Yong Shi University of Nebraska at Omaha, Omaha, NE, USA

Chinese Academy of Sciences, Beijing, China

Zhan Shi Smeal College of Business, Penn State University, University Park, PA, USA

Tung-Li Shih Department of Hospitality Management, Ming Dao University, Changhua Peetow, Taiwan

Wei K. Shih Bates White Economic Consulting, Washington, DC, USA

Andrew F. Siegel University of Washington, Seattle, WA, USA

Nicholas Sim School of Economics, University of Adelaide, Adelaide, SA, Australia

Ben J. Sopranzetti Rutgers, The State University of New Jersey, Newark, NJ, USA

Suresh Srivastava University of Alaska Anchorage, Anchorage, AK, USA

Jung-Bin Su Department of Finance, China University of Science and Technology, Nankang, Taipei, Taiwan

Edward W. Sun KEDGE Business School and BEM Management School, Bordeaux, France

Norman R. Swanson Department of Economics, Rutgers, The State University of New Jersey, New Brunswick, NJ, USA

Tzu Tai Department of Finance and Economics, Rutgers, The State University of New Jersey, Piscataway, NJ, USA

Alex P. Tang Morgan State University, Baltimore, MD, USA

Stephen J. Taylor Lancaster University Management School, Lancaster, UK

Stuart Thomas RMIT University, Melbourne, VIC, Australia

Chiung-Min Tsai Central Bank of the Republic of China (Taiwan), Taipei, Taiwan, Republic of China

Maria Vargas Universidad de Zaragoza, Zaragoza, Spain

Itzhak Venezia School of Business, The Hebrew University, Jerusalem, Israel

Bocconi University, Milan, Italy

Chia-Jane Wang Manhattan College, Riverdale, NY, USA

Cindy Shin-Huei Wang CORE, Université Catholique de Louvain and FUNDP, Academie Louvain, Louvain-la-Neuve, Belgium

Department of Quantitative Finance, National TsingHwa University, Hsinchu City, Taiwan

Kehluh Wang Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan

Shin Yun Wang National Dong Hwa University, Shou-Feng, Hualien, Taiwan

Weining Wang Ladislaus von Bortkiewicz Chair of Statistics, C.A.S.E. – Center for Applied Statistics and Economics, Humboldt-Universität zu Berlin, Berlin, Berlin, Germany

Yanzhi Wang Yuan Ze University, Taiwan

Department of Finance, College of Management, National Taiwan University, Taipei, Taiwan

Yu-Jen Wang Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan

Bruce W. Weber Lerner College of Business and Economics, University of Delaware, Newark, DE, USA

K. C. John Wei Hong Kong University of Science and Technology, Kowloon, Hong Kong

Wing-Keung Wong Department of Economics, Hong Kong Baptist University, Kowloon, Hong Kong

Po-Cheng Wu Department of Finance and Banking, Kainan University, Taoyuan, ROC, Taiwan

Yangru Wu Rutgers Business School – Newark and New Brunswick, Rutgers, The State University of New Jersey, New Brunswick, NJ, USA

Zhijie Xiao Department of Economics, Boston College, Chestnut Hill, MA, USA

Yi Meng Xie School of Business and Administration, Beijing Normal University, Beijing, China

Department of Economics, University of Southern California, Los Angeles, CA, USA

Haipeng Xing SUNY at Stony Brook, Stony Brook, NY, USA

Li Xu Washington State University, Richland, WA, USA

Chin W. Yang Clarion University of Pennsylvania, Clarion, PA, USA

National Chung Cheng University, Chia–yi, Taiwan

Li Yang University of New South Wales, Sydney, Australia

Yating Yang Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan

Shih-Kuo Yeh Department of Finance, National Chung Hsing University, Taichung 402, Taiwan, Republic of China

Hai-Chin Yu Department of International Business, Chung Yuan University, Chungli, Taiwan

Jing Rung Yu National Chi Nan University, Nantou, Taiwan

Qianni Yuan Department of Finance, Xiamen University, Xiamen, China

Kamil Yilmaz College of Administrative Sciences and Economics, Koc University, Istanbul, Turkey

Shaojun Zhang School of Accounting and Finance, Faculty of Business, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

Feng Zhao The University of Texas at Dallas, Richardson, TX, USA

Wei Zhong Wang Yanan Institute for Studies in Economics and Department of Statistics, School of Economics, Xiamen University, Xiamen, China

Chunyang Zhou Shanghai Jiaotong University, Shanghai, China

Xing Zhou Rutgers Business School – Newark and New Brunswick, Rutgers, The State University of New Jersey, New Brunswick, NJ, USA

Introduction to Financial Econometrics and Statistics

1

Cheng-Few Lee and John C. Lee

Contents

C.-F. Lee (✉)
Department of Finance and Economics, Rutgers Business School, Rutgers, The State University of New Jersey, Piscataway, NJ, USA

Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
e-mail: lee@business.rutgers.edu

J.C. Lee
Center for PBBEF Research, North Brunswick, NJ, USA
e-mail: johnleejohnlee@yahoo.com

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_1,

Abstract

The main purposes of this introduction chapter are (i) to discuss important financial econometrics and statistics which have been used in finance and accounting research and (ii) to present an overview of 98 chapters which have been included in this handbook. Sections 1.2 and 1.3 briefly review and discuss financial econometrics and statistics. Sections 1.4 and 1.5 discuss application of financial econometrics and statistics. Section 1.6 first classifies 98 chapters into 14 groups in accordance with subjects and topics. Then this section has classified the keywords from each chapter into two groups: finance and accounting topics and methodology topics. Overall, this chapter gives readers of this handbook guideline of how to apply this handbook to their research.

1.1 Introduction

Financial econometrics and statistics have become very important tools for empirical research in both finance and accounting. Econometric methods are important tools for asset-pricing, corporate finance, options, and futures, and conducting financial accounting research. Important econometric methods used in this research include: single equation multiple regression, simultaneous regression, panel data analysis, time-series analysis, spectral analysis, nonparametric analysis, semi-parametric analysis, GMM analysis, and other methods.

Portfolio theory and management research have used different statistics distributions, such as normal distribution, stable distribution, and log-normal distribution. Options and futures research have used binomial distribution, log-normal distribution, non-central Chi-square distribution, Poisson distribution, and others. Auditing research has used sampling survey techniques to determine the sampling error and non-sampling error for auditing. Risk management research has used Copula distribution and other distributions.

Section 1.1 is the introduction. Section 1.2 discusses financial econometrics. In this section, we have six subsections. These subsections include single equation regression methods, simultaneous equation models, panel data analysis, as well as alternative methods to deal with measurement error, time-series analysis, and spectral analysis. In the next section, Sect. 1.3, we discuss financial statistics. Within financial statistics, we discuss six subtopics, including statistical distributions; principle components and factor analysis; nonparametric, semi-parametric, and GMM analyses; and cluster analysis. After exploring these topics, we discuss the applications of financial econometrics and financial statistics in Sects. 1.4 and 1.5. In Sect. 1.6, we discuss the overview of all papers included in this handbook in accordance with the subject and methodologies used in the papers. Finally in Sect. 1.7, we summarize all the chapters in this handbook and add our concluding remarks.

As mentioned previously, Sect. 1.2 covers the topic of financial econometrics. We divide this section into six subsections. Within Sect. 1.2.1, we talk about single equation regression methods. We discuss some important issues related to single equation regression methods, including Heteroskedasticity, Specification Error, Measurement Error, Skewness and the Kurtosis Effect, Nonlinear Regression and Box-Cox transformation, Structural Change, the Chow Test and Moving Chow Test, Threshold Regression, Generalize Fluctuation Test, Probit and Logit Regression for Credit Risk Analysis, Poisson Regression, and Fuzzy Regression. The next subsection, Sect. 1.2.2, analyzes simultaneous equation models. Within the realm of simultaneous equation models, we discuss two-stage least squares estimation (2SLS) method, seemly unrelated regression (SUR) method, three-stage least squares estimation (3SLS) method, and disequilibrium estimation method. In Sect. 1.2.3, we study panel data analysis, in which we go over fixed effect model, random effect model, and clustering effect. The next subsection, Sect. 1.2.3, explores alternative methods to deal with measurement error. The alternative methods we look over in this section includes LISREL model, multifactor and multi-indicator (MIMIC) model, partial least square method, and grouping method. After we discuss alternative methods to deal with measurement error, we examine in Sect. 1.2.4 time-series analysis. We include in our section about time-series analysis some important models, including ARIMA, ARCH, GARCH, fractional GARCH, and combined forecasting. In Sect. 1.2.5, we look into spectral analysis.

In the following section, Sect. 1.3, we discuss financial statistics, along with four subsequent subtopics. In our first subsection, Sect. 1.3.1, we discuss some important statistical distributions. This subsection will look into the different types of distributions that are in statistics, including Binomial and Poisson distribution, normal distribution, log-normal distribution, Chi-square distribution, and non-central Chi-square distribution, Wishart distribution, symmetric and non-symmetric stable distributions, and other known distributions. Then, we talk about principal components and factor analysis in Sect. 1.3.2. In the following subsection, Sect. 1.3.3, we examine nonparametric, semi-parametric, and GMM analyses. The last subsection, Sect. 1.3.4, explores cluster analysis.

After discussing financial econometrics, we explore the applications of this topic in different types of financial and accounting field research. In Sect. 1.4, we describe these applications, including asset-pricing research, corporate finance research, financial institution research, investment and portfolio research, option pricing research, future and hedging research, mutual fund research, hedge fund research, microstructure, earnings announcements, real option research, financial accounting, managerial accounting, auditing, term structure modeling, credit risk modeling, and trading cost/transaction cost modeling.

We also discuss applications of financial statistics into different types of financial and accounting field research. Section 1.5 will include these applications in asset-pricing research, investment and portfolio research, credit risk management research, market risk research, operational risk research, option pricing research, mutual fund research, hedge fund research, value-at-risk research, and auditing.

1.2 Financial Econometrics

1.2.1 Single Equation Regression Methods

There are important issues related to single equation regression estimation method. They are (a) Heteroskedasticity, (b) Specification error, (c) Measurement error, (d) Skewness and kurtosis effect, (e) Nonlinear regression and Box-Cox transformation, (f) Structural change, (g) Chow test and moving Chow test, (h) Threshold regression, (i) Generalized fluctuation, (j) Probit and Logit regression for credit risk analysis, (k) Poisson regression, and (l) Fuzzy regression. These issues are briefly discussed as follows:

(a) Heteroskedasticity
  - White (1980) and Newvey and West (1987) are two important papers discussing how the heteroskedasticity test can be performed. The latter paper discusses heteroskedasticity when there are serial correlations.

(b) Specification error
  - Specification error occurs when there is missing variable in a regression analysis. To test the existence of specification error, we can refer to the papers by Thursby (1985), Fok et al. (1996), Cheng and Lee (1986), and Maddala et al. (1996).

(c) Measurement error
  - Management error problem is when there exists imprecise independent variable in a regression analysis. Papers by Lee and Jen (1978), Kim (1995, 1997, 2010), Miller and Modigliani (1966), and Lee and Chen (2012) have explored how measurement error methods can be applied to finance research. Lee and Chen have discussed alternative errors in variable estimation methods and their application in finance research.

(d) Skewness and kurtosis effect
  - Both skewness and kurtosis are two important measurement variables to prepare stock variation analysis. Papers by Lee (1976a), Sears and Wei (1988), and Lee and Wu (1985) discuss the skewness and kurtosis issue in asset pricing.

(e) Nonlinear regression and Box-Cox transformation
  - Nonlinear regression and Box-Cox transformation are important tools for finance, accounting, and urban economic researches. Papers by Lee (1976, 1977), Lee et al. (1990), Frecka and Lee (1983), and Liu (2006) have discussed how nonlinear regression and Box-Cox transformation techniques can be used to improve the specification of finance and accounting research. Kau and Lee (1976), and Kau et al. (1986) have explored how Box-Cox transformation can be used to conduct the empirical study of urban structure.

(f) Structural change
  - Papers by Yang (1989), Lee et al. (2011b, 2013) have discussed how the structural change model can be used to improve the empirical study of dividend policy and the issuance of new equity.

(g) Chow test and Moving Chow test
  - Chow (1960) has proposed a dummy variable approach to examine the existence of structure change for regression analysis. Zeileis et al. (2002)

have developed software programs to perform the Chow test and other structural change models which has been frequently used in finance and economic research.

(h) Threshold regression
   – Hansen (1996, 1997, 1999, 2000a, and 2000b) have explored the issue of threshold regressions and their applications in detecting structure change for regression.

(i) Generalize fluctuation test
   – Kuan and Hornik (1995) have discussed how the generalized fluctuation test can be used to perform structural change to regression.

(j) Probit and Logit regression for credit risk analysis
   – Probit and Logit regressions are frequently used in credit risk analysis. Ohlson (1980) used the accounting ratio and macroeconomic data to do credit risk analysis. Shumway (2001) has used accounting ratios and stock rate returns for credit risk analysis in terms of Probit and Logit regression techniques. Most recently, Hwang et al. (2008, 2009) and Cheng et al. (2010) have discussed Probit and Logit regression for credit risk analysis by introducing nonparametric and semi-parametric techniques into this kind of regression analysis.

(k) Poisson regression
   – Lee and Lee (2012) have discussed how the Poisson Regression can be performed, regardless of the relationship between multiple directorships, corporate ownership, and firm performance.

(l) Fuzzy regression
   – Shapiro (2005), Angrist and Lavy (1999), and Van Der Klaauw (2002) have discussed how Fuzzy Regression can be performed. This method has the potential to be used in finance accounting and research.

1.2.2 Simultaneous Equation Models

In this section, we will discuss alternative methods to deal with simultaneous equation models. There are (a) two-stage least squares estimation (2SLS) method, (b) seemly unrelated regression (SUR) method, (c) three-stage least squares estimation (3SLS) method, (d) disequilibrium estimation method, and (e) generalized method of moments.

(a) Two-stage least squares estimation (2SLS) method
   – Lee (1976a) has applied this to started market model; Miller and Modigliani (1966) have used 2SLS to study cost of capital for utility industry; Chen et al. (2007) have discuss the two-stage least squares estimation (2SLS) method for investigating corporate governance.

(b) Seemly unrelated regression (SUR) method
   – Seemly unrelated regression has frequently used in economic and financial research. Lee and Zumwalt (1981) have discussed how the seemly unrelated regression method can be applied in asset-pricing determination.

(c) Three-stage least squares estimation (3SLS) method
  – Chen et al. (2007) have discussed how the three-stage least squares estimation (3SLS) method can be applied in corporate governance research.
(d) Disequilibrium estimation method
  – Mayer (1989), Martin (1990), Quandt (1988), Amemiya (1974), and Fair and Jaffee, (1972) have discussed how alternative disequilibrium estimation method can be performed. Tsai (2005), Sealey (1979), and Lee et al. (2011a) have discussed how the disequilibrium estimation method can be applied in asset-pricing test and banking management analysis.
(e) Generalized method of moments
  – Hansen (1982) and Hamilton (1994, ▶ Chap. 14) have discussed how GMM method can be performed. Chen et al. (2007) have used the two-stage least squares estimation (2SLS), three-stage squares method, and GMM method to investigate corporate governance.

1.2.3 Panel Data Analysis

In this section, we will discuss important issues related to panel data analysis. They are (a) fixed effect model, (b) random effect model, and (c) clustering effect model.

Three well-known textbooks by Wooldridge (2010), Baltagi (2008) and Hsiao (2003) have discussed the applications of panel data in finance, economics, and accounting research. Now, we will discuss the fixed effect, random effect, and clustering effect in panel data analysis.

(a) Fixed effect model
  – Chang and Lee (1977) and Lee et al. (2011a) have discussed the role of the fixed effect model in panel data analysis of dividend research.
(b) Random effect model
  – Arellano and Bover (1995) have explored the random effect model and its role in panel data analysis. Chang and Lee (1977) have applied both fix effect and random effect model to investigating the relationshipbetween price per share, dividend per share, and retained earnings per share.
(c) Clustering effect model
  – Papers by Thompson (2011), Cameron et al. (2006), and Petersen (2009) review the clustering effect model and its impact on panel data analysis.

1.2.4 Alternative Methods to Deal with Measurement Error

In this section, we will discuss alternative methods of dealing with measurement error problems. They are (a) LISREL model, (b) multifactor and multi-indicator (MIMIC) model, and (c) partial least square method, and (d) grouping method.

(a) LISREL model
  – Papers by Titman and Wessal (1988), Chang (1999), Chang et al. (2009), Yang et al. (2010) have described the LISREL model and its way to resolve the measurement error problems of finance research.

(b) Multifactor and multi-indicator (MIMIC) model
  – Chang et al. (2009) and Wei (1984) have applied in the multifactor and multi-indicator (MIMIC) model in capital structure and asset-pricing research.

(c) Partial least square method
  – Papers by Core (2000), Ittner et al. (1997), and Lambert and Lacker (1987) have applied the partial least square method to deal with measurement error problems in accounting research.

(d) Grouping method
  – Papers by Lee (1973), Chen (2011), Lee and Chen (2013), Lee (1977b), Black et al. (1972), Blume and Friend (1973), and Fama and MacBeth (1973) analyze grouping method and its way to deal with measurement error problem in capital asset-pricing tests.

There are other errors in variable method, such as (i) Classical method, (ii) instrumental variable method, (iii) mathematical programming method, (iv) maximum likelihood method, (v) GMM method, and (vi) Bayesian Statistic Method. Lee and Chen (2012) have discussed all above-mentioned methods in details.

1.2.5 Time Series Analysis

In this section, we will discuss important models in time-series analysis. They are (a) ARIMA, (b) ARCH, (c) GARCH, (d) fractional GARCH, and (e) combined forecasting.

- Two well-known textbooks by Anderson (1994) and Hamilton (1994) have discussed the issues related to time-series analysis. We will discuss some important topics in time-series analysis in the following subsections.
- Myers (1991) discloses ARIMA's role in time-series analysis: Lien and Shrestha (2007) discuss ARCH and its impact on time-series analysis: Lien (2010) discusses GARCH and its role in time-series analysis: Leon and Vaello-Sebastia (2009) further research into GARCH and its role in time series in a model called Fractional GARCH.
- Granger and Newbold (1973), Granger and Newbold (1974), Granger and Ramanathan (1984) have theoretically developed combined forecasting methods. Lee et al. (1986) have applied combined forecasting methods to forecast market beta and accounting beta. Lee and Cummins (1998) have shown how to use the combined forecasting methods to perform cost of capital estimates.

1.2.6 Spectral Analysis

Anderson (1994), Chacko and Viceira (2003), and Heston (1993) have discussed how spectral analysis can be performed. Heston (1993) and Bakshi et al. (1997) have applied spectral analysis in the evaluation of option pricing.

1.3 Financial Statistics

1.3.1 Important Statistical Distributions

In this section, we will discuss different statistical distributions. They are: (a) Poisson distribution, (c) normal distribution, (d) log-normal distribution, (e) Chi-square distribution, (f) non-central Chi-square distribution.

Two well-known textbooks by Cox et al. (1979) and Rendleman and Barter (1979) have used binomial, normal, and lognormal distributions to develop an option pricing model. The following subsections note some famous authors that provide studies on these different statistical distributions. Black and Sholes (1973) have used lognormal distributions to derive the option pricing model. Finally, Aitchison and Brown (1973) is a well-known book to investigate lognormal distribution. Schroder (1989) has derived the option pricing model in terms of non-central Chi-square distribution.

Fama (1971) has used stable distributions to investigate the distribution of stock rate of returns. Chen and Lee (1981) have derived statistics distribution of Sharpe performance measure and found that Sharpe performance measure can be described by Wishart distribution.

1.3.2 Principle Components and Factor Analysis

Anderson's (2003) book entitled "*An Introduction to Multivariate Statistical Analysis*" has discussed principal components and factor analysis in detail. Chen and Shimerda (1981), Pinches and Mingo (1973), and Kao and Lee (2012) discuss how principal components and factor analyses can be used to do finance Lee et al. (1989) and accounting research.

1.3.3 Nonparametric and Semi-parametric Analyses

Ait-Sahalia and Lo (2000), and Hutchison et al. (1994) have discussed how nonparametric can be used in risk management and derivative securities evaluation. Hwang et al. (2010), and Hwang et al. (2007) have used semi-parametric to conduct credit risk analysis.

1.3.4 Cluster Analysis

The detailed procedures to discuss how cluster analysis can be used to find groups in data can be found in the textbook by Kaufman and Rousseeuw (1990). Brown and Goetzmann (1997) have applied cluster analysis in mutual fund research.

1.4 Applications of Financial Econometrics

In this section, we will briefly discuss how different methodologies of financial econometrics will be applied to the topics of finance and accounting.

(a) Asset-pricing Research
   – Methodologies used in asset-pricing research include (1) Heteroskedasticity, (2) Specification error, (3) Measurement error, (4) Skewness and kurtosis effect, (5) Nonlinear regression and Box-Cox transformation, (6) Structural change, (7) Two-stage least squares estimation (2SLS) method, (8) Seemly unrelated regression (SUR) method, (9) Three-stage least squares estimation (3SLS) method, (10) Disequilibrium estimation method, (11) Fixed effect model, (12) Random effect model, (13) Clustering effect model of panel data analysis, (14) Grouping method, (15) ARIMA, (16) ARCH, (17) GARCH, (18) Fractional GARCH, and (19) Wishart distribution.

(b) Corporate Finance Research:
   – Methodologies used in Corporate finance research include (1) Heteroskedasticity, (2) Specification error, (3) Measurement error, (4) Skewness and kurtosis effect, (5) Nonlinear regression and Box-Cox transformation, (6) Structural change, (7) Probit and Logit regression for credit risk analysis, (8) Poisson regression, (9) Fuzzy regression, (10) Two-stage least squares estimation (2SLS) method, (11) Seemly unrelated regression (SUR) method, (12) Three-stage least squares estimation (3SLS) method, (13) Fixed effect model, (14) Random effect model, (15) Clustering effect model of panel data analysis, and (16) GMM Analysis.

(c) Financial Institution Research
   – Methodologies used in Financial Institution research include (1) Heteroskedasticity, (2) Specification error, (3) Measurement error, (4) Skewness and kurtosis effect, (5) Nonlinear regression and Box-Cox transformation, (6) Structural change, (7) Probit and Logit regression for credit risk analysis, (8) Poisson regression, (9) Fuzzy regression, (10) Two-stage least squares estimation (2SLS) method, (11) Seemly unrelated regression (SUR) method, (12) Three-stage least squares estimation (3SLS) method, (13) Disequilibrium estimation method, (14) Fixed effect model, (15) Random effect model, (16) Clustering effect model of panel data analysis, (17) Semiparametric analysis.

(d) Investment and Portfolio Research
   – Methodologies used in investment and portfolio research include (1) Heteroskedasticity, (2) Specification error, (3) Measurement error, (4) Skewness and kurtosis effect, (5) Nonlinear regression and Box-Cox transformation, (6) Structural change, (7) Probit and Logit regression for credit risk analysis, (8) Poisson regression, and (9) Fuzzy regression.

(e) Option Pricing Research
   – Methodologies used in option pricing research include (1) ARIMA, (2) ARCH, (3) GARCH, (4) Fractional GARCH, (5) Spectral analysis, (6) Binomial distribution, (7) Poisson distribution, (8) normal distribution,

(9) log-normal distribution, (10) Chi-square distribution, (11) non-central Chi-square distribution, and (12) Nonparametric analysis.

(f) Future and Hedging Research
   – Methodologies used in future and hedging research include (1) Heteroskedasticity, (2) Specification error, (3) Measurement error, (4) Skewness and kurtosis effect, (5) Nonlinear regression and Box-Cox transformation, (6) Structural change, (7) Probit and Logit regression for credit risk analysis, (8) Poisson regression, and (9) Fuzzy regression.

(g) Mutual Fund Research
   – Methodologies used in mutual fund research include (1) Heteroskedasticity, (2) Specification error, (3) Measurement error, (4) Skewness and kurtosis effect, (5) Nonlinear regression and Box-Cox transformation, (6) Structural change, (7) Probit and Logit regression for credit risk analysis, (8) Poisson regression, (9) Fuzzy regression, and (10) Cluster analysis.

(h) Credit Risk Modeling
   – Methodologies used in credit risk modeling include (1) Heteroskedasticity, (2) Specification error, (3) Measurement error, (4) Skewness and kurtosis effect, (5) Nonlinear regression and Box-Cox transformation, (6) Structural change, (7) Two-stage least squares estimation (2SLS) method, (8) Seemly unrelated regression (SUR) method, (9) Three-stage least squares estimation (3SLS) method, (10) Disequilibrium estimation method, (11) Fixed effect model, (12) Random effect model, (13) Clustering effect model of panel data analysis, (14) ARIMA, (15) ARCH, (16) GARCH, and (17) Semiparametric analysis.

(i) Other Application
   – Financial econometrics is also important tools to conduct research in (1) Trading cost/transaction cost modeling, (2) Hedge fund research, (3) Microstructure, (4) Earnings announcement, (5) Real option research, (6) Financial accounting, (7) Managerial accounting, (8) Auditing, and (9) Term structure modeling.

1.5 Applications of Financial Statistics

Financial statistics is an important tool for research in (1) Asset-pricing research, (2) Investment and portfolio research, (3) Credit risk management research, (4) Market risk research, (5) Operational risk research, (6) Option pricing research, (7) Mutual fund research, (8) Hedge fund research, (9) Value-at-risk research, and (10) Auditing research.

1.6 Overall Discussion of Papers in this Handbook

In this section, we classify 98 papers (chapters 2–99) which have been presented in Appendix 1 in accordance with (A) Chapter titles and (B) Keywords.

(A) Chapter title classification in terms of Chapter Titles

Based on chapter titles, we classify 98 chapters into the following 14 topics:

(i) Financial Accounting (▶ Chaps. 2, 9, 10, 61, 97)
(ii) Mutual Funds (▶ Chaps. 3, 24, 25, 68, 88)
(iii) Microstructure (▶ Chaps. 4, 44, 47, 96)
(iv) Corporate Finance (▶ Chaps. 5, 21, 30, 38, 42, 46, 60, 63, 75, 79, 95)
(v) Asset Pricing (▶ Chaps. 6, 15, 22, 28, 34, 36, 39, 45, 50, 81, 85, 87, 93, 99)
(vi) Options (▶ Chaps. 7, 32, 37, 55, 65, 84, 86, 90, 98)
(vii) Portfolio Analysis (▶ Chaps. 8, 26, 35, 53, 67, 73, 80, 81, 83)
(viii) Risk Management (▶ Chaps. 11, 13, 16, 17, 23, 27, 41, 51, 54, 72, 91, 92)
(ix) International Finance (▶ Chaps. 12, 40, 43, 59, 69)
(x) Event Study (▶ Chap. 14)
(xi) Methodology (▶ Chaps. 18, 19, 20, 29, 31, 33, 46, 49, 52, 56, 57, 58, 62, 74, 76, 77, 78, 82, 89)
(xii) Banking Management (▶ Chap. 64)
(xiii) Pension Funds (▶ Chap. 66)
(xiv) Futures and Index Futures (▶ Chaps. 48, 70, 71, 94)

(B) Keywords classification

Based on the keywords in Appendix 1, we classify these keywords into two groups: (i) finance and accounting topics and (ii) methodology topics. The number behind each keyword is the chapter it is associated with.

(i) Finance and Accounting Topics

Abnormal earnings (87), Accounting earnings (87), Activity-based costing system (27), Agency costs (5, 97), Aggregation bias (43), Analyst experience (2), Analyst forecast accuracy (63), Analysts' forecast accuracy (97), Analysts' forecast bias (63, 97), Arbitrage pricing theory (APT) (6, 7, 36, 81), Asset (93), Asset allocation (45), Asset allocation fund (88), Asset pricing (34, 81), Asset return predictability (76), Asset returns (52), Asset-pricing returns (96), Asymmetric information (5), Asymmetric mean reversion (15), Asymmetric stochastic volatility (62), Asymmetric volatility response (15), Balanced scorecard (29), Bank capital (13), Bank holding companies (13), Bank risks (13), Bank stock return (6), Banks (12), Barrier option (65), Basket credit derivatives (23), Behavioral finance (55, 66, 73), Bias (57), Bias reduction (92), Bid-ask spreads (96, 99), Binomial option pricing model (37), Black-Scholes model (7, 90), Black-Sholes option pricing model (37), Board structure (42), Bond ratings (89), Bottom-up capital budgeting (75), Bounded complexity (85), Bounds (71), Brier score (72), Brokerage reputation (63), Business cycle (67), Business models (75), Business performance evaluation (29), Business value of firm, Buy-and-hold return (50), Calendar-time (50), Calendar-time portfolio approach (14), Call option (37), Capital asset-pricing model (CAPM) (6, 25, 28, 36, 81, 93), Capital budgeting (75, 29), Capital markets (25), Capital structure (5, 60), Carry trade (69), Case-Shiller home price indices (19), CEO compensation (97), CEO stock options (97), Change of measure (30), Cheapest-to-deliver bond (71), Chicago board of trade, (71), Cholesky decomposition (23), Closed-end Funds (25), Comparative financial systems (12), Composite trapezoid rule (51), Comprehensive earnings (87), Compromised solution (89), Compustat database (38), Compound sum method (46), Conditioning information (35), Constant/dynamic hedging (44), Contagious effect (11), Corner portfolio (45), Corporate earnings (9), Corporate finance (5), Corporate merger (21), Corporate ownership structure (42), Corporate policies (38), Corporation regulation (9), Correlated defaults (11), Cost of capital (93), Country funds (25), Credit rating (21), Credit rating (27), Credit risk (27, 65, 91), Credit risk index (27), Credit risk rating (16), Credit VaR (91), Creditworthiness (16), Cumulative abnormal return (50), Cumulative probability distribution (45), Currency market (58), Cyberinfrastructure (49), Daily realized volatility (40), Daily stock price (82), Debt maturity (64), Delivery options (71), Delta (45), Demand (33), Demonstration effect (17), Deterioration of bank asset quality (64), Determinants of capital structure (60), Discount cash flow model (46), Discretionary accruals (61), Discriminant power (89), Disposition effect (22), Dividends (38, 65, 79), Domestic investment companies (17), Double exponential smoothing (88), Duality (83), Dynamics (67), Earning management (61), Earnings change (10), Earnings level (10), Earnings quality (42), Earnings surprises (81), Economies of scale (21), Ederington hedging effectiveness (70), Effort allocation (97), Effort aversion (55), EGB2 distribution (80), Electricity (33), Empirical Bayes (85), Empirical corporate finance (95), Employee stock option (30), Endogeneity (38, 95), Endogeneity of variables (13), Endogenous supply (93), Equity valuation models (87), Equity value (75), European option (7), European put (5), Evaluation (34), Evaluation of funds (3), Exactly identified (93), Exceedance correlation (52), Exchange rate (43, 59), Executive compensation schemes (55), Exercise boundary (30), Expected market risk premium (15), Expected stock return (80), Expected utility (83), Experimental control (4), Experimental economics (4), Extreme events (67), Fallen angel (72), Finance panel data (24), Financial analysts (2), Financial crisis (64), Financial institutions (12), Financial leverage (75), Financial markets (12), Financial modeling (3), Financial planning and forecasting (87), Financial ratios (21), Financial returns (62), Financial service (49), Financial simulation (49), Financial statement analysis (87), Financial strength (16), Firm and time effects (24), Firm Size (9), Firm's performance score (21), Fixed operating cost (75), Flexibility hypothesis (79), Foreign exchange market (40), Foreign investment (17), Fourier inversion (84), Fourier transform (19), Free cash flow hypothesis (79), Frequentist segmentation (85), Fund management (53), Fundamental analysis (87), Fundamental asset values (73), Fundamental transform (84), Futures hedging (70), Gamma (45), Generalized (35), Generalized autoregressive conditional heteroskedasticity (51), Global investments (3), Gold (58), Green function (84), Grid and cloud computing (49), Gross return on investment (GRI) (75), Group decision making (29), Growth option (75), Growth rate (46), Hawkes process (11), Heavy-tailed data (20), Hedge ratios (98), Hedging (98), Hedging effectiveness (94), Hedging performance (98), Herding (66), Herding towards book-to-market factor (66), Herding towards momentum factor (66), Herding towards size factor (66), Herding towards the market (66), High end computing (49), High-dimensional data (77), Higher moments (80), High-frequency data (40), High-order moments (57),

Price on earnings model (10), Pricing (78), Pricing performance (98), Probability of informed trading (PIN) (96), Property (78), Property rights (12), Put option (37), Put-call parity (37), Quadratic cost (93), Quality options (71), Random number generation (49), Range (74), Rank dependent utility (83), Rating migration (72), Rational bias (2), Rational expectations (43), Real estate (78), Real sphere (59), Realized volatility (74), Recurrent event (72), Recursive (85), Reflection principle (65), Regime-switching hedging strategy (70), Registered trading firms (17), Relative value of equity (75), Research and development expense (61), Restrictions (59), Retention option (75), Return attribution (75), Return models (10), Reverse-engineering (16), Risk (83), Risk adjusted performance (3), Risk aversion (55), Risk management (41, 49, 67, 74, 80), Risk measurement (26), Risk premium (80), Risk-neutral pricing (32), Robust estimation (41), S&P 500 index (7), Sarbanes-Oxley act (63), SCAD penalty (77), Scale-by-scale decomposition (19), Seasonality (33), Semi-log (78), Sentiment (30), Shape parameter (82), Share prices (59), Share repurchases (38), Sharpe ratios (35, 53), Short run (59), Short selling (26), Short-term financing (64), Signaling hypothesis (79), Sigma (37), Smile shapes (32), Smooth transition (74), Special items (61), Speculative bubbles (73), Spot price (33), Stationarity (10), Statistical learning (77), Stochastic discount factors (25, 35), Stochastic interest rates (98), Stochastic order (83), Stochastic volatility (44, 84, 85, 92, 98), Stock market overreaction (15), Stock markets (67), Stock option (65), Stock option pricing (98), Stock price indexes (62), Stock/futures (44), Strike price (55), Structural break (31), Subjective value (30), Substantial price fluctuations (82), Sustainable growth rate, synergy (21), Synthetic utility value (68), Systematic risk (3, 79), TAIEX (45), Tail risk (67), Timberland investments (34), Time-varying risk (25), Time-varying risk aversion (40), Time-varying volatility (15), Timing options (71), Tobin's model (99), Top-down capital budgeting (75), Total risk (79), Tournament (73), Trade direction (96), Trade turnover industry (9), Transaction costs (99), Transfer pricing (29), Treasury bond futures (71), Trend extraction (18), Trust (12), Turkish economy (59), U.S. stocks (52), Ultrahigh-dimensional data (77), Unbiasedness (43), Uncertainty avoidance (12), Uncovered interest parity (69), Unexpected volatility shocks (15), Unsystematic risk (3), Utility-based hedging strategy (70), VaR-efficient frontier (45), Variability percentage adjustment (21), Visual Basic for applications (37), Volatility index (VIX) (92), Volatility (37, 80), Volatility co-persistence (44), Volatility daily effect (92), Volatility dependencies (62), Volatility feedback effect (15), Weak efficiency (43), Weak instruments (95), Wealth transfer (75), Write-downs (61), Yaari's dual utility (83), Yield curve (39), Zero-investment portfolio (50).

(ii) Methodology Topics

A mixture of Poisson distribution (98), Analyst estimation (ANOVA) (2, 28), Analytic hierarchy process (29), Analysis of variance (19), Anderson-Darling statistic (20), Anderson-Rubin statistic (95), ANST-GARCH model (asymmetric nonlinear smooth transition- GARCH model) (15), Approximately normal distribution (28), ARCH (41, 44), ARCH models (32), ARX-GARCH (autoregressive (AR) mean process with exogenous (X) variables- GARCH model) (85), Asset-pricing tests (35), Asset-pricing regression (24) Asymmetric dependence (52), Asymptotic distribution (44), Autocovariance (99), Autoregression (62), Autoregressive conditional jump intensity (82), Autoregressive model (88), Autoregressive moving average with exogenous variables (10), Autoregressive parameters (44), Bankruptcy prediction (27), Bayesian updating (2), Binomial distribution (28), Block bootstrap (56), Block granger causality (17), Bootstrap (8, 50), Bootstrap test (14), Bootstrapped critical values (24), Boundary function (31), Box-Cox (78), Bubble test (31), Change-point models (85), Clayton copula (8, 11), Cluster standard errors (24), Custering effect (79), Co-integration (76), Co-integration breakdown test (31), Combination of forecasts (88), Combinatorial optimization (16), Combined forecasting (87), Combining forecast (27), Complex logarithm (84), Conditional distribution (56), Conditional market model (50), Conditional skewness (80), Conditional value-at-risk (26, 83), Conditional variance (70), Conditional variance estimates (44), Contemporaneous jumps (85), Contingency tables (28), Contingent claim model (75), Continuous wavelet transform (19), Cook's distance (63), Copula (8, 41, 67, 74, 91), Correction method (92), Correlation (67, 73), Correlation analysis (99), CoVar (54), Covariance decomposition (72), Cox-Ingersoll-Ross (CIR) model (22, 71), Cross-sectional and time-series dependence (42), CUSUM squared test (31), Data-mining (16), Decision trees (37), Default correlation (23, 91), Dickey-Fuller test (10), Dimension reduction (77), Discrete wavelet transform (19), Discriminant analysis (89), Distribution of underlying asset (7), Double clustering (24), Downside risk model (26), Dynamic conditional correlation (52, 58, 74), Dynamic factor model (11), Dynamic random-effects models (38), Econometric methodology (38), Econometric modeling (33), Econometrics (12), Error component two-stage least squares (EC2SLS) (12), Error in variable problem (60, 96), Estimated cross-sectional standard deviations of betas (66), Event study methodology (5, 50), Ex ante probability (82), Excess kurtosis (44), Exogeneity test (95), Expectation–maximization (EM) algorithm (96), Expected return distribution (45), Explanatory power (89), Exponential trend model (88), Extended Kalman filtering (86), Factor analysis (68, 89), Factor copula (23, 91), Factor model (39, 50), Fama-French three-factor model (14), Feltham and Ohlson model (87), Filtering methods (19), Fixed effects (57, 63, 79), Forecast accuracy (2, 9), Forecast bias (2), Forecasting complexity (97), Forecasting models (27), Fourier transform method (92), Frank copula (11), GARCH (8, 40, 41, 96), GARCH hedging strategy (70), GARCH models (48, 52), GARCH-in-mean (40), GARCH-jump model (82), GARJI model (82), Gaussian copula (8), Generalized correlations (77), Generalized hyperbolic distribution (94), Generalized method of moments (13), Gibbs sampler (62), Generalized least square (GLS) (36), Generalized method of moments (GMM) (5, 25, 43, 57, 95), Generalized two-stage least squares (G2SLS) (12), Goal programming (89), Goodness-of-fit test (82, 20), Granger-causality test (76), Gumbel copula (8, 11), Hazard model (72), Heath-Jarrow-Morton model (86), Hedonic models (78), Heston (84), Heterogeneity (43), Heteroskedasticity (57), Hidden Markov models (85), Hierarchical clustering with K-Means approach (30), Hierarchical system (68), Huber estimation (66), Hyperparameter estimation (85), Hypothesis testing (53), Ibottson's RATS (50), Infinitely divisible models (48), Instrumental variable (IV) estimation (95, 57), Johnson's Skewness-adjusted *t*-test (14), Joint-normality assumption (94), Jones (1991) model (61), Jump detection (18), Jump dilution model (30), Jump process (56), Kalman filter (66), Kolmogorov-Smirnov statistic (20), Kupiec's proportion of failures test (48), Large-scale simulations (14), Latent variable (60), Least squares (78), Likelihood maximization (32), Linear filters (18), Linear trend model (88), LISREL approach (36, 60), Locally linear quantile regression (54), Logistic smooth transition regression model (69), Logit regression (27), Log-likelihood function (99), Lognormal (65), Long-horizon event study (14), Long memory process (31, 32), Loss distribution (20), Loss function (51), MAD model (26), Matching procedure (63), Mathematical optimization (55), MATLAB (90), Maximum likelihood (35, 38, 52), Maximum likelihood estimation (MLE) (36, 44, 71), Maximum sharp measure (94), Markov Chain Monte Carlo (MCMC) (62, 69, 85), Mean-variance ratio (53), Measurement error (36, 57), Method of maximum likelihood (51), Method of moments (35), A comparative study of two models SV with MCMC algorithm (62), Microsoft Excel (37), Multiple indicator multiple causes (MIMIC) (36), Minimum generalized semi-invariance (94), Minimum recording threshold (20), Minimum value of squared residuals (MSE loss function) (10), Minimum variance efficiency (35), Misspecification (44), Model formulation (38), Model selection (56, 77), Monitoring fluctuation test (31), Monte Carlo simulation (11, 23, 32, 49, 57), Moving average method (88), Moving estimates processes (79), MSE (62), Multifactor diffusion process (56), Multifactor multi-indicator approach (36), Multiple criteria and multiple constraint linear programming (29), Multiple criteria decision making (MCDM) (68), Multiple indicators and multiple causes (MIMIC) model (60), Multiple objective programming (26), Multiple regression (6, 9), Multi-resolution analysis (19), Multivariate technique (89), Multivariate threshold autoregression model (17), MV model (26), Nonlinear filters (18), Nonlinear Kalman filter (22), Nonlinear optimization (38), Non-normality (41), Nonparametric (7), Nonparametric density estimation (86), Nonparametric tests (28), Normal copula (11), Normal distribution (45), Ohlson model (87), Order flow models (4), Ordered logit (27), Ordered probit (27), Ordinary least-squares regression (63, 73), Ordinary least-squares (OLS) (90, 39, 95, 36), Orthogonal factors (6), Outlier (33), Out-of-sample forecasts (56), Panel data estimates (12, 40, 38), Panel data regressions (42, 2), Parametric approach (51), Parametric bootstrap (35), Partial adjustment (93), Partial linear model (54), Penalized least-squares (77), Prediction test (31), Principle component analysis (89, 91), Principle component factors (21), Probability density function (27), Quantile autoregression (QAR) (41), Quadratic trend model (88), Quantile dependence (67), Quantile regression (41, 54, 67), Quasi-maximum likelihood (22), Quasi-maximum likelihood estimation strategy (48), Random walk models (4), Rank regressions (63), Realized distribution (58), Rebalancing model (26), Recursive filters (85), Recursive programming (37), Reduced-form model (23, 93), Regime-switch model (69), Regression models (85, 78), Revenue surprises rotation-corrected angle (84), Ruin probability (20), Seemingly unrelated regressions (SUR) (40, 93), Semi-parametric approach (51), Semi-parametric model (54), Serial correlation (44), Shrinkage (77), Simple adjusted formula (84), Simulations (55), Simultaneous equations (60, 93, 95, 87), Single clustering (24), Single exponential smoothing (88), Skewed generalized student's t (51), Skewness (57), Specification test (56), Spectral analysis (19), Standard errors in finance panel data (30), Standardized Z (21), State-space model (39, 66), Static factor model (11), Stepwise discriminant analysis (89), Stochastic dominance (83), Stochastic frontier analysis (13), Structural change model (79), Structural equation modeling (SEM) (60), Structural VAR (17), Student's t-copula (8, 52), Student's t-distribution (62), Seemingly unrelated regression (SUR) (43), Survey forecasts (43), Survival analysis (72), SVECM models (59), Stochastic volatility (SVOL) (62), Tail dependence (52), Taylor series expansion (90), t-Copula (11), Tempered stable distribution (48), Term structure (32, 39, 71), Term structure modeling (86), Time-series analysis (18, 34, 41), Time-heterogeneity (72), Time-series and cross-sectional effects (12), Time-varying covariate (72), Time-varying dependence (8), Time-varying parameter (34), Time-varying rational expectation hypothesis (15), Trading simulations (4), Two-sector asset allocation model (45), Two-stage estimation (52), Two-stage least square (2SLS) (95), Two-way clustering method of standard errors (42), Unbounded autoregressive moving average model (88), Unconditional coverage test (51), Unconditional variance (70), Uniformly most powerful unbiased test (53), Unit root tests (10), Unit root time series (31), Unweighted GARCH (44), Value-at-risk (VAR) (45, 54, 83, 20, 26, 41, 48, 51, 76), Variable selection (77), Variance decomposition (76), Variance estimation (70), Variance reduction methods (32), Variance-gamma process (82), VG-NGARCH model (82), Visual Basic for applications (VBA) (37), Volatility forecasting (74), Volatility regime switching (15), Volatility threshold (58), Warren and Shelton model (87), Wavelet (18), Wavelet filter (19), Weighted GARCH, (44), Weighted least-squares regression (14), Wilcoxon rank test (21), Wilcoxon two-sample test (9), Wild-cluster bootstrap (24), and Winter's method (88).

1.7 Summary and Conclusion Remarks

This chapter has discussed important financial econometrics and statistics which have been used in finance and accounting research. In addition, this chapter has presented an overview of 98 chapters which have been included in this handbook. In Sect. 1.2 "Financial Econometrics," we have six subsections which are: a single equation regression methods, Simultaneous equation models, Panel data analysis, Alternative methods to deal with measurement error, Time-series analysis, and

Spectral Analysis. Section 1.3 "Financial Statistics" has four subsections: Important Statistical Distributions, Principle components and factor analysis, Nonparametric and Semi-parametric analyses, Cluster analysis review and discuss financial econometrics and statistics. In Sect. 1.4 "Applications of financial econometrics," we briefly discuss how different methodologies of financial econometrics will be applied to the topics of finance and accounting. These methods include: Asset-pricing Research, Corporate Finance Research, Financial Institution Research, Investment and Portfolio Research, Option Pricing Research, Future and Hedging Research, Mutual Fund Research, and Credit Risk Modeling. Section 1.5, "Applications of Financial Statistics," states that financial statistics is an important tool to conduct research in the areas of (1) Asset-pricing Research, (2) Investment and Portfolio Research, (3) Credit Risk Management Research, (4) Market Risk Research, (5) Operational Risk Research, (6) Option Pricing Research, (7) Mutual Fund Research, (8) Hedge Fund Research, (9) Value-at-risk Research, and (10) Auditing. Section 1.6 is an "Overall Discussion of Papers in this Handbook." It classifies 98 chapters into 14 groups in accordance to Chapter title and keywords.

Appendix 1: Brief Abstracts and Keywords for Chapters 2 to 99

Chapter 2: Experience, Information Asymmetry, and Rational Forecast Bias

This chapter uses a Bayesian model of updating forecasts in which the bias in forecast endogenously determines how the forecaster's own estimates weigh into the posterior beliefs. The model used in this chapter predicts a concave relationship between accuracy in forecast and posterior weight that is put on the forecaster's self-assessment. This chapter then uses a panel regression to test the analytical findings and find that an analyst's experience is indeed concavely related to the forecast error.

Keywords: Financial analysts, Forecast accuracy, Information asymmetry, Forecast bias, Bayesian updating, Panel regressions, Rational bias, Optional bias, Analyst estimation, Analyst experience

Chapter 3: An Overview of Modeling Dimensions for Performance Appraisal of Global Mutual Funds (Mutual Funds)

This paper examines various performance models derived by financial experts across the globe. A number of studies have been conducted to examine investment performance of mutual funds of the developed capital markets. The measure of performance of financial instruments is basically dependent on three important models derived independently by Sharpe, Jensen, and Treynor. All three models are based on the assumptions that (1) all investors are averse to risk, and are single period expected utility of terminal wealth maximizers, (2) all investors have identical decision horizons and homogeneous expectations regarding investment opportunities, (3) all investors are able to choose among portfolios solely on the basis of expected returns and variance of returns, (4) all trans-actions costs and taxes are zero, and (5) all assets are infinitely divisible. Overall, this paper has examined nine alternative mutual funds measure. The method used in this kind of research is regression analysis.

Keywords: Financial modeling, Mutual funds, Performance appraisal, Global investments, Evaluation of funds, Portfolio management, Systematic risk, Unsystematic risk, Risk adjusted performance, Prediction of price movements

Chapter 4: Simulation as a Research Tool for Market Architects

This chapter uses simulation to gain insights into trading and market structure topic by two statistical methods. The statistical methods we use include experimental design, and careful controls over experimental parameters such as the instructions given to participants. The first is discrete event simulation and the model of computer-generated trade order flow that we describe in Sect. 3. To create a realistic, but not ad hoc, market background, we use draws from a log-normal returns distribution to simulate changes in a stock's fundamental value, or P*. The model uses price-dependent Poisson distributions to generate a realistic flow of computer-generated buy and sell orders whose intensity and supply-demand balance vary over time. The order flow fluctuations depend on the difference between the current market price and the P* value. In Sect. 4, we illustrate the second method, which is experimental control to create groupings of participants in our simulations that have the same trading "assignment." The result is the ability to make valid comparisons of trader's performance in the simulations.

Keywords: Trading simulations, Market microstructure, Order flow models, Random walk models, Experimental economics, Experimental control

Chapter 5: Motivations for Issuing Putable Debt: An Empirical Analysis

This paper is the first to examine the motivations for issuing putable bonds in which the embedded put option is not contingent upon a company-related event. We find that the market favorably views the issue announcement of these bonds that we refer to as bonds with European put options or European putable bonds. This response is in contrast to the response documented by the literature to other bond issues (straight, convertible, and most studies examining poison puts), and to the response documented in the current paper to the issue announcements of poison-put bonds. Our results suggest that the market views issuing European putable bonds as helping mitigate security mispricing. Our study is an application of important statistical methods in corporate finance, namely, Event Studies and the use of General Method of Moments for cross-sectional regressions.

Keywords: Agency costs, Asymmetric information, Corporate finance, Capital structure, Event study methodology, European put, General method of moments, Management myopia, Management entrenchment, Poison put

Chapter 6: Multi Risk-Premia Model of U.S. Bank Returns: An Integration of CAPM and APT

Interest rate sensitivity of bank stock returns has been studied using an augmented CAPM: a multiple regression model with market returns and interest rate as independent variables. In this chapter, we test an asset-pricing model in which the CAPM is augmented by three orthogonal factors which are proxies for the innovations in inflation, maturity risk, and default risk. The methodologies used in this chapter are multiple regression and factor analysis.

Keywords: CAPM, APT, Bank stock return, Interest rate risk, Orthogonal factors, Multiple regression

Chapter 7: Non-parametric Bounds for European Option Prices

This chapter derives a new nonparametric lower bound and provides an alternative interpretation of Ritchken's (1985) upper bound to the price of the European option. In a series of numerical examples, our new lower bound is substantially tighter than previous lower bounds. This is prevalent especially for out-of-the-money (OTM) options where the previous lower bounds perform badly. Moreover, we present that our bounds can be derived from histograms which are completely nonparametric in an empirical study. We first construct histograms from realizations of S&P 500 index returns following Chen et al. (2006), calculate the dollar beta of the option and expected payoffs of the index and the option, and eventually obtain our bounds. We discover violations in our lower bound and show that those violations present arbitrage profits. In particular, our empirical results show that out-of-the-money calls are substantially overpriced (violate the lower bound). The methodologies used in this chapter are nonparametric, option pricing model, and histograms methods.

Keywords: Option bounds, Nonparametric, Black-Scholes model, European option, S&P 500 index, Arbitrage, Distribution of underlying asset, Lower bound, Out-of-the-money, Kernel pricing

Chapter 8: Can Time-Varying Copulas Improve Mean-Variance Portfolio?

This chapter evaluates whether constructing a portfolio using time-varying copulas yields superior returns under various weight updating strategies. Specifically, minimum-risk portfolios are constructed based on various copulas and the Pearson correlation, and a 250-day rolling window technique is adopted to derive a sequence of time-varied dependences for each dependence model. Using daily data of the G-7 countries, our empirical findings suggest that portfolios using time-varying copulas, particularly the Clayton-dependence, outperform those constructed using Pearson correlations. The above results still hold under different weight updating strategies and portfolio rebalancing frequencies. The methodologies used in this chapter are Copulas, GARCH, Student's *t*-Copula, Gumbel Copula, Clayton Copula, Time-Varying Dependence, Portfolio Optimization, and Bootstrap.

Keywords: Copulas, Time-varying dependence, Portfolio optimization, Bootstrap, Out-of-sample return, Performance evaluation, GARCH, Gaussian copula, Student's *t*-copula, Gumbel copula, Clayton copula

Chapter 9: Determinations of Corporate Earnings Forecast Accuracy: Taiwan Market Experience

This chapter examines the accuracy of the earnings forecasts by the following test methodologies. Multiple Regression Models are used to examine the effect of six factors: firm size, market volatility, trading volume turnover, corporate earnings variances, type of industry, and experience. If the two-sample groups are related, Wilcoxon Two-Sample Test will be used to determine the relative earnings forecast accuracy. Readers are well advised and referred to the chapter appendix for methodological issues such as sample selection, variable definition, regression model, and Wilcoxon tow-sample test.

Keywords: Multiple regression, Wilcoxon two-sample test, Corporate earnings, Forecast accuracy, Management earnings, Firm size, Corporation regulation, Volatility, Trade turnover, Industry

Chapter 10: Market-Based Accounting Research (MBAR) Models: A Test of ARIMAX Modeling

This study uses standard models such as earnings level and earnings changes, among others. Models that fit better to the data drawn from companies listed on the Athens Stock Exchange have been selected employing autoregressive integrated moving average with exogenous variables (ARIMAX) models. Models I (price on earnings model) "II (returns on change in earnings divided by beginning-of-period price and prior period)" V (returns on change in earnings over opening market value), VII (returns deflated by lag of 2 years on earnings over opening market value), and IX (differenced-price model) have statistically significant coefficients of explanatory variables. These models take place with backward looking information instead of forward looking information that recent literature is assessed. The methodologies used in this chapter are price on earnings model, return models, autoregressive moving average with exogenous variables (ARIMAX), minimum value of squared residuals (MSE loss function), and Dickey-Fuller test.

Keywords: Market-based accounting research (MBAR), Price on earnings model, Earnings level, Earnings change, Return models, Autoregressive moving average with exogenous variables (ARIMAX), Minimum value of squared residuals (MSE loss function), Unit root tests stationarity, Dickey-Fuller test

Chapter 11: An Assessment of Copula Functions Approach in Conjunction with Factor Model in Portfolio Credit Risk Management

This study uses a mixture of the dynamic factor model of Duffee (1999) and a contagious effect in the specification of a Hawkes process, a class of counting processes which allows intensities to depend on the timing of previous events (Hawkes 1971). Using the mixture factor- contagious-effect model, Monte Carlo simulation is performed to generate default times of two hypothesized firms. The goodness-of-fit of the joint distributions based on the most often used copula functions in literature, including the Normal, t-, Clayton, Frank, and Gumbel copula, respectively, is assessed against the simulated default times. It is demonstrated that as the contagious effect increases, the goodness-of-fit of the joint distribution functions based on copula functions decreases, which highlights the deficiency of the copula function approach.

Keywords: Static factor model, Dynamic factor model, Correlated defaults, Contagious effect, Hawkes process, Monte Carlo simulation, Normal copula, *t*-copula, Clayton copula, Frank copula, Gumbel copula

Chapter 12: Assessing Importance of Time-Series Versus Cross-Sectional Changes in Panel Data: A Study of International Variations in Ex-Ante Equity Premia and Financial Architecture

This chapter uses simultaneous equation modeling and uses Hausman test to determine whether to report fixed or random-effects estimates. We first report random-effects estimates based on the estimation procedure of Baltagi (Baltagi 1981; Baltagi and Li 1995; Baltagi and Li 1994). We consider that the error component two-stage least squares (EC2SLS) estimator of Baltagi and Li (1995) is more efficient than the generalized two-stage least squares (G2SLS) estimator of Balestra and Varadharajan-Krishnakumar (1987). For our second estimation procedure, for comparative purposes, we use the dynamic panel modeling estimates recommended by Blundell and Bond (1998). We employ the model of Blundell and Bond (1998), as these authors argue that their estimator is more appropriate than the Arellano and Bond (1991) model for smaller time periods relative to the size of the panels. We also use this two-step procedure, use as an independent variable the first lag of the dependent variable, reporting robust standard errors of Windmeijer (2005). Thus, our two different panel estimation techniques place differing emphasis on cross-sectional and time-series effects, with the Baltagi-Li estimator emphasizing cross-sectional effects and the Blundell-Bond estimator emphasizing time-series effects.

Keywords: Panel data estimates, Time-series and cross-sectional effects, Econometrics, Financial institutions, Banks, Financial markets, Comparative financial systems, Legal traditions, Uncertainty avoidance, Trust, Property rights, Error component two-stage least squares (EC2SLS), The generalized two-stage least squares (G2SLS)

Chapter 13: Does Banking Capital Reduce Risk?: An Application of Stochastic Frontier Analysis and GMM Approach

This chapter employs stochastic frontier analysis to create a new type of instrumental variable. The unrestricted frontier model determines the highest possible profitability based solely on the book value of assets employed. We develop a second frontier based on the level of bank holding company capital as well as the amount of assets. The implication of using the unrestricted model is that we are measuring the unconditional inefficiency of the banking organization. This chapter applies generalized method of moments (GMM) regression to avoid the problem caused by departure from normality. To control for the impact of size on a bank's risk-taking behavior, the book value of assets is considered in the model. The relationship between the variables specifying bank behavior and the use of equity is analyzed by GMM regression.

Keywords: Bank capital, Generalized method of moments, Stochastic frontier analysis, Bank risks, Bank holding companies, Endogeneity of variables

Chapter 14: Evaluating Long-Horizon Event Study Methodology

This chapter examines the performance of more than 20 different testing procedures that fall into two categories. First, the buy-and-hold benchmark approach uses a benchmark to measure the abnormal buy-and-hold return for every event firm, and tests the null hypothesis that the average abnormal return is zero. Second, the calendar-time portfolio approach forms a portfolio in each calendar month consisting of firms that have had an event within a certain time period prior to the month, and tests the null hypothesis that the intercept is zero in the regression of monthly portfolio returns against the factors in an asset-pricing model. This chapter also evaluates the performance of bootstrapped Johnson's skewness-adjusted t-test. This computation-intensive procedure is considered because the distribution of long-horizon abnormal returns tends to be highly skewed to the right. The bootstrapping method uses repeated random sampling to measure the significance of relevant test statistics. Due to the nature of random sampling, the resultant measurement of significance varies each time such a procedure is used. We also evaluate simple nonparametric tests, such as the Wilcoxon signed-rank test or the Fisher's sign test, which are free from random sampling variation.

Keywords: Long-horizon event study, Johnson's Skewness-adjusted *t*-test, Weighted least-squares regression, Bootstrap test, Calendar-time portfolio approach, Fama-French three-factor model, Johnson's skewness-adjusted t-statistic, Large-scale simulations

Chapter 15: Effect of Unexpected Volatility Shocks on Intertemporal Risk-Return Relation

This chapter employs the ANST-GARCH model that is capable of capturing the asymmetric volatility effect of a positive and negative return shock. The key feature of the model is the regime-shift mechanism that allows a smooth, flexible transition of the conditional volatility between different states of volatility persistence. The regime-switching mechanism is governed by a logistic transition function that changes values depending on the level of the previous return shock. With a negative (positive) return shock, the conditional variance process is described as a high (low)-persistence-in-volatility regime. The ANST-GARCH model describes the heteroskedastic return dynamics more accurately and generates better volatility forecasts.

Keywords: Intertemporal risk-return relation, Unexpected volatility shocks, Time-varying rational expectation hypothesis, Stock market overreaction, Expected market risk premium, Volatility feedback effect, Asymmetric mean reversion, Asymmetric volatility response, Time-varying volatility, Volatility regime switching, ANST-GARCH model

Chapter 16: Combinatorial Methods for Constructing Credit Risk Ratings

This chapter uses a novel method, the Logical Analysis of Data (LAD), to reverse-engineer and construct credit risk ratings which represent the creditworthiness of financial institutions and countries. LAD is a data-mining method based on combinatorics, optimization, and Boolean logic that utilizes combinatorial search techniques to discover various combinations of attribute values that are characteristic of the positive or negative character of observations. The proposed methodology is applicable in the general case of inferring an objective rating system from archival data, given that the rated objects are characterized by vectors of attributes taking numerical or ordinal values. The proposed approaches are shown to generate transparent, consistent, self-contained, and predictive credit risk rating models, closely approximating the risk ratings provided by some of the major rating agencies. The scope of applicability of the proposed method extends beyond the rating problems discussed in this study, and can be used in many other contexts where ratings are relevant. This study also uses multiple linear regression to derive the logical rating scores.

Keywords: Credit risk rating, Reverse-engineering, Logical analysis of data, Combinatorial optimization, Data-mining, Creditworthiness, Financial strength, Internal rating, Preorder, Logical rating score

Chapter 17: Dynamic Interactions in the Taiwan Stock Exchange: A Threshold VAR Models

This chapter constructs a six-variable VAR model (including NASDAQ returns, TSE returns, NT/USD returns, net foreign purchases, net domestic investment companies (dic) purchases, and net registered trading firms (rtf) purchases) to examine: (i) the interaction among three types of institutional investors, particularly to test whether net foreign purchases lead net domestic purchases by dic and rtf (the so-called demonstration effect); (ii) whether net institutional purchases lead market returns or vice versa, and (iii) whether the corresponding lead-lag relationship is positive or negative? Readers are well advised to refer to chapter appendix for detailed discussion of the unrestricted VAR model, the structural VAR model, and the threshold VAR analysis. The methodologies used in this chapter are multivariate threshold autoregression model, structural VAR, and Block Granger Causality.

Keywords: Demonstration effect, Multivariate threshold autoregression model, Foreign investment, Lead-lag relationship, Structural VAR, Block Granger causality, Institutional investors, Domestic investment companies, Registered trading firms, Qualified foreign institutional investors

Chapter 18: Methods of Denoising Financial Data

This chapter uses denoising analysis which imposes new challenges for financial data mining due to the irregularities and roughness observed in financial data, particularly, for instantaneously collected massive amounts of tick-by-tick data from financial markets for information analysis and knowledge extraction. Inefficient decomposition of the systematic pattern (the trend) and noises of financial data will lead to erroneous conclusions since irregularities and roughness of the financial data make the application of traditional methods difficult. The methodologies used in this chapter are linear filters, nonlinear filters, time-series analysis, trend extraction, and wavelet.

Keywords: Jump detection, Linear filters, Nonlinear filters, Time-series analysis, Trend extraction, Wavelet

Chapter 19: Analysis of Financial Time: Series Using Wavelet Methods

This chapter presents a set of tools, which allow gathering information about the frequency components of a time-series. In the first step, we discuss spectral analysis and filtering methods. Spectral analysis can be used to identify and to quantify the different frequency components of a data series. Filters permit to capture specific components (e.g. trends, cycles, seasonalities) of the original time-series. In the second step, we introduce wavelets, which are relatively new tools in economics and finance. They take their roots from filtering methods and Fourier analysis, but overcome most of the limitations of these two methods. Their principal advantages derive from: (i) combined information from both time-domain and frequency-domain and (ii) their flexibility as they do not make strong assumptions concerning the data generating process for the series under investigation.

Keywords: Filtering methods, Spectral analysis, Fourier transform, Wavelet filter, Continuous wavelet transform, Discrete wavelet transform, Multi-resolution analysis, Scale-by-scale decomposition, Analysis of variance, Case-Shiller home price indices

Chapter 20: Composite Goodness-of-Fit Tests for Left Truncated Loss Sample

This chapter derives the exact formulae for several goodness-of-fit statistics that should be applied to loss models with left-truncated data where the fit of a distribution in the right tail of the distribution is of central importance. We apply the proposed tests to real financial losses, using a variety of distributions fitted to operational loss and the natural catastrophe insurance claims data. The methodologies discussed in this chapter are goodness-of-fit tests, loss distribution, ruin probability, value-at-risk, Anderson-Darling statistic, Kolmogorov-Smirnov statistic.

Keywords: Goodness-of-fit tests, Left-truncated data, Minimum recording threshold, Loss distribution, Heavy-tailed data, Operational risk, Insurance, Ruin probability, Value-at-risk, Anderson-Darling statistic, Kolmogorov-Smirnov statistic

Chapter 21: Effect of Merger on the Credit Rating and Performance of Taiwan Security Firms

This chapter identifies and defines variables for merger synergy analysis followed by principal component factor analysis, variability percentage adjustment, and performance score calculation. Finally, Wilcoxon sign rank test is used for hypothesis testing. We extract principle component factors from a set of financial ratios. Percentage of variability explained and factor loadings are adjusted to get a modified average weight for each financial ratio. This weight is multiplied by the standardized Z value of the variable, and summed a set of variables get a firm's performance score. Performance scores are used to rank the firm. Statistical significance of difference in pre- and post-merger rank is tested using the Wilcoxon sign rank.

Keywords: Corporate merger, Financial ratios, Synergy, Economies of scale, Credit rating, Variability percentage adjustment, Principle component factors, Firm's performance score, Standardized Z, Wilcoxon rank test

Chapter 22: On-/Off-the-Run Yield Spread Puzzle: Evidence from Chinese Treasury Markets

This chapter uses on-/off-the-run yield spread to describe "on-/off-the-run yield spread puzzle" in Chinese treasury markets. To explain this puzzle, we introduce a latent factor in the pricing of Chinese off-the-run government bonds and use this factor to model the yield difference between Chinese on-the-run and off-the-run issues. We use the nonlinear Kalman filter approach to estimate the model. The methodologies used in this chapter are CIR model, nonlinear Kalman filter and Quasi-maximum likelihood model.

Keywords: On-/off-the-run yield spread, Liquidity, Disposition effect, CIR model, Nonlinear Kalman filter, Quasi-maximum likelihood

Chapter 23: Factor Copula for Defaultable Basket Credit Derivatives

This chapter uses a factor copula approach to evaluate basket credit derivatives with issuer default risk and demonstrate its application in a basket credit linked note (BCLN). We generate the correlated Gaussian random numbers by using the Cholesky decomposition, and then, the correlated default times can be decided by these random numbers and the reduced-form model. Finally, the fair BCLN coupon rate is obtained by the Monte Carlo simulation. We also discuss the effect of issuer default risk on BCLN. We show that the effect of issuer default risk cannot be accounted for thoroughly by considering the issuer as a new reference entity in the widely used one factor copula model, in which constant default correlation is often assumed. A different default correlation between the issuer and the reference entities affects the coupon rate greatly and must be taken into account in the pricing model.

Keywords: Factor copula, Issuer default, Default correlation, Reduced-form model, Basket credit derivatives, Cholesky decomposition, Monte Carlo simulation

Chapter 24: Panel Data Analysis and Bootstrapping: Application to China Mutual Funds

This chapter estimates double- and single-clustered standard errors by wild-cluster bootstrap procedure. To obtain the wild bootstrap samples in each cluster, we reuse the regressors (X), but modify the residuals by transforming the OLS residuals with weights which follow the popular two-point distribution suggested by Mammen (1993) and others. We then compare them with other estimates in a set of asset-pricing regressions. The comparison indicates that bootstrapped standard errors from double clustering outperform those from single clustering. They also suggest that bootstrapped critical values are preferred to standard asymptotic *t*-test critical values to avoid misleading test results.

Keywords: Asset-pricing regression, Bootstrapped critical values, Cluster standard errors, Double clustering, Firm and time effects, Finance panel data, Single clustering, Wild-cluster bootstrap

Chapter 25: Market Segmentation and Pricing of Closed-End Country Funds: An Empirical Analysis

This chapter finds that for closed-end country funds, the international CAPM can be rejected for the underlying securities (NAVs) but not for the share prices. This finding indicates that country fund share prices are determined globally, whereas the NAVs reflect both global and local prices of risk. Cross-sectional variations in the discounts or premiums for country funds are explained by the differences in the risk exposures of the share prices and the NAVs. Finally, this chapter shows that the share price and NAV returns exhibit predictable variation, and country fund premiums vary over time due to time-varying risk premiums. The chapter employs Generalized Method of Moments (GMM) to estimate stochastic discount factors and examines if the price of risk of closed-end country fund shares and NAVs is identical.

Keywords: Capital markets, Country funds, CAPM, Closed-end funds, Market segmentation, GMM, Net asset value, Stochastic discount factors, Time-varying risk, International asset pricing

Chapter 26: A Comparison of Portfolios Using Different Risk Measurements

This study uses three different risk measurements: the Mean-variance model, the Mean Absolute Deviation model, and the Downside Risk model. Meanwhile short selling is also taken into account since it is an important strategy that can bring a portfolio much closer to the efficient frontier by improving a portfolio's risk-return trade-off. Therefore, six portfolio rebalancing models, including the MV model, MAD model and the Downside Risk model, with/without short selling, are compared to determine which is the most efficient. All models simultaneously consider the criteria of return and risk measurement. Meanwhile, when short selling is allowed, models also consider minimizing the proportion of short selling. Therefore, multiple objective programming is employed to transform multiple objectives into a single objective in order to obtain a compromising solution. An example is used to perform simulation, and the results indicate that the MAD model, incorporated with a short selling model, has the highest market value and lowest risk.

Keywords: Portfolio selection, Risk measurement, Short selling, MV model, MAD model, Downside risk model, Multiple objective programming, Rebalancing model, Value-at-risk, Conditional value-at-risk

Chapter 27: Using Alternative Models and a Combining Technique in Credit Rating Forecasting: An Empirical Study

This chapter first utilizes the ordered logit and the ordered probit models. Then, we use ordered logit combining method to weight different techniques' probability measures, as described in Kamstra and Kennedy (1998) to form the combining model. The samples consist of firms in the TSE and the OTC market, and are divided into three industries for analysis. We consider financial variables, market variables as well as macroeconomic variables and estimate their parameters for out-of-sample tests. By means of Cumulative Accuracy Profile, the Receiver Operating Characteristics, and McFadden, we measure the goodness-of-fit and the accuracy of each prediction model. The performance evaluations are conducted to compare the forecasting results, and we find that combing technique does improve the predictive power.

Keywords: Bankruptcy prediction, Combining forecast, Credit rating, Credit risk, Credit risk index, Forecasting models, Logit regression, Ordered logit, Ordered probit, Probability density function

Chapter 28: Can We Use the CAPM as an Investment Strategy?: An Intuitive CAPM and Efficiency Test

The aim of this chapter is to check whether certain playing rules, based on the undervaluation concept arising from the CAPM, could be useful as investment strategies, and can therefore be used to beat the Market. If such strategies work, we will be provided with a useful tool for investors, and, otherwise, we will obtain a test whose results will be connected with the efficient Market hypothesis (EMH) and with the CAPM. The methodology used is both intuitive and rigorous: analyzing how many times we beat the Market with different strategies, in order to check whether when we beat the Market, this happens by chance.

Keywords: ANOVA, Approximately normal distribution, Binomial distribution, CAPM, Contingency tables, Market efficiency, Nonparametric tests, Performance measures

Chapter 29: Group Decision Making Tools for Managerial Accounting and Finance Applications

This chapter adopts an Analytic Hierarchy Process (AHP) approach to solve various accounting or finance problems such as developing a business performance evaluation system and developing a banking performance evaluation system. AHP uses hierarchical schema to incorporate nonfinancial and external performance measures. Our model has a broader set of measures that can examine external and nonfinancial performance as well as internal and financial performance. While AHP is one of the most popular multiple goals decision-making tools, Multiple Criteria and Multiple Constraint (MC^2) Linear Programming approach also can be used to solve group decision-making problems such as transfer pricing and capital budgeting problems. The methodologies used in this chapter are Analytic Hierarchy Process, multiple criteria and multiple constraint linear programming, and balanced scorecard and business performance evaluation.

Keywords: Analytic hierarchy process, Multiple criteria and multiple constraint linear programming, Business performance evaluation, Activity-based costing system, Group decision making, Optimal trade-offs, Balanced scorecard, Transfer pricing, Capital budgeting

Chapter 30: Statistics Methods Applied in Employee Stock Options

This study provides model-based and compensation-based approaches to price subjective value of employee stock options (ESOs). In model-based approach, we consider a utility-maximizing model that the employee allocates his wealth among the company stock, market portfolio, and risk-free bond, and then derive the ESO formulae which take into account illiquidity and sentiment effects. By using the method of change of measure, the derived formulae are simply like that of the market values with altered parameters. To calculate compensation-based subjective value, we group employees by hierarchical clustering with K-Means approach and back out the option value in an equilibrium competitive employment market. Further, we test illiquidity and sentiment effects on ESO values by running the regressions which consider the problem of standard errors in finance panel data.

Keywords: Employee stock option, Sentiment, Subjective value, Illiquidity, Change of measure, Hierarchical clustering with K-Means approach, Standard errors in finance panel data, Exercise boundary, Jump diffusion model

Chapter 31: Structural Change and Monitoring Tests

This chapter focuses on various structural change and monitoring tests for a class of widely used time-series models in economics and finance, including $I(0)$, $I(1)$, $I(d)$ processes and the co-integration relationship. In general, structural change tests can be categorized into two types: One is the classical approach to testing for structural change, which employs retrospective tests using a historical data set of a given length; the other one is the fluctuation-type test in a monitoring scheme, which means for given a history period for which a regression relationship is known to be stable, we then test whether incoming data are consistent with the previously established relationship. Several structural changes such as CUSUM squared tests, the QLR test, the prediction test, the multiple break test, bubble tests, co-integration breakdown tests, and the monitoring fluctuation test are discussed in this chapter, and we further illustrate all details and usefulness of these tests.

Keywords: Co-integration breakdown test, Structural break, Long memory process, Monitoring fluctuation test, Boundary function, CUSUM squared test, Prediction test, Bubble test, Unit root time series, Persistent change

Chapter 32: Consequences of Option Pricing of a Long Memory in Volatility

This chapter use conditionally heteroskedastic time-series models to describe the volatility of stock index returns. Volatility has a long memory property in the most general models and then the autocorrelations of volatility decay at a hyperbolic rate; contrasts are made with popular, short memory specifications whose autocorrelations decay more rapidly at a geometric rate. Options are valued for ARCH volatility models by calculating the discounted expectations of option payoffs for an appropriate risk-neutral measure. Monte Carlo methods provide the expectations. The speed and accuracy of the calculations is enhanced by two variance reduction methods, which use antithetic and control variables. The economic consequences of a long memory assumption about volatility are documented, by comparing implied volatilities for option prices obtained from short and long memory volatility processes.

Keywords: ARCH models, Implied volatility, Index options, Likelihood maximization, Long memory, Monte Carlo, Option prices, Risk-neutral pricing, Smile shapes, Term structure, Variance reduction methods

Chapter 33: Seasonal Aspects of Australian Electricity Market

This chapter develops econometric models for seasonal patterns in both price returns and proportional changes in demand for Australian electricity. Australian Electricity spot prices differ considerably from equity spot prices in that they contain an extremely rapid mean reversion process. The electricity spot price could increase to a market cap price of AU$12,500 per Megawatt Hour (MWh) and revert back to a mean level (AUD$30) within a half hour interval. This has implications for derivative pricing and risk management. We also model extreme spikes in the data. Our study identifies both seasonality effects and dramatic price reversals in the Australian electricity market. The pricing seasonality effects include time-of-day, day-of-week, monthly, and yearly effects. There is also evidence of seasonality in demand for electricity.

Keywords: Electricity, Spot price, Seasonality, Outlier, Demand, Econometric modeling

Chapter 34: Pricing Commercial Timberland Returns in the United States

This chapter uses both parametric and nonparametric approaches to evaluate private- and public-equity timberland investments in the United States. Private-equity timberland returns are proxied by the NCREIF Timberland Index, whereas public-equity timberland returns are proxied by the value-weighted returns on a dynamic portfolio of the US publicly traded forestry firms that had or have been managing timberlands. Static estimations of the capital asset-pricing model and Fama-French three-factor model are obtained by ordinary least squares, whereas dynamic estimations are obtained by state-space specifications with the Kalman filter. In estimating the stochastic discount factors, linear programming is used.

Keywords: Alternative asset class, Asset pricing, Evaluation, Fama-French three-factor model, Nonparametric analysis, State-space model, Stochastic discount factor, Timberland investments, Time series, Time-varying parameter

Chapter 35: Optimal Orthogonal Portfolios with Conditioning Information

This chapter derives and characterizes optimal orthogonal portfolios in the presence of conditioning information in the form of a set of lagged instruments. In this setting, studied by Hansen and Richard (1987), the conditioning information is used to optimize with respect to the unconditional moments. We present an empirical illustration of the properties of the optimal orthogonal portfolios. The methodology in this chapter includes regression and maximum likelihood parameter estimation, as well as method of moments estimation. We form maximum likelihood estimates of nonlinear functions as the functions evaluated at the maximum likelihood parameter estimates.

Keywords: Asset-pricing tests, Conditioning information, Minimum variance efficiency, Optimal portfolios, Predicting returns, Portfolio management, Stochastic discount factors, Generalized, Method of moments, Maximum likelihood, Parametric bootstrap, Sharpe ratios

Chapter 36: Multi-factor, Multi-indicator Approach to Asset Pricing: Method and Empirical Evidence

This chapter uses a multifactor, multi-indicator approach to test the capital asset-pricing model (CAPM) and the arbitrage pricing theory (APT). This approach is able to solve the measuring problem in the market portfolio in testing CAPM, and it is also able to directly test APT by linking the common factors to the macroeconomic indicators. We propose a MIMIC approach to test CAPM and APT. The beta estimated from the MIMIC model by allowing measurement error on the market portfolio does not significantly improve the OLS beta, while the MLE estimator does a better job than the OLS and GLS estimators in the cross-sectional regressions because the MLE estimator takes care of the measurement error in beta. Therefore, the measurement error problem on beta is more serious than that on the market portfolio.

Keywords: Capital asset-pricing model, CAPM, Arbitrage pricing theory, Multifactor multi-indicator approach, MIMIC, Measurement error, LISREL approach, Ordinary least square, OLS, General least square, GLS, Maximum likelihood estimation, MLE

Chapter 37: Binomial OPM, Black-Scholes OPM and Their Relationship: Decision Tree and Microsoft Excel Approach

This chapter will first demonstrate how Microsoft Excel can be use to create the Decision Trees for the Binomial Option Pricing Model. At the same time, this chapter will discuss the Binomial Option Pricing Model in a less mathematical fashion. All the mathematical calculations will be taken care by the Microsoft Excel program that is presented in this chapter. Finally, this chapter also uses the Decision Tree approach to demonstrate the relationship between the Binomial Option Pricing Model and the Black-Scholes Option Pricing Model.

Keywords: Binomial option pricing model, Decision trees, Black-Sholes option pricing model, Call option, Put option, Microsoft Excel, Visual Basic for applications, VBA, Put-call parity, Sigma, Volatility, Recursive programming

Chapter 38: Dividend Payments and Share Repurchases of U.S. Firms: An Econometric Approach

This chapter uses the econometric methodology to deal with the dynamic interrelationships between dividend payments and share repurchases and investigate endogeneity of certain explanatory variables. Identification of the model parameters is achieved in such models by exploiting the cross-equations restrictions on the coefficients in different time periods. Moreover, the estimation entails using nonlinear optimization methods to compute the maximum likelihood estimates of the dynamic random-effects models and for testing statistical hypotheses using likelihood ratio tests. This study also highlights the importance of developing comprehensive econometric models for these interrelationships. It is common in finance research to spell out "specific hypotheses" and conduct empirical research to investigate validity of the hypotheses.

Keywords: Compustat database, Corporate policies, Dividends, Dynamic random-effects models, Econometric methodology, Endogeneity, Maximum likelihood, Intangible assets, Model formulation, Nonlinear optimization, Panel data, Share repurchases

Chapter 39: Term Structure Modeling and Forecasting Using the Nelson-Siegel Model

In this chapter, we illustrate some recent developments in the yield curve modeling by introducing a latent factor model called the dynamic Nelson-Siegel model. This model not only provides good in-sample fit, but also produces superior out-of-sample performance. Beyond Treasury yield curve, the model can also be useful for other assets such as corporate bond and volatility. Moreover, the model also suggests generalized duration components corresponding to the level, slope, and curvature risk factors. The dynamic Nelson-Siegel model can be estimated via a one-step procedure, like the Kalman filter, which can also easily accommodate other variables of interests. Alternatively, we could estimate the model through a two-step process by fixing one parameter and estimating with ordinary least squares. The model is flexible and capable of replicating a variety of yield curve shapes: upward sloping, downward sloping, humped, and inverted humped. Forecasting the yield curve is achieved through forecasting the factors and we can impose either a univariate autoregressive structure or a vector autoregressive structure on the factors.

Keywords: Term structure, Yield curve, Factor model, Nelson-Siegel curve, State-space model

Chapter 40: The Intertemporal Relation Between Expected Return and Risk On Currency

The literature has so far focused on the risk-return trade-off in equity markets and ignored alternative risky assets. This chapter examines the presence and significance of an intertemporal relation between expected return and risk in the foreign exchange market. This chapter tests the existence and significance of a daily risk-return trade-off in the FX market based on the GARCH, realized, and range volatility estimators. Our empirical analysis relies on the maximum likelihood estimation of the GARCH-in-mean models, as described in Appendix A. We also use the seemingly unrelated (SUR) regressions and panel data estimation to investigate the significance of a time-series relation between expected return and risk on currency.

Keywords: GARCH, GARCH-in-mean, Seemingly unrelated regressions (SUR), Panel data estimation, Foreign exchange market, ICAPM, High-frequency data, Time-varying risk aversion, High-frequency data, Daily realized volatility

Chapter 41: Quantile Regression and Value-at-Risk

This chapter studies quantile regression (QR) estimation of Value-at-Risk (VaR). VaRs estimated by the QR method display some nice properties. In this chapter, different QR models in estimating VaRs are introduced. In particular, VaR estimation based on quantile regression of the QAR models, Copula models, ARCH models, GARCH models, and the CaViaR models is systematically introduced. Comparing the proposed QR method with traditional methods based on distributional assumptions, the QR method has the important property that it is robust to non-Gaussian distributions. Quantile estimation is only influenced by the local behavior of the conditional distribution of the response near the specified quantile. As a result, the estimates are not sensitive to outlier observations. Such a property is especially attractive in financial applications since many financial data like, say, portfolio returns (or log returns), are usually not normally distributed. To highlight the importance of the QR method in estimating VaR, we apply the QR techniques to estimate VaRs in International Equity Markets. Numerical evidence indicates that QR is a robust estimation method for VaR.

Keywords: ARCH, Copula, GARCH, Non-normality, QAR, Quantile regression, Risk management, Robust estimation, Time series, Value-at-risk

Chapter 42: Earnings Quality and Board Structure: Evidence from South East Asia

Using a sample of listed firms in Southeast Asia countries, this chapter examines the association among board structure and corporate ownership structure in affecting earnings quality. The econometric method employed is regressions of panel data. In a panel data setting, I address both cross-sectional and time-series dependence. Following Gow et al. (2010), I employ the two-way clustering method where the standard errors are clustered by both firm and year in my regressions of panel data.

Keywords: Earnings quality, Board structure, Corporate ownership structure, Panel data regressions, Cross-sectional and time-series dependence, Two-way clustering method of standard errors

Chapter 43: Rationality and Heterogeneity of Survey Forecasts of the Yen-Dollar Exchange Rate: A Reexamination

This chapter examines the rationality and diversity of industry-level forecasts of the yen-dollar exchange rate collected by the Japan Center for International Finance. We compare three specifications for testing rationality: the "conventional" bivariate regression, the univariate regression of a forecast error on a constant and other information set variables, and an error correction model (ECM). We extend the analysis of industry-level forecasts to a SUR-type structure using an innovative GMM technique (Bonham and Cohen 2001) that allows for forecaster cross-correlation due to the existence of common shocks and/or herd effects. Our GMM tests of micro-homogeneity uniformly reject the hypothesis that forecasters exhibit similar rationality characteristics.

Keywords: Rational expectations, Unbiasedness, Weak efficiency, Micro-homogeneity, Heterogeneity, Exchange rate, Survey forecasts, Aggregation bias, GMM, SUR

Chapter 44: Stochastic Volatility Structures and Intra-day Asset Price Dynamics

This chapter uses conditional volatility estimators as special cases of a general stochastic volatility structure. The theoretical asymptotic distribution of the measurement error process for these estimators is considered for particular features observed in intraday financial asset price processes. Specifically, I consider the effects of (i) induced serial correlation in returns processes, (ii) excess kurtosis in the underlying unconditional distribution of returns, (iii) market anomalies such as market opening and closing effects, and (iv) failure to account for intraday trading patterns. These issues are considered with applications in option pricing/trading strategies and the constant/dynamic hedging frameworks in mind. The methodologies used in this chapter are ARCH, maximum likelihood method, and unweighted GARCH.

Keywords: ARCH, Asymptotic distribution, Autoregressive parameters, Conditional variance estimates, Constant/dynamic hedging, Excess kurtosis, Index futures, Intraday returns, Market anomalies, Maximum likelihood estimates, Misspecification, Mis-specified returns, Persistence, Serial correlation, Stochastic volatility, Stock/futures, Unweighted GARCH, Volatility co-persistence

Chapter 45: Optimal Asset Allocation Under VaR Criterion: Taiwan Stock Market

This chapter examines the riskiness of the Taiwan stock market by determining the VaR from the expected return distribution generated by historical simulation. Value-at-risk (VaR) measures the worst expected loss over a given time horizon under normal market conditions at a specific level of confidence. VaR is determined by the left tail of the cumulative probability distribution of expected returns. Our result indicates the cumulative probability distribution has a fatter left tail, compared with the left tail of a normal distribution. This implies a riskier market. We also examined a two-sector asset allocation model subject to a target VaR constraint. The VaR-efficient frontier of the TAIEX traded stocks recommended, mostly, a corner portfolio.

Keywords: Value-at-risk, Asset allocation, Cumulative probability distribution, Normal distribution, VaR-efficient frontier, Historical simulation, Expected return distribution, Two-sector asset allocation model, Delta, gamma, Corner portfolio, TAIEX

Chapter 46: Alternative Methods for Estimating Firm's Growth Rate

The most common valuation model is the dividend growth model. The growth rate is found by taking the product of the retention rate and the return on equity. What is less well understood are the basic assumptions of this model. In this paper, we demonstrate that the model makes strong assumptions regarding the financing mix of the firm. In addition, we discuss several methods suggested in the literature on estimating growth rates and analyze whether these approaches are consistent with the use of using a constant discount rate to evaluate the firm's assets and equity. The literature has also suggested estimating growth rate by using the average percentage change method, compound-sum method, and/or regression methods. We demonstrate that the average percentage change is very sensitive to extreme observations. Moreover, on average, the regression method yields similar but somewhat smaller estimates of the growth rate compared to the compound-sum method. We also discussed the inferred method suggested by Gordon and Gordon (1997) to estimate the growth rate. Advantages, disadvantages, and the interrelationship among these estimation methods are also discussed in detail.

Keywords: Compound sum method, Discount cash flow model, Growth rate, Internal growth rate, Sustainable growth rate

Chapter 47: Econometric Measures of Liquidity

A security is liquid to the extent that an investor can trade significant quantities of the security quickly, at or near the current market price, and bearing low transaction costs. As such, liquidity is a multidimensional concept. In this chapter, I review several widely used econometrics or statistics-based measures that researchers have developed to capture one or more dimensions of a security's liquidity (i.e., limited dependent variable model (Lesmond et al. 1999) and autocovariance of price changes (Roll 1984)). These alternative proxies have been designed to be estimated using either low-frequency or high-frequency data, so I discuss four liquidity proxies that are estimated using low-frequency data and two proxies that require high-frequency data. Low-frequency measures permit the study of liquidity over relatively long time horizons; however, they do not reflect actual trading processes. To overcome this limitation, high-frequency liquidity proxies are often used as benchmarks to determine the best low-frequency proxy. In this chapter, I find that estimates from the effective tick measure perform best among the four low-frequency measures tested.

Keywords: Liquidity, Transaction costs, Bid-ask spread, Price impact, Percent effective spread, Market model, Limited dependent variable model, Tobin's model, Log-likelihood function, Autocovariance, Correlation analysis

Chapter 48: A Quasi-Maximum Likelihood Estimation Strategy for Value-at-Risk Forecasting: Application to Equity Index Futures Markets

The chapter uses GARCH model and quasi-maximum likelihood estimation strategy to investigate equity index futures markets. We present the first empirical evidence for the validity of the ARMA-GARCH model with tempered stable innovations to estimate 1-day-ahead value-at-risk in futures markets for the S&P 500, DAX, and Nikkei. We also provide empirical support that GARCH models based on the normal innovations appear not to be as well suited as infinitely divisible models for predicting financial crashes. In our empirical analysis, we forecast 1 % value-at-risk in both spot and futures markets using normal and tempered stable GARCH models following a quasi-maximum likelihood estimation strategy. In order to determine the accuracy of forecasting for each specific model, backtesting using Kupiec's proportion of failures test is applied.

Keywords: Infinitely divisible models, Tempered stable distribution, GARCH models, Value-at-risk, Kupiec's proportion of failures test, Quasi-maximum likelihood estimation strategy

Chapter 49: Computer Technology for Financial Service

This chapter examines the core computing competence for financial services. Securities trading is one of the few business activities where a few seconds of processing delay can cost a company big fortune. Grid and Cloud computing will be briefly described. How the underlying algorithm for financial analysis can take advantage of Grid environment is chosen and presented. One of the most popular practiced algorithms Monte Carlo Simulation is used in our cases study for option pricing and risk management. The various distributed computational platforms are carefully chosen to demonstrate the performance issue for financial services.

Keywords: Financial service, Grid and cloud computing, Monte Carlo simulation, Option pricing, Risk management, Cyberinfrastructure, Random number generation, High end computing, Financial simulation, Information technology

Chapter 50: Long-Run Stock Return and the Statistical Inference

This chapter introduces the long-run stock return methodologies and their statistical inference. The long-run stock return is usually computed by using a holding strategy more than 1 year but up to 5 years. Two categories of long-run return methods are illustrated in this chapter: the event-time approach and calendar-time approach. The event-time approach includes cumulative abnormal return, buy-and-hold abnormal return, and abnormal returns around earnings announcements. In former two methods, it is recommended to apply the empirical distribution (from the bootstrapping method) to examine the statistical inference, whereas the last one uses classical *t*-test. In addition, the benchmark selections in the long-run return literature are introduced. Moreover, the calendar-time approach contains mean monthly abnormal return, factor models, and Ibbotson's RATS, which could be tested by time-series volatility.

Keywords: Long-run stock return, Buy-and-hold return, Factor model, Event-time, Calendar-time, Cumulative abnormal return, Ibottson's RATS, Conditional market model, Bootstrap, Zero-investment portfolio

Chapter 51: Value-at-Risk Estimation via a Semi-Parametric Approach: Evidence from the Stock Markets

This study utilizes the parametric approach (GARCH-based models) and the semi-parametric approach of Hull and White (1998) (HW-based models) to estimate the Value-at-Risk (VaR) through the accuracy evaluation of accuracy for the eight stock indices in Europe and Asia stock markets. The measure of accuracy includes the unconditional coverage test by Kupiec (1995) as well as two loss functions, quadratic loss function and unexpected loss. As to the parametric approach, the parameters of generalized autoregressive conditional heteroskedasticity (GARCH) model are estimated by the method of maximum likelihood and the quantiles of asymmetric distribution like skewed generalized student's t (SGT) can be solved by composite trapezoid rule. Sequentially, the VaR is evaluated by the framework proposed by Jorion (2000). Turning to the semi-parametric approach of Hull and White (1998), before performing the traditional historical simulation, the raw return series is scaled by a volatility ratio where the volatility is estimated by the same procedure of parametric approach.

Keywords: Value-at-risk, Semi-parametric approach, Parametric approach, Generalized autoregressive conditional heteroskedasticity, Skewed generalized student's *t*, Composite trapezoid rule, Method of maximum likelihood, Unconditional coverage test, Loss function

Chapter 52: Modeling Multiple Asset Returns by a Time-Varying *t* Copula Model

This chapter illustrates a framework to model joint distributions of multiple asset returns using a time-varying Student's *t* copula model. We model marginal distributions of individual asset returns by a variant of GARCH models and then use a Student's *t* copula to connect all the margins. To build a time-varying structure for the correlation matrix of *t* copula, we employ a dynamic conditional correlation (DCC) specification. We illustrate the two-stage estimation procedures for the model and apply the model to 45 major US stocks returns selected from nine sectors. As it is quite challenging to find a copula function with very flexible parameter structure to account for difference dependence features among all pairs of random variables, our time-varying *t* copula model tends to be a good working tool to model multiple asset returns for risk management and asset allocation purposes. Our model can capture time-varying conditional correlation and some degree of tail dependence, while it also has limitations of featuring symmetric dependence and inability of generating high tail dependence when being used to model a large number of asset returns.

Keywords: Student's *t* copula, GARCH models, Asset returns, U.S. stocks, Maximum likelihood, Two-stage estimation, Tail dependence, Exceedance correlation, Dynamic conditional correlation, Asymmetric dependence

Chapter 53: Internet Bubble Examination with Mean-Variance Ratio

This chapter illustrates the superiority of the mean-variance ratio (MVR) test over the traditional SR test by applying both tests to analyze the performance of the S&P 500 index and the NASDAQ 100 index after the bursting of the Internet bubble in 2000s. This shows the superiority of the MVR test statistic in revealing short-term performance and, in turn, enables investors to make better decisions in their investments. The methodologies used in this chapter are mean-variance ratio, Sharpe ratio, hypothesis testing, and uniformly most powerful unbiased test.

Keywords: Mean-variance ratio, Sharpe ratio, Hypothesis testing, Uniformly most powerful unbiased test, Internet bubble, Fund management

Chapter 54: Quantile Regression in Risk Calibration

This chapter uses the CoVaR (Conditional VaR) framework to obtain accurate information on the interdependency of risk factors. The basic technical elements of CoVaR estimation are two levels of quantile regression: one on market risk factors; another on individual risk factor. Tests on the functional form of the two-level quantile regression reject the linearity. A flexible semi-parametric modeling framework for CoVaR is proposed. A partial linear model (PLM) is analyzed. In applying the technology to stock data covering the crisis period, the PLM outperforms in the crisis time, with the justification of the backtesting procedures. Moreover, using the data on global stock markets indices, the analysis on marginal contribution of risk (MCR) defined as the local first order derivative of the quantile curve sheds some light on the source of the global market risk.

Keywords: CoVAR, Value-at-risk, Quantile regression, Locally linear quantile regression, Partial linear model, Semi-parametric model

Chapter 55: Strike Prices of Options for Overconfident Executives

This chapter uses Monte Carlo simulation to investigate the impacts of managerial overconfidence on the optimal strike prices of executive incentive options. Although it has been shown that optimally managerial incentive options should be awarded in-the-money, in practice, most firms award them at-the-money. We show that the optimal strike prices of options granted to overconfident executive are directly related to their overconfidence level, and that this bias brings the optimal strike prices closer to the institutionally prevalent at-the-money prices. The Monte Carlo simulation procedure uses a Mathematica program to find the optimal effort by managers and the optimal (for stockholders) contract parameters. An expanded discussion of the simulations, including the choice of the functional forms and the calibration of the parameters, is provided.

Keywords: Overconfidence, Managerial effort, Incentive options, Strike price, Simulations, Behavioral finance, Executive compensation schemes, Mathematica optimization, Risk aversion, Effort aversion

Chapter 56: Density and Conditional Distribution Based Specification Analysis

This chapter uses densities and conditional distributions analysis to carry out consistent specification testing and model selection among multiple diffusion processes. In this chapter, we discuss advances to this literature introduced by Corradi and Swanson (2005), who compare the cumulative distribution (marginal or joint) implied by a hypothesized null model with corresponding empirical distributions of observed data. In particular, parametric specification tests in the spirit of the conditional Kolmogorov test of Andrews (1997) that rely on block bootstrap resampling methods in order to construct test critical values are discussed. The methodologies used in this chapter are continuous time simulation methods, single process specification testing, multiple process model selection, and multi-factor diffusion process, block bootstrap, and jump process.

Keywords: Multifactor diffusion process, Specification test, Out-of-sample forecasts, Conditional distribution, Model selection, Block bootstrap, Jump process

Chapter 57: Assessing the Performance of Estimators Dealing with Measurement Errors

This chapter describes different procedures to deal with measurement error in linear models, and assess their performance in finite samples using Monte Carlo simulations, and data on corporate investment. We consider the standard instrumental variables approach proposed by Griliches and Hausman (1986) as extended by Biorn (2000) [OLS-IV], the Arellano and Bond (1991) instrumental variable estimator, and the higher-order moment estimator proposed by Erickson and Whited (2000, 2002). Our analysis focuses on characterizing the conditions under which each of these estimators produces unbiased and efficient estimates in a standard “errors in variables” setting. In the presence of fixed effects, under heteroscedasticity, or in the absence of a very high degree of skewness in the data, the EW estimator is inefficient and returns biased estimates for mismeasured and perfectly measured regressors. In contrast to the EW estimator, IV-type estimators (OLS-IV and AB-GMM) easily handle individual effects, heteroskedastic errors, and different degrees of data skewness. The IV approach, however, requires assumptions about the autocorrelation structure of the mismeasured regressor and the measurement error. We illustrate the application of the different estimators using empirical investment models.

Keywords: Investment equations, Measurement error, Monte Carlo simulations, Instrumental variables, GMM, Bias, Fixed effects, Heteroscedasticity, Skewness, High-order moments

Chapter 58: Realized Distributions of Dynamic Conditional Correlation and Volatility Thresholds in the Crude Oil, Gold, and Dollar/Pound Currency Markets

This chapter proposes a modeling framework for the study of co-movements in price changes among crude oil, gold, and dollar/pound currencies that are conditional on volatility regimes. Methodologically, we extend the Dynamic Conditional Correlation (DCC) multivariate GARCH model to examine the volatility and correlation dynamics depending on the variances of price returns involving a threshold structure. The results indicate that the periods of market turbulence are associated with an increase in co-movements in commodity (gold and oil) prices. The results imply that gold may act as a safe haven against major currencies when investors face market turmoil.

Keywords: Dynamic conditional correlation, Volatility threshold, Realized distribution, Currency market, Gold, Oil

Chapter 59: Pre-IT Policy, Post-IT Policy, and the Real Sphere in Turkey

We estimate Two SVECM (Structural Vector Error Correction) Models for the Turkish economy based on imposing short run and Long-run restrictions that accounts for examining the behavior of the real sphere in the Pre-IT policy (before Inflation-Targeting adoption) and Post-IT policy (after Inflation-Targeting Adoption). Responses reveals that an expansionary interest policy shock leads to a decrease in price level, a fall in output, an appreciation in the exchange rate, an improvement in the share prices in the very short run for the most of Pre-IT period.

Keywords: SVECM models, Turkish economy, Short run, Long run, Restrictions, Inflation targeting, Pre-IT policy, Post-IT policy, Share prices, Exchange rate, Monetary policy shock, Output, Price level, Real sphere

Chapter 60: Determination of Capital Structure: A LISREL Model Approach

In this chapter, we employ structural equation modeling (SEM) in LISREL system to solve the measurement errors problems in the analysis of the determinants of capital structure and find the important factors consistent with capital structure theory by using date from 2002 to 2010. The purpose of this chapter is to investigate whether the influences of accounting factors on capital structure change and whether the important factors are consistent with the previous literature. The methodologies discussed in this chapter are structural equation modeling (SEM), multiple indicators and multiple causes (MIMIC) model, LISREL system, simultaneous equations, and SEM with confirmatory factor analysis (CFA) approach.

Keywords: Capital structure, Structural equation modeling (SEM), Multiple indicators and multiple causes (MIMIC) model, LISREL system, Simultaneous equations, Latent variable, Determinants of capital structure, Error in variable problem

Chapter 61: Evaluating the Effectiveness of Futures Hedging

This chapter examines the Ederington hedging effectiveness (EHE) comparisons between unconditional OLS hedge strategy and other conditional hedge strategies. It is shown that OLS hedge strategy outperforms most of the optimal conditional hedge strategies when EHE is used as the hedging effectiveness criteria. Before concluding that OLS hedge is better than the others; however, we need to understand under what circumstances the result is derived. We explain why OLS is the best hedge strategy under EHE criteria in most cases, and how most conditional hedge strategies are judged as inferior to OLS hedge strategy by an EHE comparison.

Keywords: Futures hedging, Portfolio management, Ederington hedging effectiveness, Variance estimation, Unconditional variance, Conditional variance, OLS hedging strategy, GARCH hedging strategy, Regime-switching hedging strategy, Utility-based hedging strategy

Chapter 62: Evidence on Earning Management by Integrated Oil and Gas Companies

This chapter uses Jones Model (1991) which projects the expected level of discretionary accruals and demonstrates specific test methodology for detection of earnings management in the oil and gas industry. This study utilized several parametric and nonparametric statistical methods to test for such earnings management. By comparing actuals versus projected accruals, we are able to compute the total unexpected accruals. We also correlate unexpected total accruals with several difficult to manipulate indicators that reflect company's level of activities.

Keywords: Earning management, Jones (1991) model, Discretionary accruals, Income from operations, Nonrecurring items, Special items, Research and development expense, Write-downs, Political cost, Impression management, Oil and gas industry

Chapter 63: A Comparative Study of Two Models SV with MCMC Algorithm

This chapter examines two asymmetric stochastic volatility models used to describe the volatility dependencies found in most financial returns. The first is the autoregressive stochastic volatility model with Student's t-distribution (ARSV-t), and the second is the basic Svol of JPR (1994). In order to estimate these models, our analysis is based on the Markov Chain Monte Carlo (MCMC) method. Therefore, the technique used is a Metropolishastings (Hastings 1970), and the Gibbs sampler (Casella and George 1992; Gelfand and Smith 1990; Gilks et al. 1993). The empirical results concerned on the Standard and Poor's 500 composite Index (S&P), CAC40, Nasdaq, Nikkei, and Dow-Jones stock price indexes reveal that the ARSV-t model provides a better performance than the Svol model on the Mean Squared Error (MSE) and the Maximum Likelihood function.

Keywords: Autoregression, Asymmetric stochastic volatility, MCMC, Metropolishastings, Gibbs sampler, Volatility dependencies, Student's t-distribution, SVOL, MSE, Financial returns, Stock price indexes

Chapter 64: Internal Control Material Weakness, Analysts' Accuracy and Bias, and Brokerage Reputation

This chapter uses the Ordinary Least-Squares (OLS) methodology in the main tests to examine the impact of internal control material weaknesses (ICMW hereafter) on sell side analysts. We match our ICMW firms with non-ICMWs based on industry, sales, and assets. We re-estimate the models using rank regression technique to assess the sensitivity of the results to the underlying functional form assumption made by OLS. We use Cook's distance to test the outliers.

Keywords: Internal control material weakness, Analyst forecast accuracy, Analyst forecast bias, Brokerage reputation, Sarbanes-Oxley act, Ordinary least squares regressions, Rank regressions, Fixed effects, Matching procedure, Cook's distance

Chapter 65: What Increases Banks' Vulnerability to Financial Crisis: Short-Term Financing or Illiquid Assets?

This chapter applies Logit and OLS econometric techniques to analyze the Federal Reserve Y-9C report data. We show that short-term financing is a response to the adverse economic shocks rather than a cause of the recent crisis. The likelihood of financial crisis actually stems from the illiquidity and low creditworthiness of the investment. Our results are robust to endogeneity concerns when we use a difference-in-differences (DiD) approach with the Lehman bankruptcy in 2008 proxying for an exogenous shock.

Keywords: Financial crisis, Short-term financing, Debt maturity, Liquidity risk, Deterioration of bank asset quality

Chapter 66: Accurate Formulae for Evaluating Barrier Options with Dividends Payout and the Application in Credit Risk Valuation

This chapter approximates the discrete dividend payout by a stochastic continuous dividend yield, so the post dividend stock price process can be approximated by another log-normally diffusive stock process with a stochastic continuous payout ratio up to the ex-dividend date. Accurate approximation analytical pricing formulae for barrier options are derived by repeatedly applying the reflection principle. Besides, our formulae can be applied to extend the applicability of the first passage model – a branch of structural credit risk model. The stock price falls due to the dividend payout in the option pricing problem is analog to selling the firm's asset to finance the loan repayment or dividend payout in the first passage model. Thus, our formulae can evaluate vulnerable bonds or the equity values given that the firm's future loan/dividend payments are known.

Keywords: Barrier option, Option pricing, Stock option, Dividend, Reflection principle, Lognormal, Credit risk

Chapter 67: Pension Funds: Financial Econometrics on the Herding Phenomenon in Spain and the United Kingdom

This chapter uses the estimated cross-sectional standard deviations of betas to analyze if manager's behavior enhances the existence of herding phenomena and the impact of the Spanish and UK pension funds investment on the market efficiency. We also estimate the betas with an econometric technique less applied in the financial literature: state-space models and the Kalman filter. Additionally, in order to obtain a robust estimation, we apply the Huber estimator. Finally, we apply several models and study the existence of herding toward the market, size, book-to-market, and momentum factors.

Keywords: Herding, Pension funds, State-space models, Kalman filter, Huber estimation, Imitation, Behavioral finance, Estimated cross-sectional standard deviations of betas, Herding toward the market, Herding toward size factor, Herding toward book-to-market factor, and Herding toward momentum factor

Chapter 68: Estimating the Correlation of Asset Returns: A Quantile Dependence Perspective

This chapter uses the *Copula Quantile-on-Quantile Regression* (C-QQR) approach to construct the correlation between the conditional quantiles of stock returns. This new approach of estimating correlation utilizes the idea that the condition of a stock market is related to its return performance, particularly to the conditional quantile of its return, as the lower return quantiles reflect a weak market while the upper quantiles reflect a bullish one. The C-QQR approach uses the copula to generate a regression function for modeling the dependence between the conditional quantiles of the stock returns under consideration. It is estimated using a two-step quantile regression procedure, where in principle, the first step is implemented to model the conditional quantile of one stock return, which is then related in the second step to the conditional quantile of another return.

Keywords: Stock markets, Copula, Correlation, Quantile regression, Quantile dependence, Business cycle, Dynamics, Risk management, Investment, Tail risk, Extreme events, Market uncertainties

Chapter 69: Multi-criteria Decision Making for Evaluating Mutual Funds Investment Strategies

This chapter uses the criteria measurements to evaluate investment style and investigate multiple criteria decision-making (MCDM) problem. To achieve this objective, first, we employ factor analysis to extract independent common factors from those criteria. Second, we construct the evaluation frame using hierarchical system composed of the above common factors with evaluation criteria, and then derive the relative weights with respect to the considered criteria. Third, the synthetic utility value corresponding to each investment style is aggregated by the weights with performance values. Finally, we compare with empirical data and find that the model of MCDM predicts the rate of return.

Keywords: Investment strategies, Multiple Criteria Decision Making (MCDM), Hierarchical system, Investment style, Factor analysis, Synthetic utility value, Performance values

Chapter 70: Econometric Analysis of Currency Carry Trade

This chapter investigates carry trade strategy in the currency markets whereby investors fund positions in high interest rate currencies by selling low interest rate currencies to earn the interest rate differential. In this chapter, we first provide an overview of the risk and return profile of currency carry trade; second, we introduce two popular models, the regime-switch model and the logistic smooth transition regression model, to analyze carry trade returns because the carry trade returns are highly regime dependent. Finally, an empirical example is illustrated.

Keywords: Carry trade, Uncovered interest parity, Markov chain Monte Carlo, Regime-switch model, Logistic smooth transition regression model

Chapter 71: Analytical Bounds for Treasury Bond Futures Prices

This study employs a maximum likelihood estimation technique presented by Chen and Scott (1993) to estimate the parameters for two-factor Cox-Ingersoll-Ross models of the term structure. Following the estimation, the factor values are solved for by matching the short rate with the cheapest-to-deliver bond price. Then, upper bounds and lower bounds for Treasury bond futures prices can be calculated. This study first shows that the popular preference-free, closed-form cost of carry model is an upper bound for the Treasury bond futures price. Then, the next step is to derive analytical lower bounds for the futures price under one- and two-factor Cox-Ingersoll-Ross models of the term structure.

Keywords: Treasury bond futures, Delivery options, Cox-Ingersoll-Ross models, Bounds, Maximum likelihood estimation, Term structure, Cheapest-to-deliver bond, Timing options, Quality options, Chicago board of trade

Chapter 72: Rating Dynamics of Fallen Angels and Their Speculative Grade-Rated Peers: Static Versus Dynamic Approach

This study adopts the survival analysis framework (Allison 1984) to examine issuer-heterogeneity and time-heterogeneity in the rating migrations of fallen angels (FAs) and their speculative grade-rated peers (FA peers). Cox's hazard model is considered the pre-eminent method to estimate the probability that an issuer survives in its current rating grade at any point in time t over the time horizon T. In this study, estimation is based on two Cox's hazard models, including a proportional hazard model (Cox 1972) and a dynamic hazard model. The first model employs a static estimation approach and time-independent covariates, whereas the second uses a dynamic estimation approach and time-dependent covariates. To allow for any dependence among rating states of the same issuer, the marginal event-specific method (Wei et al. 1989) was used to obtain robust variance estimates. For validation purpose, the Brier score (Brier 1950) and its covariance decomposition (Yates 1982) were applied to assess the forecast performance of estimated models in forming time-varying survival probability estimates for issuers out-of-sample.

Keywords: Survival analysis, Hazard model, Time-varying covariate, Recurrent event, Brier score, Covariance decomposition, Rating migration, Fallen angel, Markov property, Issuer-heterogeneity, Time-heterogeneity

Chapter 73: Creation and Control of Bubbles: Managers Compensation Schemes, Risk Aversion, and Wealth and Short Sale Constraints

This chapter takes an alternative approach of inquiry – that of using laboratory experiments – to study the creation and control of speculative bubbles. The following three factors are chosen for analysis: the compensation scheme of portfolio managers, wealth and supply constraints, and the relative risk aversion of traders. Under a short investment horizon induced by a tournament compensation scheme, speculative bubbles are observed in markets of speculative traders and in mixed markets of conservative and speculative traders. The primary method of analysis is to use live subjects in a laboratory setting to generate original trading data, which are compared to their fundamental values. Standard statistical techniques are used to supplement analysis in explaining the divergence of asset prices from their fundamental values.

Keywords: Speculative bubbles, Laboratory experimental asset markets, Fundamental asset values, Tournament, Market efficiency, Behavioral finance, Ordinary least squares regression, Correlation

Chapter 74: Range Volatility: A Review of Models and Empirical Studies

In this chapter, we survey the significant development of range-based volatility models, beginning with the simple random walk model up to the conditional autoregressive range (CARR) model. For the extension to range-based multivariate volatilities, some approaches developed recently are adopted, such as the dynamic conditional correlation (DCC) model, the double smooth transition conditional correlation (DSTCC) GARCH model, and the copula method. At last, we introduce different approaches to build bias-adjusted realized range to obtain a more efficient estimator.

Keywords: Range, Volatility forecasting, Dynamic conditional correlation, Smooth transition, Copula, Realized volatility, Risk management

Chapter 75: Business Models: Applications to Capital Budgeting, Equity Value, and Return Attribution

This chapter describes a business model in a contingent claim modeling framework. The chapter then provides three applications of the business model. Firstly, the chapter determines the optimal capital budgeting decision in the presence of fixed operating costs, and shows how the fixed operating cost should be accounted by in an NPV calculation. Secondly, the chapter determines the values of equity value, the growth option, the retention option as the building blocks of primitive firm value. Using a sample of firms, the chapter illustrates a method in comparing the equity values of firms in the same business sector. Thirdly, the chapter relates the change in revenue to the change in equity value, showing how the combined operating leverage and financial leverage may affect the firm valuation and risks.

Keywords: Bottom-up capital budgeting, Business model, Capital budgeting, Contingent claim model, Equity value, Financial leverage, Fixed operating cost, Gross return on investment (GRI), Growth option, Market performance measure, NPV, Operating leverage, Relative value of equity, Retention option, Return attribution, Top-down capital budgeting, Wealth transfer

Chapter 76: VAR Models: Estimation, Inferences, and Applications

This chapter provides a brief overview of the basic Vector autoregression (VAR) approach by focusing on model estimation and statistical inferences. VAR models have been used extensive in finance and economic analysis. Applications of VAR models in some finance areas are discussed, including asset pricing, international finance, and market microstructure. It is shown that such approach provides a powerful tool to study financial market efficiency, stock return predictability, exchange rate dynamics, and information content of stock trades and market quality.

Keywords: VAR, Granger-causality test, Impulse response, Variance decomposition, Co-integration, Asset return predictability, Market quality, Information content of trades, Informational efficiency

Chapter 77: Model Selection for High-Dimensional Problems

This chapter introduces penalized least squares, which seek to keep important predictors in a model, while penalizing coefficients associated with irrelevant predictors. As such, under certain conditions, penalized least squares can lead to a sparse solution for linear models and achieve asymptotic consistency in separating relevant variables from irrelevant ones. We then review independence screening, a recently developed method for analyzing ultrahigh-dimensional data where the number of variables or parameters can be exponentially larger than the sample size. Independence screening selects relevant variables based on certain measures of marginal correlations between candidate variables and the response. Finally, we discuss and advocate multistage procedures that combine independence screening and variable selection and that may be especially suitable for analyzing high-frequency financial data.

Keywords: Model selection, Variable selection, Dimension reduction, Independence screening, High-dimensional data, Ultrahigh-dimensional data, Generalized correlations, Penalized least squares, Shrinkage, Statistical learning, SCAD penalty, Oracle property

Chapter 78: Hedonic Regression Models

The chapter examines three specific, different hedonic specifications: the linear, semi-log, and Box-Cox transformed hedonic models and applies them to real estate data. It also discusses recent innovations related to hedonic models and how these models are being used in contemporary studies. This provides a basic overview of the nature and variety of hedonic empirical pricing models that are employed in the economics literature. It explores the history of hedonic modeling and summarizes the field's utility-theory-based, microeconomic foundations. It also provides a discussion of and potential solutions for common problems associated with hedonic modeling.

Keywords: Hedonic models, Regression, Real estate, Box-Cox, Pricing, Price indexes, Semi-log, Least squares, Housing, Property

Chapter 79: Optimal Payout Ratio Under Uncertainty and the Flexibility Hypothesis: Theory and Empirical Evidence

We theoretically extend the proposition of DeAngelo and DeAngelo's (2006) optimal payout policy in terms of the flexibility dividend hypothesis. We also introduce growth rate, systematic risk, and total risk variables into the theoretical model. We use a panel data collected in the USA from 1969 to 2009 to empirically investigate the impact of growth rate, systematic risk, and total risk on the optimal payout ratio in terms of the fixed-effects model. Furthermore, we implement the moving estimates process to find the empirical breakpoint of the structural change for the relationship between the payout ratio and risks and confirm that the empirical breakpoint is not different from our theoretical breakpoint. Our theoretical model and empirical results can therefore be used to identify whether flexibility or the free cash flow hypothesis should be used to determine the dividend policy.

Keywords: Dividends, Payout policy, Optimal payout ratio, Flexibility hypothesis, Free cash flow hypothesis, Signaling hypothesis, Fixed effect, Clustering effect, Structural change model, Moving estimates processes, Systematic risk, Total risk, Market perfection

Chapter 80: Modeling Asset Returns with Skewness, Kurtosis, and Outliers

This chapter uses an exponential generalized beta distribution of the second kind (EGB2) to model the returns on 30 Dow-Jones industrial stocks. The model accounts for stock return characteristics, including fat tails, peakedness (leptokurtosis), skewness, clustered conditional variance, and leverage effect. The goodness-of-fit statistic provides supporting evidence in favor of EGB2 distribution in modeling stock returns. The EGB2 distribution used in this chapter is a four parameter distribution. It has a closed-form density function, and its higher-order moments are finite and explicitly expressed by its parameters. The EGB2 distribution nests many widely used distributions such as normal distribution, log-normal distribution, Weibull distribution, and standard logistic distribution.

Keywords: Expected stock return, Higher moments, EGB2 distribution, Risk management, Volatility, Conditional skewness, Risk premium

Chapter 81: Does Revenue Momentum Drive or Ride Earnings or Price Momentum?

This chapter performs dominance test to show that revenue surprises, earnings surprises, and prior returns, each lead to significant momentum returns that cannot be fully explained by the others, suggesting that each convey some exclusive and unpriced information content. Also, the joint implications of revenue surprises, earnings surprises, and prior returns are underestimated by investors, particularly when information variables point in the same direction. Momentum cross-contingencies are observed in that momentum profits driven by firm fundamental information positively depend on the accompanying firm market information, and vice versa. A three-way combined momentum strategy may offer monthly return as high as 1.44%.

Keywords: Earnings surprises, Momentum strategies, Post-earnings-announcement drift, Revenue surprises

Chapter 82: A VG-NGARCH Model for Impacts of Extreme Events on Stock Returns

This chapter compares two types of GARCH models, namely, the VG-NGARCH and the GARCH-jump model with autoregressive conditional jump intensity, i.e., the GARJI model, to make inferences on the log of stock returns when there are irregular substantial price fluctuations. The VG-NGARCH model imposes a nonlinear asymmetric structure on the conditional shape parameters in a variance-gamma process, which describes the arrival rates for news with different degrees of influence on price movements, and provides an ex ante probability for the occurrence of large price movements. On the other hand, the GARJI model, a mixed GARCH-jump model proposed by Chan and Maheu (2002), adopts two independent autoregressive processes to model the variances corresponding to moderate and large price movements, respectively.

Keywords: VG-NGARCH model, GARCH-jump model, Autoregressive conditional jump intensity, GARJI model, Substantial price fluctuations, Shape parameter, Variance-gamma process, Ex ante probability, Daily stock price, Goodness-of-fit

Chapter 83: Risk-Averse Portfolio Optimization via Stochastic Dominance Constraints

This chapter presents a new approach to portfolio selection based on stochastic dominance. The portfolio return rate in the new model is required to stochastically dominate a random benchmark. We formulate optimality conditions and duality relations for these models and construct equivalent optimization models with utility functions. Two different formulations of the stochastic dominance constraint, primal and inverse, lead to two dual problems which involve von Neuman–Morgenstern utility functions for the primal formulation and rank dependent (or dual) utility functions for the inverse formulation. We also discuss the relations of our approach to value-at-risk and conditional value-at-risk.

Keywords: Portfolio optimization, Stochastic dominance, Stochastic order, Risk, Expected utility, Duality, Rank dependent utility, Yaari's dual utility, Value-at-risk, Conditional value-at-risk

Chapter 84: Implementation Problems and Solutions in Stochastic Volatility Models of the Heston Type

This chapter compares three major approaches to solve the numerical instability problem inherent in the fundamental solution of the Heston model. In this chapter, we used the fundamental transform method proposed by Lewis to reduce the number of variables from two to one and separate the payoff function from the calculation of the Green function for option pricing. We show that the simple adjusted-formula method is much simpler than the rotation-corrected angle method of Kahl and Jäckel and also greatly superior to the direct integration method of Shaw if taking computing time into consideration.

Keywords: Heston, Stochastic volatility, Fourier inversion, Fundamental transform, Complex logarithm, Rotation-corrected angle, Simple adjusted formula, Green function

Chapter 85: Stochastic Change-Point Models of Asset Returns and Their Volatilities

This chapter considers two time-scales and uses the "short" time-scale to define GARCH dynamics and the "long" time-scale to incorporate parameter jumps. This leads to a Bayesian change-point ARX-GARCH model, whose unknown parameters may undergo occasional changes at unspecified times and can be estimated by explicit recursive formulas when the hyperparameters of the Bayesian model are specified. Efficient estimators of the hyperparameters of the Bayesian model can be developed. The empirical Bayes approach can be applied to the frequentist problem of partitioning the time series into segments under sparsity assumptions on the change-points.

Keywords: ARX-GARCH, Bounded complexity, Contemporaneous jumps, Change-point models, Empirical Bayes, Frequentist segmentation, Hidden Markov models, Hyperparameter estimation, Markov chain Monte Carlo, Recursive filters, Regression models, Stochastic volatility

Chapter 86: Unspanned Stochastic Volatilities and Interest Rate Derivatives Pricing

This chapter first reviews the recent literature on the Unspanned Stochastic Volatilities (USV) documented in the interest rate derivatives markets. The USV refers to the volatilities factors implied in the interest rate derivatives prices that have little correlation with the yield curve factors. We then present the result in Li and Zhao (2006) that a sophisticated DTSM without USV feature can have serious difficulties in hedging caps and cap straddles, even though they capture bond yields well. Furthermore, at-the-money straddle hedging errors are highly correlated with cap-implied volatilities and can explain a large fraction of hedging errors of all caps and straddles across moneyness and maturities. We also present a multifactor term structure model with stochastic volatility and jumps that yields a closed-form formula for cap prices from Jarrow et al. (2007). The three-factor stochastic volatility model with Poisson jumps can price interest rate caps well across moneyness and maturity. The econometric methods in this chapter include extended Kalman filtering, maximum likelihood estimation with latent variables, local polynomial method, and nonparametric density estimation.

Keywords: Term structure modeling, Interest rate volatility, Heath-Jarrow-Morton model, Nonparametric density estimation, Extended Kalman filtering

Chapter 87: Alternative Equity Valuation Models

This chapter examines alternative equity valuation models and their ability to forecast future stock prices. We use simultaneous equations estimation technique to investigate the stock price forecast ability of Ohlson's model, Feltham and Ohlson's Model, and Warren and Shelton's (1971) model. Moreover, we use the combined forecasting methods proposed by Granger and Newbold (1973) and Granger and Ramanathan (1984) to form combined stock price forecasts from individual models. Finally, we examine whether comprehensive earnings can provide incremental price-relevant information beyond net income.

Keywords: Ohlson model, Feltham and Ohlson model, Warren and Shelton model, Equity valuation models, Simultaneous equations estimation, Fundamental analysis, Financial statement analysis, Financial planning and forecasting, Combined forecasting, Comprehensive earnings, Abnormal earnings, Operating earnings, Accounting earnings

Chapter 88: Time Series Models to Predict the Net Asset Value (NAV) of an Asset Allocation Mutual Fund VWELX

This research examines the use of various forms of time-series models to predict the total net asset value (NAV) of an asset allocation mutual fund. The first set of model structures included simple exponential smoothing, double exponential smoothing, and the Winter's method of smoothing. The second set of predictive models used represented trend models. They were developed using regression estimation. They included linear trend model, quadratic trend model, and an exponential model. The third type of method used was a moving average method. The fourth set of models incorporated the Box-Jenkins method, including an autoregressive model, a moving average model, and an unbounded autoregressive and moving average method.

Keywords: NAV of a mutual fund, Asset allocation fund, Combination of forecasts, Single exponential smoothing, Double exponential smoothing, Winter's method, Linear trend model, Quadratic trend model, Exponential trend model, Moving average method, Autoregressive model, Moving average model, Unbounded autoregressive moving average model

Chapter 89: Discriminant Analysis and Factor Analysis: Theory and Method

This chapter discusses three multivariate techniques in detail: discriminant analysis, factor analysis, and principal component analysis. In addition, the stepwise discriminant analysis by Pinches and Mingo (1973) is improved using a goal programming technique. These methodologies are applied to determine useful financial ratios and the subsequent bond ratings. The analysis shows that the stepwise discriminant analysis fails to be an efficient solution as the hybrid approach using the goal programming technique outperforms it, which is a compromised solution for the maximization of the two objectives, namely, the maximization of the explanatory power and the maximization of discriminant power.

Keywords: Multivariate technique, Discriminant analysis, Factor analysis, Principle component analysis, Stepwise discriminant analysis, Goal programming, Bond ratings, Compromised solution, Explanatory power, Discriminant power

Chapter 90: Implied Volatility: Theory and Empirical Method

This chapter reviews the different theoretical methods used to estimate implied standard deviation and to show how the implied volatility can be estimated in empirical work. The OLS method for estimating implied standard deviation is first introduced and the formulas derived by applying a Taylor series expansion method to Black-Scholes option pricing model are also described. Three approaches of estimating implied volatility are derived from one, two, and three options, respectively. Because of these formulas with the remainder terms, the accuracy of these formulas depends on how an underlying asset is close to the present value of exercise price in an option. The formula utilizing three options for estimating implied volatility is more accurate rather than other two approaches. In this chapter, we use call options on S&P 500 index futures in 2010 and 2011 to illustrate how MATLAB can be used to deal with the issue of convergence in estimating implied volatility of future options.

Keywords: Implied volatility, Implied standard deviation (ISD), Option pricing model, MATLAB, Taylor series expansion, Ordinary least-squares (OLS), Black-Scholes Model, Options on S&P 500 index futures

Chapter 91: Measuring Credit Risk in a Factor Copula Model

This chapter uses a new approach to estimate future credit risk on target portfolio based on the framework of CreditMetricsTM by J.P. Morgan. However, we adopt the perspective of factor copula and then bring the principal component analysis concept into factor structure to construct a more appropriate dependence structure among credits. In order to examine the proposed method, we use real market data instead of a virtual one. We also develop a tool for risk analysis which is convenient to use, especially for banking loan businesses. The results show the fact that people assume dependence structures are normally distributed will indeed lead to risks underestimate. On the other hand, our proposed method captures better features of risks and shows the fat-tail effects conspicuously even though assuming the factors are normally distributed.

Keywords: Credit risk, Credit VaR, Default correlation, Copula, Factor copula, Principal component analysis

Chapter 92: Instantaneous Volatility Estimation by Nonparametric Fourier Transform Methods

This chapter conducts some simulation tests to justify the effectiveness of the Fourier transform method. Malliavin and Mancino (2009) proposed a nonparametric Fourier transform method to estimate the instantaneous volatility under the assumption that the underlying asset price process is a semi-martingale. Two correction schemes are proposed to improve the accuracy of volatility estimation. By means of these Fourier transform methods, some documented phenomena such as volatility daily effect and multiple risk factors of volatility can be observed. Then, a linear hypothesis between the instantaneous volatility and VIX derived from Zhang and Zhu (2006) is investigated.

Keywords: Information content, Instantaneous volatility, Fourier transform method, Bias reduction, Correction method, Local volatility, Stochastic volatility, VIX, Volatility daily effect, Online estimation

Chapter 93: A Dynamic CAPM with Supply Effect: Theory and Empirical Results

This chapter first theoretically extends Black's CAPM, and then uses price, dividend per share, and earnings per share to test the existence of supply effect with US equity data. A simultaneous equation system is constructed through a standard structural form of a multi-period equation to represent the dynamic relationship between supply and demand for capital assets. The equation system is exactly identified under our specification. Then, two hypotheses related to supply effect are tested regarding the parameters in the reduced-form system. The equation system is estimated by the Seemingly Unrelated Regression (SUR) method, since SUR allows one to estimate the presented system simultaneously while accounting for the correlated errors.

Keywords: CAPM, Asset, Endogenous supply, Simultaneous equations

Chapter 94: A Generalized Model for Optimum Futures Hedge Ratio

This chapter proposes the generalized hyperbolic distribution as the joint log-return distribution of the spot and futures. Using the parameters in this distribution, we derive several most widely used optimal hedge ratios: minimum variance, maximum Sharpe measure, and minimum generalized semivariance. To estimate these optimal hedge ratios, we first write down the log-likelihood functions for symmetric hyperbolic distributions. Then, we estimate these parameters by maximizing the log-likelihood functions. Using these MLE parameters for the generalized hyperbolic distributions, we obtain the minimum variance hedge ratio and the optimal Sharp hedge ratio. Also based on the MLE parameters and the numerical method, we can calculate the minimum generalized semivariance hedge ratio.

Keywords: Optimal hedge ratio, Generalized hyperbolic distribution, Martingale property, Minimum variance hedge ratio, Minimum generalized semi-invariance, Maximum Sharp measure, Joint-normality assumption, Hedging effectiveness

Chapter 95: Instrument Variable Approach to Correct for Endogeneity in Finance

This chapter reviews the instrumental variables (IV) approach to endogeneity from the point of view of a finance researcher who is implementing instrumental variable methods in empirical studies. This chapter is organized into two parts. Part I discusses the general procedure of the instrumental variable approach, including Two-Stage Least Square (2SLS) and Generalized Method of Moments (GMM), the related diagnostic statistics for assessing the validity of instruments, which are important but not used very often in finance applications, and some recent advances in econometrics research on weak instruments. Part II surveys corporate finance applications of instrumental variables. We found that the instrumental variables used in finance studies are usually chosen arbitrarily, and very few diagnostic statistics are performed to assess the adequacy of IV estimation. The resulting IV estimates thus are questionable.

Keywords: Endogeneity, OLS, Instrumental variable (IV) estimation, Simultaneous equations, 2SLS, GMM, Overidentifying restrictions, Exogeneity test, Weak instruments, Anderson-Rubin statistic, Empirical corporate finance

Chapter 96: Application of Poisson Mixtures in the Estimation of Probability of Informed Trading

This research first discusses the evolution of probability of informed trading in the finance literature. Motivated by asymmetric effects, e.g., return and trading volume in up and down markets, this study modifies a mixture of the Poisson distribution model by different arrival rates of informed buys and sells to measure the probability of informed trading proposed by Easley et al. (1996). By applying the expectation–maximization (EM) algorithm to estimate the parameters of the model, we derive a set of equations for maximum likelihood estimation and these equations are encoded in a SAS Macro utilizing SAS/IML for implementation of the methodology.

Keywords: Probability of informed trading (*PIN*), Expectation–maximization (EM) algorithm, A mixture of Poisson distribution, Asset-pricing returns, Order imbalance, Information asymmetry, Bid-ask spreads, Market microstructure, Trade direction, Errors in variables

Chapter 97: CEO Stock Options and Analysts' Forecast Accuracy and Bias

This chapter uses ordinary least squares estimation to investigate the relations between CEO stock options and analysts' earnings forecast accuracy and bias. Our OLS models relate forecast accuracy and forecast bias (the dependent variables) to CEO stock options (the independent variable) and controls for earnings characteristics, firm characteristics, and forecast characteristics. In addition, the models include controls for industry and year. We use four measures of options: new options, existing exercisable options, existing unexercisable options, and total options (sum of the previous three), all scaled by total number of shares outstanding, and estimate two models for each dependent variable, one including total options and the other including new options, existing exercisable options, and existing unexercisable options. We also use both contemporaneous as well as lagged values of options in our main tests.

Keywords: CEO stock options, Analysts' forecast accuracy, Analysts' forecast bias, CEO compensation, Agency costs, Investment risk taking, Effort allocation, Opportunistic earnings management, Opportunistic disclosure management, Forecasting complexity

Chapter 98: Option Pricing and Hedging Performance Under Stochastic Volatility and Stochastic Interest Rates

This chapter fills this gap by first developing an implementable option model in closed form that admits both stochastic volatility and stochastic interest rates and that is parsimonious in the number of parameters. Based on the model, both delta-neutral and single-instrument minimum variance hedging strategies are derived analytically. Using S&P 500 option prices, we then compare the pricing and hedging performance of this model with that of three existing ones that, respectively, allow for (i) constant volatility and constant interest rates (the Black-Scholes), (ii) constant volatility but stochastic interest rates, and (iii) stochastic volatility but constant interest rates. Overall, incorporating stochastic volatility and stochastic interest rates produces the best performance in pricing and hedging, with the remaining pricing and hedging errors no longer systematically related to contract features. The second performer in the horse-race is the stochastic volatility model, followed by the stochastic interest rates model and then by the Black-Scholes.

Keywords: Stock option pricing, Stochastic volatility, Stochastic interest rates, Hedge ratios, Hedging, Pricing performance, and Hedging performance

Chapter 99: The Le Châtelier Principle of the Capital Market Equilibrium

This chapter purports to provide a theoretical underpinning for the problem of the Investment Company Act. The theory of the Le Chatelier Principle is well known in thermodynamics: The system tends to adjust itself to a new equilibrium as far as possible. In capital market equilibrium, added constraints on portfolio investment in each stock can lead to inefficiency manifested in the right-shifting efficiency frontier. According to the empirical study, the potential loss can amount to millions of dollars coupled with a higher risk-free rate and greater transaction and information costs.

Keywords: Markowitz model, Efficient frontiers, With constraints, Without constraints, Le Chatelier principle, Thermodynamics, Capital market equilibrium, Diversified mutual funds, Quadratic programming, Investment company act

References

Aitchison, J., & Brown, J. A. C. (1973). *Lognormal distribution, with special reference to its uses in economics*. London: Cambridge University Press.

Ait-Sahalia, Y., & Lo, A. W. (2000). Nonparametric risk management and implied risk aversion. *Journal of Econometrics, 94*, 9–51.

Amemiya, T. (1974). A note on a fair and Jaffee model. *Econometrica, 42*, 759–762.

Anderson, T. W. (1994). *The statistical analysis of time series*. New York: Wiley-Interscience.

Anderson, T. W. (2003). *An introduction to multivariate statistical analysis*. New York: Wiley-Interscience.

Angrist, J. D., & Lavy, V. (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics, 14*(2), 533–575.

Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics, 68*(1), 29–51.

Bakshi, G., Cao, C., & Chen, Z. (1997). Empirical performance of alternative option pricing models. *Journal of Finance, 52*(5), 2003–2049.

Baltagi, B. (2008). *Econometric analysis of panel data* (4th ed.). Chichester: Wiley.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy, 81*(3), 637–654.

Black, F., Jensen, M. C., & Scholes, M. (1972). The capital asset pricing model: Some empirical tests. In M. C. Jensen (Ed.), *Studies in the theory of capital markets*. New York: Praeger.

Blume, M. E., & Friend, I. (1973). A new look at the capital asset pricing model. *Journal of Finance, 28*, 19–33.

Brown, S. J., & Goetzmann, W. N. (1997). Mutual fund styles. *Journal of Financial Economics, 43*(3), 373–399.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2006). *Robust inference with multiway clustering* (Working Paper).

Chacko, G., & Viceira, L. M. (2003). Spectral GMM estimation of continuous-time processes. *Journal of Econometrics, 116*(1), 259–292.

Chang, C. F. (1999). *Determinants of capital structure and management compensation: The partial least squares approach*. Ph.D. dissertation, Rutgers University.

Chang, H. S., & Lee, C. F. (1977). Using pooled time-series and cross section data to test the firm and time effects in financial analysis. *Journal of Financial and Quantitative Analysis, 12*, 457–471.

Chang, C., Lee, A. C., & Lee, C. F. (2009). Determinants of capital structure choice: A structural equation modeling approach. *Quarterly Review of Economic and Finance, 49*(2), 197–213.
Chen, H. Y. (2011). *Momentum strategies, dividend policy, and asset pricing test*. Ph.D. dissertation, Rutgers University.
Chen, S. N., & Lee, C. F. (1981). The sampling relationship between Sharpe's performance measure and its risk proxy: Sample size, investment horizon and market conditions. *Management Science, 27*(6), 607–618.
Chen, K. H., & Shimerda, T. A. (1981). An empirical analysis of useful finance ratios. *Financial Management, 10*(1), 51–60.
Chen, W. P., Chung, H., Lee, C. F., & Liao, W. L. (2007). Corporate governance and equity liquidity: Analysis of S&P transparency and disclosure rankings. *Corporate Governance: An International Review, 15*(4), 644–660.
Cheng, D. C., & Lee, C. F. (1986). Power of alternative specification errors tests in identifying misspecified market models. *The Quarterly Review of Economics and Business, 16*(3), 6–24.
Cheng, K. F., Chu, C. K., & Hwang, R. C. (2010). Predicting bankruptcy using the discrete-time semiparametric hazard model. *Quantitative Finance, 10*, 1055–1066.
Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica, 28*, 591–605.
Core, J. E. (2000). The directors' and officers' insurance premium: An outside assessment of the quality of corporate governance. *Journal of Law, Economics, and Organization, 16*(2), 449–477.
Cox, J. C., Ross, S. A., & Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of Financial Economics, 7*(3), 229–263.
Fair, R. C., & Jaffee, D. M. (1972). Methods of estimation of market in disequilibrium. *Econometrica, 40*, 497–514.
Fama, E. F. (1971). Parameter estimates for symmetric stable distributions. *Journal of the American Statistical Association, 66*, 331–338.
Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: Empirical tests. *Journal of Political Economy, 81*, 607–636.
Fok, R. C. W., Lee, C. F., & Cheng, D. C. (1996). Alternative specifications and estimation methods for determining random beta coefficients: Comparison and extensions. *Journal of Financial Studies, 4*(2), 61–88.
Frecka, T. J., & Lee, C. F. (1983). Generalized financial ratio adjustment processes and their implications. *Journal of Accounting Research, 21*, 308–316.
Gordon, J., & Gordon, M. (1997). The fiinite horizon expected return model. *The Financial Analyst Journal, 53*, 52–61.
Granger, C. W. J., & Newbold, P. (1973). Some comments on the evaluation of economic forecasts. *Applied Economics, 5*(1), 35–47.
Granger, C. W. J., & Newbold, P. (1974). Spurious regressions in econometrics. *Journal of Econometrics, 2*, 111–120.
Granger, C. W. J., & Ramanathan, R. (1984). Improved methods of combining forecasts. *Journal of Forecasting, 3*, 197–204.
Hamilton, J. D. (1994). *Time series analysis*. Princeton: Princeton University Press.
Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica, 50*(4), 1029–1054.
Hansen, B. E. (1996). Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica, 64*(2), 413–430.
Hansen, B. E. (1997). Approximate asymptotic p-values for structural change tests. *Journal of Business and Economic Statistics, 15*, 60–67.
Hansen, B. E. (1999). Threshold effects in non-dynamic panels: Estimation, testing, and inference. *Journal of Econometrics, 93*, 345–368.
Hansen, B. E. (2000a). Sample splitting and threshold estimation. *Econometrica, 68*(3), 575–603.
Hansen, B. E. (2000b). Testing for structural change in conditional models. *Journal of Econometrics, 97*, 93–115.

Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies, 6*(2), 327–343.

Hsiao, C. (2003). *Analysis of panel data* (2nd ed.). Cambridge MA: Cambridge University Press.

Hutchinson, J. M., Lo, A. W., & Poggio, T. (1994). A nonparametric approach to pricing and hedging derivative securities via learning networks. *Journal of Finance, 49*(3), 851–889.

Hwang, R. C., Cheng, K. F., & Lee, J. C. (2007). A semiparametric method for predicting bankruptcy. *Journal of Forecasting, 26*, 317–342.

Hwang, R. C., Wei, H. C., Lee, J. C., & Lee, C. F. (2008). On prediction of financial distress using the discrete-time survival model. *Journal of Financial Studies, 16*, 99–129.

Hwang, R. C., Cheng, K. F., & Lee, C. F. (2009). On multiple-class prediction of issuer crediting ratings. *Journal of Applied Stochastic Models in Business and Industry, 25*, 535–550.

Hwang, R. C., Chung, H., & Chu, C. K. (2010). Predicting issuer credit ratings using a semiparametric method. *Journal of Empirical Finance, 17*(1), 120–137.

Ittner, C. D., Larcker, D. F., & Rajan, M. V. (1997). The choice of performance measures in annual bonus contracts. *Accounting Review, 72*(2), 231–255.

Jones, J. (1991). Earnings management during import relief investigations. *Journal of Accounting Research, 29*, 193–228.

Kao, L. J., & Lee, C. F. (2012). Alternative method for determining industrial bond ratings: theory and empirical evidence. *International Journal of Information Technology & Decision Making, 11*(6), 1215–1236.

Kau, J. B., & Lee, C. F. (1976). Functional form, density gradient and the price elasticity of demand for housing. *Urban Studies, 13*(2), 181–192.

Kau, J. B., Lee, C. F., & Sirmans, C. F. (1986). *Urban econometrics: Model developments and empirical results*. Greenwich: JAI Press. Research in Urban Economics, Vol. 6.

Kaufman, L., & Rousseeuw, P. J. (1990). *Finding groups in data: An introduction to cluster analysis* (9th ed.). New York: Wiley-Interscience.

Kim, D. (1995). The errors in the variables problem in the cross-section of expected stock returns. *Journal of Finance, 50*(5), 1605–1634.

Kim, D. (1997). A reexamination of firm size, book-to-market, and earnings price in the cross-section of expected stock returns. *Journal of Financial and Quantitative Analysis, 32*(4), 463–489.

Kim, D. (2010). Issues related to the errors-in-variables problems in asset pricing tests. In *Handbook of quantitative finance and risk management* (Part V, Chap. 70, pp. 1091–1108).

Kuan, C. M., & Hornik, K. (1995). The generalized fluctuation test: A unifying view. *Econometric Reviews, 14*, 135–161.

Lambert, R., & Larcker, D. (1987). An analysis of the use of accounting and market measures of performance in executive compensation contracts. *Journal of Accounting Research, 25*, 85–125.

Lee, C. F. (1973). *Errors-in-variables estimation procedures with applications to a capital asset pricing model*. Ph.D. dissertation, The State University of New York at Buffalo.

Lee, C. F. (1976a). A note on the interdependent structure of security returns. *Journal of Financial and Quantitative Analysis, 11*, 73–86.

Lee, C. F. (1976b). Functional form and the dividend effect of the electric utility industry. *Journal of Finance, 31*(5), 1481–1486.

Lee, C. F. (1977a). Functional form, skewness effect and the risk-return relationship. *Journal of Financial and Quantitative Analysis, 12*, 55–72.

Lee, C. F. (1977b). Performance measure, systematic risk and errors-in-variable estimation method. *Journal of Economics and Business, 39*, 122–127.

Lee, A. (1996). *Cost of capital and equity offerings in the insurance industry*. Ph.D. dissertation, The University of Pennsylvania in Partial.

Lee, C. F., & Chen, H. Y. (2013). *Alternative errors-in-variable models and their application in finance research* (Working Paper). Rutgers University.

Lee, A. C., & Cummins, J. D. (1998). Alternative models for estimating the cost of capital for property/casualty insurers. *Review of Quantitative Finance and Accounting, 10*(3), 235–267.

Lee, C. F., & Jen, F. C. (1978). Effects of measurement errors on systematic risk and performance measure of a portfolio. *Journal of Financial and Quantitative Analysis, 13*(2), 299–312.
Lee, K. W., & Lee, C. F. (2012). *Multiple directorships, corporate ownership, firm performance* (Working Paper). Rutgers University.
Lee, C. F., & Wu, C. C. (1985). The impacts of kurtosis on risk stationarity: Some empirical evidence. *Financial Review, 20*(4), 263–269.
Lee, C. F., & Zumwalt, J. K. (1981). Associations between alternative accounting profitability measures and security returns. *Journal of Financial and Quantitative Analysis, 16*, 71–93.
Lee, C. F., Newbold, P., Finnerty, J. E., & Chu, C. C. (1986). On accounting-based, market-based and composite-based beta predictions: Method and implications. *Financial Review, 21*, 51–68.
Lee, C. F., Wei, K. C. J., & Bubnys, E. L. (1989). The APT versus the multi-factor CAPM: Empirical evidence. *Quarterly Review of Economics and Business, 29*, 6–16.
Lee, C. F., Wu, C. C., & Wei, K. C. J. (1990). Heterogeneous investment horizon and capital asset pricing model: Theory and implications. *Journal of Financial and Quantitative Analysis, 25*, 361–376.
Lee, C. F., Tsai C. M., & Lee A. C. (2011a). Asset pricing with disequilibrium price adjustment: Theory and empirical evidence. *Quantitative Finance*.
Lee, C. F., Gupta, M. C., Chen, H. Y., & Lee, A. C. (2011b). Optimal payout ratio under uncertainty and the flexibility hypothesis: Theory and empirical evidence. *Journal of Corporate Finance, 17*(3), 483–501.
Lee, C. F., Chien, C. C., Lin, H. C., Lin, Y. C., & Lin, F. C. (2013). *Tradeoff between reputation concerns and economic dependence for auditors – Threshold regression approach* (Working Paper). Rutgers University.
Leon, A., & Vaello-Sebastia, A. (2009). American GARCH employee stock option valuation. *Journal of Banking and Finance, 33*(6), 1129–1143.
Lien, D. (2010). An note on the relationship between the variability of the hedge ratio and hedging performance. *Journal of Futures Market, 30*(11), 1100–1104.
Lien, D., & Shrestha, K. (2007). An empirical analysis of the relationship between hedge ratio and hedging horizon using wavelet analysis. *Journal of Futures Market, 27*(2), 127–150.
Liu, B. (2006). *Two essays in financial econometrics: I: Functional forms and pricing of country funds. II: The term structure model of inflation risk premia*. Ph.D. dissertation, Rutgers University.
Maddala, G. S., Rao, C. R., & Maddala, G. S. (1996). *Handbook of statistics 14: Statistical methods in finance*. Amsterdam: Elsevier Science & Technology.
Martin, C. (1990). Corporate borrowing and credit constrains: Structural disequilibrium estimates for the U.K. *Review of Economics and Statistics, 72*(1), 78–86.
Mayer, W. J. (1989). Estimating disequilibrium models with limited a priori price-adjustment information. *Journal of Econometrics, 41*(3), 303–320.
Miller, M. H., & Modigliani, F. (1966). Some estimates of the cost of capital to the utility industry, 1954-7. *American Economic Review, 56*(3), 333–391.
Myers, R. J. (1991). Estimating time-varying optimal hedge ratios on futures markets. *Journal of Futures Markets, 11*(1), 39–53.
Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica, 55*(3), 703–708.
Ohlson, J. S. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research, 19*, 109–131.
Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies, 22*, 435–480.
Pinches, G. E., & Mingo, K. A. (1973). A multivariate analysis of industrial bond ratings. *Journal of Finance, 28*(1), 1–18.
Quandt, R. E. (1988). *The econometrics of disequilibrium*. New York: Basil Blackwell Inc.

Rendleman, R. J., Jr., & Barter, B. J. (1979). Two-state option pricing. *Journal of Finance, 24*, 1093–1110.

Rubinstein, M. (1994). Implied binomial trees. *Journal of Finance, 49*, 771–818.

Schroder, M. (1989). Computing the constant elasticity of variance option pricing formula. *Journal of Finance, 44*(1), 211–219.

Sealey, C. W., Jr. (1979). Credit rationing in the commercial loan market: Estimates of a structural model under conditions of disequilibrium. *Journal of Finance, 34*, 689–702.

Sears, R. S., & John Wei, K. C. (1988). The structure of skewness preferences in asset pricing model with higher moments: An empirical test. *Financial Review, 23*(1), 25–38.

Shapiro, A. F. (2005). *Fuzzy regression models* (Working Paper). Penn State University.

Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *The Journal of Business, 74*, 101–124.

Thompson, S. B. (2011). Simple formulas for standard errors that cluster by both firm and time. *Journal of Financial Economics, 99*(1), 1–10.

Thursby, J. G. (1985). The relationship among the specification error tests of Hausman, Ramsey and Chow. *Journal of the American Statistical Association, 80*(392), 926–928.

Titman, S., & Wessels, R. (1988). The determinants of capital structure choice. *Journal of Finance, 43*(1), 1–19.

Tsai, G. M. (2005). *Alternative dynamic capital asset pricing models: Theories and empirical results*. Ph.D. dissertation, Rutgers University.

Van Der Klaauw, W. (2002). Estimating the effect of financial aid offers on college enrollment: A regression-discontinuity approach. *International Economic Review, 43*(4), 1249–1287.

Wei, K. C. (1984). *The arbitrage pricing theory versus the generalized intertemporal capital asset pricing model: Theory and empirical evidence*. Ph.D. dissertation, University of Illinois at Urbana-Champaign.

White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica, 48*(4), 817–838.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). Cambridge, MA: The MIT Press.

Yang, C. C. (1989). *The impact of new equity financing on firms' investment, dividend and debt-financing decisions*. Ph.D. dissertation, The University of Illinois at Urbana-Champaign.

Yang, C. C., Lee, C. F., Gu, Y. X., & Lee, Y. W. (2010). Co-determination of capital structure and stock returns – A LISREL approach: An empirical test of Taiwan stock markets. *Quarterly Review of Economic and Finance, 50*(2), 222–233.

Zeileis, A., Leisc, F., Hornik, K., & Kleiber, C. (2002). Strucchange: An R package for testing for structural change in linear regression models. *Journal of Statistical Software, 7*, 1–38.

Experience, Information Asymmetry, and Rational Forecast Bias

2

April Knill, Kristina L. Minnick, and Ali Nejadmalayeri

Contents

Abstract

We use a Bayesian model of updating forecasts in which the bias in forecast endogenously determines how the forecaster's own estimates weigh into the posterior beliefs. Our model predicts a concave relationship between accuracy in forecast and posterior weight that is put on the forecaster's self-assessment.

A. Knill (✉)
The Florida State University, Tallahassee, FL, USA
e-mail: aknill@cob.fsu.edu

K.L. Minnick
Bentley University, Waltham, MA, USA
e-mail: kminnick@bentley.edu

A. Nejadmalayeri
Department of Finance, Oklahoma State University, Oklahoma, OK, USA
e-mail: ali.nejadmalayeri@okstate.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_2,

We then use a panel regression to test our analytical findings and find that an analyst's experience is indeed concavely related to the forecast error.

This study examines whether it is ever rational for analysts to post biased estimates and how information asymmetry and analyst experience factor into the decision. Using a construct where analysts wish to minimize their forecasting error, we model forecasted earnings when analysts combine private information with consensus estimates to determine the optimal forecast bias, i.e., the deviation from the consensus. We show that the analyst's rational bias increases with information asymmetry, but is concavely related with experience. Novice analysts post estimates similar to the consensus but as they become more experienced and develop private information channels, their estimates become biased as they deviate from the consensus. Highly seasoned analysts, who have superior analytical skills and valuable relationships, need not post biased forecasts.

Keywords

Financial analysts • Forecast accuracy • Information asymmetry • Forecast bias • Bayesian updating • Panel regressions • Rational bias • Optional bias • Analyst estimation • Analyst experience

2.1 Introduction

Extant evidence suggests an intimate link between an analyst's experience and her forecasting performance. Analysts who are experienced and highly specialized often forecast better than others (Clement and Tse 2005; Bernhardt et al. 2006). One way they do so is by posting an optimistic bias (Mest and Plummer 2003; Gu and Xue 2007). Novice analysts with limited resources tend to herd with others, which results in almost no bias (Bernhardt et al. 2006). In theory, superior forecasters produce better estimates either by resolving information asymmetry or by offering a better assessment. Lim (2001) suggests that analysts can improve forecast accuracy by strategically biasing their forecasts upwards, which placates management, and in essence purchases additional information.[1] Analysts with

[1]Beyer (2008) argues that, even without incentives to appease management, analysts may still post forecasts that exceed median earnings because managers can manipulate earnings upward to prevent falling short of earnings forecasts. Moreover, Conrad et al. (2006) find support for the idea that analysts' "... recommendation changes are "sticky" in one direction, with analysts reluctant to downgrade." Evidence also indicates that analysts rarely post sell recommendations for a stock, suggesting that losing a firm's favor can be viewed as a costly proposition. At the extreme, firms even pursue legal damages for an analyst's unfavorable recommendations. In a 2001 congressional hearing, president and chief executive officer of the Association for Investment Management and Research told the US House of Representatives Committee on Financial Services, Capital Markets Subcommittee, that "...In addition to pressures within their firms, analysts can also be, and have been, pressured by the executives of corporate issuers to issue favorable reports and recommendations. Regulation Fair Disclosure notwithstanding, recent history...has shown that companies retaliate against analysts who issue 'negative'

long histories of examining firms in a particular industry can also offer a unique perspective, and they signal their ability by posting biased estimates to signal superior ability (Bernhardt et al. 2006).

Of course, analysts do not indefinitely and indiscriminately bias forecasts to appease the firm or signal their ability. This is mainly because analysts can learn from the forecasts of other analysts (Chen and Jiang 2006). By incorporating information from other forecasts, analysts can improve the accuracy of their own forecasts without posting biased estimates. The important question then is: given the analyst's own assessment ability and her efficacy in procuring private information versus using consensus information, how should she construct an optimal forecast? Does that optimal forecast ever include bias, and how does information asymmetry affect this decision? We address these questions by analytically and empirically examining how an analyst's experience and information asymmetry affect her forecasting. In so doing, we also account for the role of the consensus estimate in an analyst's forecast. We begin by modeling the problem of optimal forecasting. To be specific, we combine key features of current rational forecasting models by Lim (2001) and Chen and Jiang (2006). As in Chen and Jiang (2006), analysts in our model form rational (i.e., minimum squared error) forecasts by weighing both public and private information.[2] Following Lim (2001), our analysts post rational forecasts that deviate from the consensus to purchase private information from managers. Motivated by Lim (2001) and Bernhardt et al. (2006), we also allow analysts to post biased forecasts because they have more expertise than the consensus. The novelty of our approach is that we directly model how information asymmetry and analyst experience combine to affect the purchase of private information for use in the forecast deviation. We are also able to model how analysts with different levels of experience, i.e., novice, moderately experienced, and highly seasoned analysts, construct their forecast.

We analytically derive the optimal deviation from the consensus as one that minimizes the mean squared error while allowing for rational Bayesian updating based on public and private knowledge. Our analysis shows that even in a rational forecast framework, analysts' forecast deviation depends on the observed consensus deviation of other analysts. When analysts observe others deviating from the consensus (especially those with more experience), they gain enough insight to avoid posting a large deviation themselves. Our results confirm the findings of both Bernhardt et al. (2006) and Chen and Jiang (2006) – that analysts can *rationally* herd. In the presence of the informative consensus, analysts choose to herd with each other, rather than post estimates that are biased.

Our theory suggests that the likelihood of posting biased estimates, conditional on the consensus, is significantly influenced by the analyst's ability to process

recommendations by denying them direct access to company executives and to company-sponsored events that are important research tools. Companies have also sued analysts personally, and their firms, for negative coverage...."

[2]See also Han et al. (2001).

information. Consistent with Hong et al. 2000), we show that novice analysts essentially herd. Without either honed analytical abilities or valuable relationships, these analysts follow the premise that the consensus is more accurate than their own private information. Second, we show that a moderately experienced analyst relies more on her sources inside the firm than her analytical skills. This manifests itself as a biased forecast since she must appease management to tap into those sources. This link between experience and deviation is not monotonic. Highly seasoned analysts do not purchase as much additional information (or they purchase the information at a reduced price), either because they possess superior analytical skills or because their valuable relationships with firms afford them beneficial information without the optimistic forecast. This preferential treatment is akin to that which is afforded to companies in relationship lending with banks (Petersen and Rajan 1994). Indeed, Carey et al. (1998) argue that the relationship lending of banks can also be ascribed to some nonbank financial intermediaries. Although the firms that an analyst covers are not necessarily financial, the same relationship could certainly exist and is based on information asymmetry.

We further demonstrate that as the analyst-firm information asymmetry increases, so does the bias. Similar to the results found in Mest and Plummer (2003), analysts find private channels and analytical ability valuable mitigating factors when faced with information asymmetry. Moderately experienced analysts, who begin to tap into reliable private channels without the valuable relationships that might afford preferential treatment, post larger deviations with the hope of ascertaining better information. This suggests that both information asymmetry and experience interactively affect the way analysts balance public and private information to form forecasts. Our model also shows that the effect of information asymmetry and analyst experience on rational deviation depends on the dispersion of and the correlation between public and private signals. The quality of private information channels, the informativeness of consensus, and the connectedness of public and private signals significantly affect how analysts form forecasts.

To examine the validity of our analytical findings, we empirically investigate how analyst experience and information asymmetry affect forecast deviation from the consensus.[3] Our empirical results confirm our theoretical predictions that rational bias (i.e., deviation from the consensus) is concavely related to analyst experience and positively associated with information asymmetry. Novice analysts and highly seasoned analysts post forecasts with smaller bias, while moderately seasoned analysts post estimates that deviate further from the consensus. Moderately seasoned analysts can benefit from a positive bias if they have the confidence to separate from the herd and access reliable private sources of information.

[3]Clement and Tse (2005) are the closest to our analysis; however while they admit that the observed link between inexperience and herding can be a complex issue that might have other roots than just career concerns, they do not provide detailed insight as to what and how this complexity develops.

These results are stronger for earlier forecasts versus later ones. As analysts become highly seasoned, with external information networks and superior skills in forecasting, they find appeasing management to purchase information less useful (or less costly in the same way that relationship lending affords firms cheaper capital). As one might expect, when information asymmetry increases, rational deviation from the consensus also increases. The degree of information asymmetry faced by seasoned analysts has a significant positive effect on the forecast deviation (information asymmetry also affects novice analysts but not as extensively).

This study contributes to a growing literature on rational bias in analyst forecasting. Motivated by recent works by Bernhardt et al. (2006), Beyer (2008), Chen and Jiang (2006), Lim (2001), and Mest and Plummer (2003), we take an integrated modeling approach to arrive at a rational forecast bias, which is based on Bayesian updating of public consensus and endogenously acquired private information. Our approach is unique in that the forecast bias affects how public and private information are combined. Specifically, analysts can use bias to improve forecast accuracy by purchasing private information but can also learn from other analysts by following the consensus. Our analytical findings, which are empirically confirmed, show that unlike behavioral models, analysts can rationally herd. Unlike signaling models, seasoned analysts do not always post biased estimates. It also shows the value of the relationships that both moderately experienced and highly seasoned analysts leverage to gain reliable private information. These results are of practical interest because they provide evidence that analysts can and may optimally bias their earnings estimates and that this optimal bias differs across both analyst experience and information asymmetry. This knowledge may be useful to brokerage houses for training purposes and/or evaluation of analyst performance, particularly across industries with different levels of information asymmetry.

2.2 Theoretical Design

Recent studies show that earnings forecasts reflect public information and private assessments (Boni and Womack 2006; Chen and Jiang 2006; Lim 2001; Ramnath 2002). As Bernhardt et al. (2006) note, analysts' forecasts of public information partly relate to the consensus, common information, and unanticipated market-wide shocks. We thus model analyst earnings forecasts using two components of earnings information: common and idiosyncratic, with some uncertainty about each component. In doing so, we also assume, as does Lim (2001), that the idiosyncratic component directly relates to the private information analysts can obtain from the firm by posting an optimistic view of the firm's prospects (Nutt et al. 1999). Our typical analyst also observes a previously posted forecast whereby the analyst can learn about both the common and idiosyncratic components of earnings by incorporating previously disclosed information. Since we assume analysts are Bayesians, they can learn from previous forecasts as an alternative to posting a positive deviation to purchase private signals from the firm. Since these prior estimates partially reflect private information, the analysts may heavily weigh them into their assessment, depending on the perceived information asymmetry and experience of the previous analysts.[4]

In this section, we model analyst earnings forecasting in the presence of forecast uncertainty. Following recent studies (Lim 2001; Ramnath 2002), our analyst has an unconditional estimate, $E = X + \varepsilon$, about earnings, $X \sim N(0, \sigma^2)$, with some uncertainty, $\varepsilon \sim N(0, {\tau_e}^{-2})$. As in Chen and Jiang (2006), our analyst also observes a noisy consensus forecast, $E_c = X + \varepsilon_c$, with a consensus uncertainty, $\varepsilon_c \sim N(0, {\tau_c}^{-2})$. As a Bayesian, our analyst forms a conditional forecast, $F = w\,E + (1 - w)\,E_c$, by combining her unconditional estimate with the consensus. The optimal weight in her conditional forecast minimizes her squared forecast error.[5]

As in Lim (2001), the optimal forecasting, however, is endogenous to private information acquisition. Our analyst's forecast precision relates to the private information analysts can obtain from the firm by posting an optimistic view of the firm's prospects. That is to say, while forecast precision, τ_e, consists of a self-accuracy component, τ_0, reflecting the analyst expertise, the forecast precision can be improved by $\tau(b)$ through placating managers by posting positively biased, b, forecasts. The marginal precision per bias, $\partial\tau/\partial b$, reflects the analyst's information asymmetry. This is because an analyst faced with greater information asymmetry should derive a larger marginal benefit from a biased forecast. Since the conditional forecast partially reflects private information, the analyst's optimal rational forecast (and forecast bias) depends on the analyst's expertise and information asymmetry. As noted before, the objective of the analyst is to minimize her squared forecast error, $F - X$:

$$\min_{w|b} \mathbf{E}\left[(F - X)^2\right] = \min_{w|b} \left[(wE + (1 - w)E_c)^2\right] \tag{2.1}$$

where $\mathbf{E}[\bullet]$ is the expectation operator. Let's assume that the analyst's estimate and the consensus are correlated, $\mathbf{E}[\varepsilon, \varepsilon_c] = \rho\ (\tau\ \tau_c)^{-1}$. This correlation reflects the extent to which the analyst uses common private channels. The analyst objective, Eq. 2.1, can be rewritten as

[4]Here, we focus only on the case of one-period sequential forecasting. However, we believe that the main implications of our model hold true for a multi-period sequential forecasting setting. Since we assume that the probabilistic characteristics of different components are known and analysts can gauge each others' experience and the amount of information asymmetry perfectly, there would be no incentive to deviate from posting commensurate optimal, rational forecasts. If expert analysts intentionally deviate from their optimal forecasts, no other analyst can compensate for their experience or information asymmetry [for more discussion, see Trueman (1990)].

[5]For more details on Bayesian methods of inference and decision making, see Winkler (1972).

$$\min_{w|b}\left[w^2(\tau_0+\tau(b))^{-2}+(1-w)^2\tau_c{}^{-2}+2\rho w(1-w)(\tau_0+\tau(b))^{-1}\tau_c^{-1}\right]. \quad (2.2)$$

The optimal weight which solves the aforementioned objective function is

$$w=\frac{(\tau_0+\tau(b))^2-\rho\tau_c(\tau_0+\tau(b))}{(\tau_0+\tau(b))^2+\tau_c^2-2\rho\tau_c(\tau_0+\tau(b))} \quad (2.3)$$

Proposition 1 *Assuming positively biasing forecasts have increasing but diminishing returns*, i.e., $\partial\tau/\partial b > 0$ *and* $\partial^2\tau/\partial b^2 < 0$, *the optimal weight on analyst's own unconditional estimate is concavely related to the analyst's expertise* (i.e., *self-accuracy*, τ_0). *For analysts with self-accuracy less* (*more*) *than* $-\tau(b)+0.5\ \rho^{-1}\ \tau_c$, *the optimal weight on the analyst's own unconditional estimate increases* (*decreases*) *with the expertise. At the optimal weight*, *the analyst's conditional precision*, $\tau_0+\tau(b)$, *equals* $0.5\ \rho^{-1}\ \tau_c$.

Proof See the Appendix 1.

As the analyst's expertise, τ_0, increases, her unconditional estimate precision rises relative to the consensus estimate precision. The analyst would then gain more accuracy by placing more weight on her own assessment, consistent with the findings of Chen and Jiang (2006). Forecast bias, however, affects how the analyst's own unconditional estimate weighs into the optimal forecast. As forecast bias increases, the analyst relies more on her own assessment because larger bias improves precision through more privately acquired information. However, this reliance on private information is limited. When the precision of private information reaches a certain limit, $0.5\ \rho^{-1}\ \tau_c$, the analyst starts to reduce her reliance on her own assessment. Since the consensus contains information, an analyst does not need to rely solely on private information to improve her overall accuracy.

Interestingly, the threshold on the reliance of private information is inversely related to the correlation between analyst's own and consensus precisions. When the signals from the consensus and analyst are highly correlated, the analyst need not post largely biased forecasts to improve her accuracy. At low correlations, however, the analyst almost exclusively relies on her own assessment and uses private channels heavily by posting biased forecasts. As Mest and Plummer (2003) show, when uncertainty about the company is high, management becomes a more important source of information. This confirms Bernhardt et al. (2006) contention that signal "correlatedness" affects analyst forecasting.

Proposition 1 has an interesting testable implication. As noted, following previous studies (Lim 2001; Chen and Jiang 2006), our analysts minimize the squared error to arrive at the optimal bias. In a cross section of analysts, this implies that as long as all analysts follow the squared-error-minimization rule, then the implication of Proposition 1 holds empirically: there would be a concave relation between experience and the weight an analyst places on her own unconditional assessment. However, since the weighted average of self-assessment and consensus belief is theoretically identical with forecast error, this then also implies that in the cross section, analysts' forecast errors and their experience are also related concavely.

Note that our analysts choose how to combine their own unconditional forecast and the consensus forecast to arrive at their reported forecast. They do so by accounting for the inherent error in each of these forecasts and choose a weight that minimizes the squared forecast error: the difference between reported forecast (i.e., conditional forecast) and the observed earnings. So for a novice analyst, the choice is what weight to put on her unconditional forecast knowing that the forecasts reported by other analysts contain much more experience and perhaps less information asymmetry. In the extreme case, where the novice analyst has no confidence on her own assessment, the optimal weight for her unconditional forecast is zero. She will fully herd. A highly seasoned analyst does the same thing: she also chooses the weight she puts on her assessment vis-à-vis consensus. In the alternate extreme case, where the highly seasoned analyst has utmost confidence on her assessment, she puts 100 % weight on her unconditional forecast. Moderately seasoned analysts thus fall somewhere in between; the weight they put on their own unconditional forecasts is between zero and one. In such a setting, all analysts arrive at their own squared-error-minimizing forecast. From the onset, however, as econometricians, we can only observe their reported conditional forecast. This means that we can only focus on the implication of analyst squared-error-minimizing exploiting cross-sectional differences in the data. As noted, the observed error is equal to the weighted average of unconditional forecast and the consensus belief. This implies that observed error is directly linked with the unobservable weight. However from Proposition 1, we know that the unobservable optimal weight is concavely related to experience, which in turn, implies that, in the cross section of analysts, the observed forecast error is concavely linked with the experience as well.

Proposition 2 *Assuming positively biasing forecasts have increasing but diminishing returns*, i.e., $\partial\tau/\partial b > 0$ *and* $\partial^2\tau/\partial b^2 < 0$, *the optimal weight on an analyst's own unconditional estimate increases monotonically with information asymmetry* (i.e., *the marginal accuracy for bias*, $\partial\tau/\partial b$).

Proof See the Appendix 1.

As the efficacy of the analyst's private information acquisition increases, that is, as $\partial\tau/\partial b$ rises, the analyst gains more precision with every cent of bias. More resourceful analysts, such as those employed by large investment houses, either have better private channels or can gain more from the same channels (see, e.g., Chen and Jiang 2006; Clement and Tse 2005). As such, these analysts' estimates become more accurate more rapidly as they bias their forecasts to purchase information from the firm. From proposition 1, we know that at the optimal weight, the sum of the analyst's own estimate and consensus precision is constant.

As information asymmetry increases, i.e., marginal precision per bias rises, a larger bias is needed to obtain the optimal weight of the analyst's estimate and the consensus. Interestingly, as an analyst becomes more experienced, her optimal weight on her own estimate becomes less sensitive to information asymmetry.

2.3 Empirical Method

We empirically test selected comparative statics to draw conclusions about the validity of our analytical results. We focus our attention on analytical predictions that result directly from the endogeneity of private information acquisition. Specifically, we test three of our model's implications: (1) whether analyst experience and forecast deviation are concavely related, (2) whether information asymmetry and analyst forecast deviation are positively related, and (3) whether there is interplay between the effects of experience and information asymmetry. We follow Chen and Jiang (2006) and empirically define bias as the difference between an analyst's forecast and the consensus estimate. We calculate consensus using only the most recent estimate from each analyst and include only those estimates that are less than 90 days old (Lim 2001).

We first examine the impact of analyst experience on forecast deviation. As analysts become more experienced, they post a more positive (i.e., increasing) deviation to improve their information gathering, resulting in better forecasts. Proposition 1 from the theoretical model suggests that this relationship is nonlinear. Highly seasoned analysts achieve forecast accuracy on their own without relying on procured information, or they possess valuable relationship with management that does not necessitate purchasing information, causing this nonlinearity. Empirically, we expect the relationship between deviation and experience to be concave (i.e., a positive coefficient on experience and a negative coefficient on experience squared). To test our contentions, we estimate the following OLS panel regression:

$$\begin{aligned}\mathrm{DFC}_{i,t} = {} & \alpha_0 + \eta_0\,\mathrm{Experience}_{i,t} + \eta_1\,\mathrm{Experience}_{i,t}{}^2 + \beta_0\,\mathrm{Num\ Revisions}_{i,t} \\ & + \beta_1\mathrm{Same\ Quarter}_{i,t} + \beta_2\,\mathrm{Reg.FD}_{i,t} + \boldsymbol{\Phi}_0\mathbf{X}_{t-4} + \boldsymbol{\Phi}_1\mathbf{I} + \boldsymbol{\Phi}_2\mathbf{t} + \varepsilon,\end{aligned} \tag{2.4}$$

where DFC is the Deviation from the consensus for analyst i in quarter t. We have several analyst-specific controls: Experience, Experience squared, and the NumRevisions. Following our analytical model, analysts with more experience should be better at forecasting earnings and should have better channels of private information. Following Leone and Wu (2007), we measure Experience in two ways: using an absolute measure (the natural log of the quarters of experience) and a relative measure (the natural log of the quarters of experience divided by the average of this measure for analysts following firms in the same industry). We include Experience^2 to test for a nonlinear relationship between Deviation and Experience. To control for varying degrees of accuracy due to intertemporal Bayesian updating, we include NumRevisions, which is the number of times the analyst revises her estimate. Studies such as Givoly and Lakonishok (1979) suggest that information has been acquired (i.e., purchased). This suggests a positive relationship between the number of revisions and the resulting deviation; thus we expect the coefficient, β_0, to be positive. Same quarter is an indicator variable for horizon value, where the variable is equal to one if the estimate occurs in the same quarter as the actual earnings report is released and zero otherwise. Including Same quarter can help us understand whether analysts are more likely to be optimistic early in the period than late in the period (Mikhail et al. 1997; Clement 1999; Clement and Tse 2005).[6] We expect the coefficient on this variable, β_1, to be negative. Reg. *FD* is an indicator variable equal to one if the quarter date is after Reg. FD was passed (October 23, 2000), and zero otherwise.[7] Although extant literature is divided on whether or not Reg. FD has decreased the information asymmetry for analysts, if information asymmetry is decreased and more public information available, analysts will be less likely to purchase private information, and the coefficient on Reg. FD, β_2, would be negative. Indeed, Zitzewitz (2002) shows that although private information has decreased post Reg. FD, the amount of public information has improved. Both Brown et al. (2004) and Irani (2004) show that information asymmetry decreased following Reg. FD.[8]

We control for firm effects through the inclusion of firm-specific variables such as Brokerage reputation, Brokerage size, Accruals, Intangible assets, and Return st. dev. (vector X in Eq. 2.4). Following Barber et al. (2000), we use the number of companies a brokerage house follows per year in total as a proxy for brokerage size. Brokerage house reputation may play an integral role in accessing to private information (Agrawal and Chen 2012). To calculate broker reputation, we start with the Carter and Manaster (1990), Carter et al. (1998) and the Loughran and Ritter (2004) rankings. When a firm goes public, the prospectus lists all of the firms that are in the syndicate, along with their shares. More prestigious underwriters are listed higher in the underwriting section. Based upon where the underwriting brokerage firm is listed, they are assigned a value of 0–9, where nine is the highest ranking. As Carter and Manaster (1990) suggest that prestigious financial institutions provide a lower level of risk (i.e., lower information asymmetry), we include control variables for these characteristics of the firm and expect the relationships to be negative.[9]

---

[6]Horizon value and the NumRevisions are highly correlated at 65 %. We therefore orthogonalize horizon value in the equation to ensure that multicollinearity is not a problem between these two variables.

[7]http://www.sec.gov/rules/final/33-7881.htm.

[8]See Lin and Yang (2010) for a study of how Reg. FD affects analyst forecasts of restructuring firms.

[9]Brokerage reputation and Brokerage size are highly correlated at 67 %. We therefore orthogonalize brokerage reputation in the equation to ensure that multicollinearity is not a problem between these two variables.

As suggested by Bannister and Newman (1996) and Dechow et al. (1998), companies that consistently manage their earnings are easier to forecast. We include Accruals (in $ millions) to control for the possibility that earnings are easier to forecast for companies that manage their earnings, resulting in less information "bought" from the company.[10] We expect the coefficient on this control variable to be negative. We include Intangible assets based on the fact that companies with greater levels of intangible assets are more difficult to forecast based on uncertainty of future performance (Hirschey and Richardson 2004).[11] Thus, we expect the coefficient to be positive. We include Return standard deviation to proxy for firm risk. More volatile firms are less likely to voluntarily issue public disclosures, making it necessary for analysts to purchase private information (Waymire 1986). The standard deviation is defined as the monthly standard deviation of stock returns over a calendar year. All firm-level variables are lagged one year (i.e., four quarters). We include both industry (one-digit SIC code) and time indicators to control for industries/times where it is easier to forecast (see Kwon (2002) and Hsu and Chiao (2010), e.g., to see how analyst accuracy differs across industry).[12] Finally, a formal fixed effects treatment around analysts is taken to ensure that standard errors are not understated.

Our theoretical analysis also examines the effect of information asymmetry on forecast deviation. In an environment with high information asymmetry, analysts without adequate, reliable resources need to post a positive deviation to access information, as shown in Proposition 2. We expect a positive coefficient on Information asymmetry. We test this relationship by estimating the following OLS panel regression:

$$\begin{aligned}\mathrm{DFC}_{\mathrm{i,t}} = {} & \alpha_0 + \lambda_0\,\text{Information Asymmetry}_{i,t} + \beta_0\,\text{Num Revisions}_{\mathrm{i,t}} + \beta_1\,\text{Same Quarter}_{\mathrm{i,t}} \\ & + \beta_2\,\text{Reg.FD}_{\mathrm{i,t}} + \boldsymbol{\Phi}_0\mathbf{X}_{t-4} + \boldsymbol{\Phi}_1\mathbf{I} + \boldsymbol{\Phi}_2\mathbf{t} + \varepsilon,\end{aligned} \tag{2.5}$$

where t is the quarter in which we measure deviation and i denotes the ith analyst. We define Information asymmetry three different ways: (1) the inverse of analyst coverage, i.e., 1/(number of brokerage houses following the firm), (2) the standard deviation of the company's forecasts ($\times 10^2$), and (3) the relative firm size, i.e., the difference between the firm's assets and the quarterly median assets for the industry. These definitions are constructed such that the direction of the expected marginal coefficient is congruent with that of information asymmetry.

[10]Following Stangeland and Zheng (2007), we measure Accruals as income before extraordinary items (Data #237) minus cash flow from operations, where cash flow from operations is defined as net cash flow from operating activities (Data #308) minus extraordinary items and discontinued operations (Data #124).

[11]Following Hirschey and Richardson (2004), we calculate Intangibles as intangible assets to total assets (Data 33/Data #6).

[12]As an alternate proxy for industry fixed effects, Fama-French 12-industry classifications (Fama and French 1997) are used. Results using these proxies are available upon request.

The first definition, the inverse of analyst coverage, is based on evidence that the level of financial analyst coverage affects how efficiently the market processes information (Bhattacharya 2001). Analyst coverage proxies for the amount of public and private information available for a firm and, therefore, captures a firm's information environment (Zhang 2006). The second definition, Forecast dispersion, is supported in papers such as Krishnaswami and Subramaniam (1998) and Thomas (2002), among others. Lastly, we use Relative firm size as a proxy for information asymmetry. This is equally supported by the literature, including but not limited to Petersen and Rajan (1994) and Sufi (2007). The firm-specific control variables are the same as in Eq. 2.4, and a fixed effects treatment around analysts is taken.

As previously explained, our model allows for the interaction of experience and information asymmetry; the concavity of the relationship between experience and Deviation may change at different levels of information asymmetry. Thus, we estimate Eq. 2.4 for analysts with both high and low levels of information asymmetry (i.e., based on relation to the industry-quarter median forecast dispersion). We expect the coefficients on Experience and Experience squared to increase when there is more information asymmetry.

Similarly, we demonstrate that experience affects the link between information asymmetry and deviation. While the deviation of the novice and highly seasoned analysts is only marginally affected by information asymmetry, the deviation of the moderately experienced analyst is highly affected by information asymmetry. Since Proposition 2 suggests that information asymmetry will affect analysts differently based on their experience, we segment our sample into three groups: novice, moderately experienced, and highly seasoned analysts. To form these experience groups, we create quarterly terciles of analysts based on their experience. The bottom third comprises the novice analysts; the middle third, moderately seasoned analysts; and the top third, highly seasoned analysts. We examine the model in Eq. 2.5 separately for our three experience subsamples. The coefficient on information asymmetry should be larger for experienced analysts.

It is possible that a resolution of idiosyncratic uncertainty makes deviation insensitive to both analyst experience and information asymmetry. Regulations such as Reg. FD were enacted to reduce the uncertainty around companies (De Jong and Apilado 2008). If Reg. FD reduced firm-specific uncertainty more than common uncertainty, then the coefficients on Experience, Experience squared, and information asymmetry should be smaller after Reg. FD. In other words, if Reg. FD was effective in "leveling the playing field" for analysts with regard to preferential treatment of some analysts over others, the implications of this model should no longer exist.

Extant evidence suggests that by prohibiting exclusive private communication of pertinent information, Reg. FD would cause overall earnings uncertainty to decline (Baily et al. 2003). In Eqs. 2.4 and 2.5, we simply control for any effect that Reg. FD might have. By using interactive terms, however, we can explore this relationship more fully by offering a comparison of the relationship between Experience/ Information asymmetry and Deviation both with and without Reg. FD. If most of the decrease in uncertainty comes from improving the common component of earnings (or equivalently, reducing the amount of private information), then the enactment of Reg. FD should make any remaining private information even more valuable and should lead to an increase in the importance of information asymmetry on deviation. Further, this effect should be disparate based on analyst experience. Specifically, experienced analysts (moderately experienced more so than highly seasoned) should see a drop in their deviation from consensus based on an increase in information asymmetry since their private information is no longer quite so private, i.e., there is less information to buy. Novice analysts, on the other hand, will likely be relatively unaffected since they depend mostly on consensus anyway. In short, since the intent of the regulation is to even the playing field for analysts, the effect of Reg. FD on the impact of experience could be a reduction in its importance. This would especially be the case with relative experience. If Reg. FD achieved what it set out to achieve, we should see a reduction (i.e., flattening) of the concavity of the experience relationship with deviation from consensus.

2.4 Data

Consistent with Brown and Sivakumar (2003) and Doyle et al. (2003), we define earnings as the First Call reported actual earnings per share.[13] Our sample includes companies based in the United States with at least two analysts, consists of 266,708 analyst-firm-quarter forecasts from 1995 to 2007, and includes all analysts' revisions. Variables are winsorized at the 1 % level to ensure that results are not biased by outliers.

First Call data is well suited for examining analyst revisions because most of the analysts' estimates in the data have the date that they were published by the broker. These revisions are reflected daily, which aids in understanding the changes in deviation based on changes in the information environment. One limitation with First Call data is that it identifies only brokerage houses, not individual analysts. Following Leone and Wu (2007), we make the assumption that for each firm quarter there is only one analyst in each brokerage house following the firm. It is noteworthy that this biases our study against finding a nonlinear impact for analyst's experience.

Panel A of Table 2.1 shows the summary statistics of the variables used in our analysis. The average Deviation from the consensus is 5.16¢. However, there is wide dispersion, from −54.67¢ to 112.50¢, as compared to an average Forecast error of −3.79¢, with a range of −131–72¢. The analysts in our sample have on average 12.71 quarters of Experience (the most experienced analyst has 13 years experience, and the least experienced analyst has no prior experience following the firm). Analysts revise their estimates 2.49 times per quarter. The companies they

[13]As a robustness test, we use I/B/E/S data. Results may be found in Appendix 2.

Table 2.1 Data characteristics

Panel A. Summary statistics

Variable	N	Source	Mean	Median	Std. dev.	Min	Max
DFC (¢)	266,708	First call; own calculation	5.16***	1***	20.43	−54.67	112.5
Quarters follow	266,708	First call; own calculation	12.71***	10***	10.37	1	52
Experience	266,708	First call; own calculation	2.15***	2.30***	0.98	0	3.81
Relative experience	266,708	First call; own calculation	0.97***	1***	0.51	0.07	2.40
Analyst coverage^{-1}	266,708	First call; own calculation	0.21***	0.13***	0.22	0.03	1
Forecast dispersion (¢)	253,673	First call; own calculation	7.90***	4.19***	10.66	0	67.44
Relative firm size	217,126	COMPUSTAT; own calculation	−9.68***	0***	60.80	−1430.35	671.78
Forecast error (¢)	266,708	First call; own calculation	−3.79***	0***	24.00	−131	72
Same quarter	266,708	First call; own calculation	0.21***	0***	0.41	0	1
Broker reputation	266,708	Carter and Manaster (1990)	0.02***	0.20***	1.46	−4.80	3.77
Broker size	266,708	First call; own calculation	6.36***	6.63***	1.03	0	7.79
Reg. FD	266,708	Securities Exchange Commission	0.63***	1***	0.48	0	1
Number of revisions	266,708	First call; own calculation	2.49***	2***	2.60	0	12
Accruals	266,708	COMPUSTAT	−0.44***	−0.09***	1.09	−7.52	0.90
Intangible assets	266,708	COMPUSTAT	0.14***	0.06***	0.18	0	0.92
Return std. dev.	266,708	CRSP	0.13***	0.11***	0.08	0.03	0.47

Panel B. Correlation

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
DFC (1)	1.00															
Quarters follow (2)	**−0.02**	1.00														
Experience (3)	**0.01**	**0.89**	1.00													
Relative experience (4)	**0.01**	**0.70**	**0.72**	1.00												

Analyst coverage^{-1} (5)	**0.01**	**−0.19**	**−0.18**	**−0.01**	1.00											
Forecast dispersion (6)	**0.32**	**0.06**	**0.05**	0.00	**−0.11**	1.00										
Relative firm size (7)	**0.01**	**−0.10**	**−0.08**	**−0.01**	**0.10**	**−0.03**	1.00									
Forecast error (8)	**−0.89**	**0.04**	0.00	**−0.01**	**−0.03**	**−0.27**	**−0.03**	1.00								
Same quarter (9)	**−0.10**	**−0.14**	**−0.26**	**−0.13**	**−0.02**	**−0.03**	0.00	**0.09**	1.00							
Broker reputation (10)	**0.02**	**−0.01**	**−0.03**	**0.04**	**−0.02**	**−0.01**	**−0.01**	**−0.02**	0.00	1.00						
Broker size (11)	0.00	**0.30**	**0.32**	**0.19**	**−0.18**	**0.03**	**−0.05**	**0.01**	**−0.05**	**0.08**	1.00					
Reg. FD (12)	**−0.05**	**0.33**	**0.28**	**0.01**	**−0.17**	**0.02**	**−0.06**	**0.07**	**−0.05**	**−0.18**	**0.19**	1.00				
NumRevisions (13)	**0.14**	**0.21**	**0.27**	**0.12**	**−0.15**	**0.21**	**−0.03**	**−0.11**	0.00	**−0.03**	**0.10**	**0.09**	1.00			
Accruals (14)	**−0.01**	**−0.17**	**−0.14**	**−0.02**	**0.19**	**−0.11**	**0.41**	**0.00**	**0.01**	**−0.03**	**−0.08**	**−0.09**	**−0.06**	1.00		
Intangible assets (15)	**0.00**	**0.02**	**0.03**	**0.01**	**−0.02**	**−0.10**	**0.06**	**−0.01**	**−0.02**	**−0.02**	**0.07**	**0.15**	**−0.04**	**−0.01**	1.00	
Return std. dev. (16)	**0.13**	**−0.21**	**−0.18**	**−0.04**	**0.09**	**0.12**	**0.10**	**−0.12**	**−0.05**	**−0.03**	**−0.08**	**−0.05**	**0.05**	**0.10**	**−0.03**	1.00

Analyst information comes from First Call for the term 1995–2007. Company information comes from Compustat/CRSP. DFC is the bias, defined as the difference between an analyst's forecast and the consensus estimate, where we calculate consensus using only the most recent estimate from each analyst and include only those estimates that are less than 90 days old. Quarters follow is the number of quarters a brokerage firm has been following a company. Experience is the natural log of the quarters of experience. Relative experience is the natural log of the quarters of experience divided by the average of this measure for analysts following firms in the same industry. Analyst coverage^{-1} is equal to 1/(number of brokerage houses following the firm), and forecast dispersion is the standard deviation of the company's forecasts ($\times 10^2$). Relative firm size is the difference between the firm's assets and the quarterly median assets for the industry. Forecast error is the difference between the estimate and the actual earnings. Same quarter is a dummy variable equal to one if the estimate is in the same quarter as the actual and zero otherwise. Brokerage reputation is a ranking of brokerage reputation, where 0 is the worst and 9 is the best. Brokerage size is the natural log of the number of companies per quarter a brokerage house follows. Reg. FD is an indicator variable that takes on a value of 1 if Reg. FD is in effect and 0 otherwise. NumRevisions is the total number of revisions the analyst made during the quarter for the quarter-end earnings. Accruals are income before extraordinary items minus cash flow from operations, where cash flow from operations is defined as net cash flow from operating activities minus extraordinary items and discontinued operations. Intangible assets are intangible assets to total assets. Return std. dev. is the standard deviation of the 12-month returns. Panel A shows the univariate statistics and Panel B shows the pairwise correlations

Asterisks in Panel A represent statistical significance relative to zero

Bolded numbers in Panel B represent 5 % or 1 % significance follow have a Return standard deviation of 13 %, Intangibles of 14 %, and Accruals of −$0.44 (in millions) on average.

Panel B of Table 2.1 shows the correlation matrix for the variables used in the analysis. There exists notable significant relation in the variables: Forecast error and Experience with Deviation from consensus. The relation actually foreshadows one of the main results of the paper, which is that private information can be used to decrease forecast error. The correlation found in this table, −0.89, is not a problem econometrically because fitted values are used in the specification found in Table 2.5. The only other correlations that would be considered a problem econometrically are for variables not used in the same specification, i.e., Experience and Relative experience.

2.5 Empirical Findings

Table 2.2 presents the results of our estimation from Eq. 2.4, which examines how analyst experience affects analyst deviation from the consensus. The coefficients on our control variables all exhibit the expected signs. There is a negative relation between the horizon value (Same quarter) and Deviation, which suggests that analysts are more likely to be optimistic early in the period than late in the period. Supporting the contentions of Carter and Manaster (1990), there is a negative relationship between the brokerage characteristics – Reputation and Size – with Deviation. There is also a negative relationship between Reg. FD and Deviation, suggesting that the enactment of Reg. FD and its mandatory indiscriminant information revelation have made forecasting easier. NumRevisions is positively related to Deviation; revisions are made when valuable information content is received (Givoly and Lakonishok 1979), suggesting that private information is paid for through incremental deviation of their forecasts over time. We find that higher Accruals lead to less deviation from the consensus. Firms who actively manage earnings make forecasting easier; thus analysts do not have to purchase private information. The positive sign on Intangible assets is expected as more intangible assets make forecasting more difficult, necessitating the procurement of private information. Finally, as the standard deviation of Returns increases, future firm performance is more difficult to predict.

Turning to our variables of interest, we find that there is a highly significant positive relationship between Experience and Deviation, as evidenced by a 0.829¢ increase in deviation for every additional unit of experience (specification 1) and a 0.581¢ increase in deviation for every additional unit of relative experience (specification 2).[14] The results in specifications (3) and (4) confirm our contention that this relationship is not linear. The squared Experience variable

[14]Inasmuch as the Experience variable is transformed using the natural logarithm, one unit of experience is approximately equal to two quarters of experience. For tractability, we refer to this as a unit in the empirical results.

Table 2.2 Impact of experience on analyst forecasting

	All estimates				Early estimates (horizon is longer)		Late estimates (horizon is shorter)		Estimates in low IA environment		Estimates in high IA environment	
	1	2	3	4	5	6	7	8	9	10	11	12
Experience	0.829***		3.024***		6.751***		1.945***		0.950***		4.745***	
	[0.049]		[0.153]		[0.571]		[0.137]		[0.109]		[0.283]	
Experience^2			−0.595***		−1.341***		−0.360***		−0.169***		−0.986***	
			[0.039]		[0.118]		[0.038]		[0.028]		[0.073]	
Relative experience		0.581***		4.250***		2.961***		2.712***		1.781***		5.572***
		[0.081]		[0.263]		[0.301]		[0.167]		[0.186]		[0.488]
RelExp^2				−1.735***		−0.953***		−0.764***		−0.712***		−2.267***
				[0.118]		[0.087]		[0.055]		[0.083]		[0.221]
Same quarter	−2.372***	−2.472***	−1.992***	−2.364***					−0.833***	−0.935***	−4.181***	−4.791***
	[0.107]	[0.107]	[0.110]	[0.107]					[0.078]	[0.076]	[0.205]	[0.200]
NumRevisions	1.204***	1.232***	1.189***	1.227***	0.958***	0.965***	1.545***	1.569***	0.499***	0.512***	1.388***	1.444***
	[0.016]	[0.016]	[0.016]	[0.016]	[0.025]	[0.025]	[0.030]	[0.030]	[0.013]	[0.013]	[0.027]	[0.027]
Reg. FD	−2.458***	−2.648***	−2.275***	−2.403***	0.731*	0.591	1.687***	1.467***	−0.368**	−0.391**	−3.394***	−3.613***
	[0.227]	[0.227]	[0.228]	[0.228]	[0.388]	[0.387]	[0.406]	[0.405]	[0.167]	[0.167]	[0.408]	[0.408]
Broker reputation	−0.508***	−0.529***	−0.495***	−0.499***	−0.642***	−0.573***	−0.300***	−0.242**	−0.205***	−0.206***	−0.714***	−0.714***
	[0.103]	[0.103]	[0.103]	[0.103]	[0.185]	[0.185]	[0.104]	[0.104]	[0.075]	[0.075]	[0.186]	[0.186]
Broker size	−0.453***	−0.493***	−0.690***	−0.518***	−0.695**	−0.696**	−0.489***	−0.391**	−0.519***	−0.476***	−0.759***	−0.451*
	[0.151]	[0.151]	[0.152]	[0.151]	[0.275]	[0.275]	[0.153]	[0.152]	[0.112]	[0.112]	[0.270]	[0.268]
Accruals	−0.122^{***c}	−0.186***	−0.165***	−0.159***	−0.363***	−0.373***	0.114***	0.112***	0.204***	0.205***	−0.465***	−0.452***
	[0.037]	[0.037]	[0.037]	[0.037]	[0.061]	[0.061]	[0.041]	[0.041]	[0.028]	[0.028]	[0.066]	[0.066]
Intangible assets	1.370***	1.244***	1.237***	1.264***	2.718***	2.718***	0.097	0.146	1.415***	1.411***	1.974***	2.054***
	[0.232]	[0.232]	[0.232]	[0.231]	[0.393]	[0.393]	[0.241]	[0.242]	[0.161]	[0.161]	[0.440]	[0.440]
Return std. dev.	19.384***	17.800***	19.063***	17.700***	26.067***	26.308***	7.387***	7.448***	7.944***	7.265***	17.102***	15.640***
	[0.578]	[0.571]	[0.578]	[0.571]	[0.975]	[0.976]	[0.612]	[0.612]	[0.448]	[0.441]	[1.003]	[0.991]
Constant	6.239***	8.269***	6.501***	6.913***	−2.193	4.345*	6.268***	6.633***	6.286***	6.423***	13.895***	14.185***
	[1.439]	[1.434]	[1.438]	[1.437]	[2.682]	[2.621]	[1.467]	[1.469]	[0.975]	[0.974]	[3.041]	[3.038]
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes

(*continued*)

Table 2.2 (continued)

	All estimates				Early estimates (horizon is longer)		Late estimates (horizon is shorter)		Estimates in low IA environment		Estimates in high IA environment	
	1	2	3	4	5	6	7	8	9	10	11	12
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	266,708	266,708	266,708	266,708	134,296	134,296	132,412	132,412	133,574	133,574	133,134	133,134
# Brokerage houses	130	130	130	130	121	121	130	130	128	128	128	128
R-squared	0.06	0.06	0.06	0.06	0.07	0.07	0.04	0.04	0.04	0.04	0.08	0.08

Table 2.2 presents the results of our estimation from Eq. 2.4. Experience is the natural log of the number of quarters the analyst has followed the firm for which forecast error is calculated. Relative experience is the analyst experience scaled by the average analyst experience (same industry). Information asymmetry (IA) is proxied based on the forecast dispersion – the standard deviation of estimates. Same quarter is a dummy variable equal to one if the estimate is in the same quarter as the actual and zero otherwise. Same quarter is orthogonalized (on NumRevisions) to ensure that multicollinearity is not a problem between these two variables. NumRevisions is the total number of revisions the analyst made during the quarter for the quarter-end earnings. Reg. FD is an indicator variable that takes on a value of 1 if Reg. FD is in effect and 0 otherwise. X is a vector of firm-specific variables including Broker reputation, Broker size, Accruals, Intang. assets, and Return std dev. Brokerage reputation is a ranking of brokerage reputation, where 0 is the worst and 9 is the best. Brokerage size is the natural log of the number of companies per quarter a Brokerage house follows. Brokerage reputation is orthogonalized (on brokerage size) to ensure that multicollinearity is not a problem between these two variables. Accruals are the accrued revenue/liabilities utilized for earnings smoothing. Intang. assets are the covered firm's intangible assets value relative to its total assets. Return std. dev. is the standard deviation of the covered firm's return. I is a vector of one-digit SIC industry dummies. T is a vector of time dummies. A formal fixed effects treatment around analysts is employed. Standard errors are reported in brackets

*, **, and *** indicate significance levels of 10 %, 5 %, and 1 %, respectively (two-tailed test). In columns 1–4 we use the full sample. In columns 5–8, we use the median value of horizon value to divide the firms into early (columns 5–6) and late estimates (7–8) groups to perform the analysis. In columns 9–12, we use the median value of forecast dispersion to divide the firms into low information asymmetry (columns 9–10) and high information asymmetry (11–12) groups to perform the analysis (both definitions) is significantly and negatively related to Deviation (−0.595¢ in specification 3 for Experience and −1.735¢ in specification 4 for Relative experience). Confirming the analytical results, the empirical results suggest that there exists a trade-off between public and private information. The results suggest that when analysts first begin to follow a firm and gain experience, they are more likely to post small deviations, indicating that they weigh public information (i.e., previous forecasts and the consensus), more heavily than private information. A possible explanation is that they have not yet acquired the preferred access to information that long-term relationships afford experienced analysts or the analytical expertise to wade through noisy information. As analysts gain confidence in their analytical ability and relationships with key personnel within the firm, however, they are more likely to post large deviations to gain private information. Highly seasoned analysts put almost no weight on the consensus in creating their earnings forecasts and rely almost exclusively on their own private information. This information costs them very little, either because they have honed their analytical ability so well that they don't need the information or because the cost of said information is reduced due to their preferential relationship with the firm.

If our hypotheses are true, we would expect that analysts are more optimistic early in the period and less so (and perhaps even pessimistic) later in the period (shortly before earnings are announced).[15] In order to test this, we segment our sample by the median horizon value into two categories, early estimates (i.e., longer horizon) and late estimates (i.e., shorter horizon). The results are substantially stronger for earlier estimates as compared to late estimates. The coefficient on the Experience variable is 6.751 when the horizon value in long and 1.945 when there is a short horizon value (specification 5 vs. 7). The coefficients on the squared Experience variable are also more extreme for the earlier analysts (−1.341 vs. −0.360), showing that the concavity of the function is more pronounced early in the forecasting period and less so later. This pronounced concavity suggests that the distinction between moderately experienced analysts and novice/highly experienced analysts is more pronounced earlier in the estimation period. Essentially, for every additional quarter of experience, there is a 5.410¢ increase in deviation when horizon is long and a 1.585¢ increase in deviation when horizon is short. Results for Relative experience are qualitatively identical, albeit with a reduced difference between early and late. Results are qualitatively identical when we divide horizon into terciles and define early as the highest quartile and late as the lowest. We can conclude from these results that analysts are indeed more optimistic earlier in the estimation period.

Next, we rerun the model in Eq. 2.4 on low and high information asymmetry subsamples (using forecast dispersion and segmenting at the median), shown in Table 2.2, columns 9–12. As Proposition 2 suggests, the empirical results indicate that the impact of experience on optimal deviation is smaller for analysts when information asymmetry is low. One additional quarter of Experience increases the

[15]We are grateful to an anonymous referee for this point.

analyst deviation by 0.950¢ (specification 9) in a low information asymmetry environment versus 4.745¢ (specification 11) in a high information asymmetry environment. One additional unit of Relative experience increases the analyst Deviation by 1.781¢ (specification 10) in a low information asymmetry environment versus 5.572¢ (specification 12) in a high information asymmetry environment. We note once again that this relationship is not linear, as evidenced by the statistically significant negative squared term. This implies that moderately experienced analysts post higher deviations from consensus than do their less and more experienced colleagues.

Table 2.3 provides alternate specifications around the Reg. FD control variable. Specifications (1) through (4) exclude the Reg. FD control to ensure that results are not reliant on its inclusion. Overall, results from Table 2.2 remain consistent. That said, we note that Relative experience in Specification (2) is no longer statistically significant. The substantiation of the nonlinearity of this proxy for Experience in Specification (4), albeit muted, suggests that this is likely due to the fact that we are no longer controlling for the impact of Reg. FD on the deviation from consensus, which we would expect would affect Relative Experience more so than experience (i.e., no longer benefitting highly seasoned analysts for the preferred relationships through better access to information) through the concavity of the function. In other words, without controlling for Reg. FD independently, the "average" impact of the proxy (i.e., before and after Reg. FD) is muted.

Specifications (4) through (8) explore further the effect Reg. FD has on the relationship between Experience and Deviation. We first note that the nonlinearity of the relationship between Experience and forecast deviation (Deviation) is intact, indicating that leveling the playing field through Reg. FD has not (completely) changed how forecasts are fundamentally linked with the information. Focusing on the more potent results, we look to the Relative experience results (Specifications 7 and 8), and we note that Reg. FD *is* effective in reducing the private information component, which in turn reduces the importance of relative experience. Empirically, this translates into a reduction in the concavity of the impact of Relative experience; the "leveling of the playing field" flattens out that curve. We see this in the Reg. FD interaction terms: specifically, a positive and significant marginal effect on (Reg. FD * Relative experience) and negative and significant marginal effect on (Reg. FD * Relative experience^2). More concretely stated, when access to private information is reduced, the relationship-based competitive advantage of relative experience – preferential access to private information – is taken away.

These results highlight a distinction between precision and the value of preferential relationships between veteran analysts and management of the firm. Since all analysts gain analytical expertise over time following the firm (though the marginal effect of this would seem to wane for the most experienced, thus explaining the maintained nonlinearity), but only some may gain preferential treatment by the firm, one could argue that Experience is a better proxy for analytical precision from the model and Relative experience is a better proxy for the preferential treatment provided to veteran analysts that possess valuable relationships with top management of the firms they are covering (i.e., management will provide private

Table 2.3 Time impact of Reg. FD on the relationship between experience/relative experience and deviation

Dependent variable = DFC (deviation from consensus)								
	1	2	3	4	5	6	7	8
Experience	0.302***		2.116***		1.528***		2.764***	
	[0.050]		[0.157]		[0.088]		[0.271]	
Experience^2			−0.485***				−0.400***	
			[0.040]				[0.082]	
Reg. FD * experience					−0.962***		0.026	
					[0.100]		[0.326]	
Reg. FD * experience^2							−0.168*	
							[0.092]	
Relative experience		−0.072		0.511*		0.822***		7.019***
		[0.082]		[0.269]		[0.149]		[0.503]
Relative experience^2				−0.274**				−3.128***
				[0.120]				[0.243]
Reg. FD * relative experience						−0.331*		−3.715***
						[0.172]		[0.586]
Reg. FD * relative experience^2								1.831***
								[0.277]
Reg. FD					−0.319	−2.308***	−1.398***	−0.975***
					[0.318]	[0.288]	[0.362]	[0.358]

(*continued*)

Table 2.3 (continued)

Dependent variable = DFC (deviation from consensus)								
	1	2	3	4	5	6	7	8
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	266,708	266,708	266,708	266,708	266,708	266,708	266,708	266,708
# Brokerage houses	130	130	130	130	130	130	130	130
Model R^2	0.06	0.06	0.06	0.06	0.06	0.06	0.06	0.06

Table 2.3 provides alternate specifications around the Reg. FD control variable. Experience is the natural log of the number of quarters the analyst has followed the firm for which forecast error is calculated. Relative experience is the analyst experience scaled by the average analyst experience (same industry). Reg. FD is an indicator variable that takes on a value of 1 if Reg. FD is in effect and 0 otherwise. Control variables included in the analysis but not shown above include same quarter, NumRevisions, Reg. FD, Broker reputation, Broker size, Accruals, Intang. assets, and Return std. dev. Same quarter is a dummy variable equal to one if the estimate is in the same quarter as the actual and zero otherwise. Same quarter is orthogonalized (on NumRevisions) to ensure that multicollinearity is not a problem between these two variables. NumRevisions is the total number of revisions the analyst made during the quarter for the quarter-end earnings. Brokerage reputation is a ranking of Brokerage reputation, where 0 is the worst and 9 is the best. Brokerage size is the natural log of the number of companies per quarter a brokerage house follows. Brokerage reputation is orthogonalized (on brokerage size) to ensure that multicollinearity is not a problem between these two variables. Accruals are the accrued revenue/liabilities utilized for earnings smoothing. Intang. assets are the covered firm's intangible assets value relative to its total assets. Return standard deviation is the standard deviation of the covered firm's return. I is a vector of one-digit SIC industry dummies. T is a vector of time dummies. A formal fixed effects treatment around analysts is employed. Standard errors are reported in brackets

*, **, and *** indicate significance levels of 10 %, 5 %, and 1 %, respectively (two-tailed test) information to only analysts with the oldest relationships). In short, Reg. FD reduces the importance of the competitive advantage of preferential treatment (proxied by Relative experience) while retaining the importance of analytical ability (proxied by Experience). This is seen empirically in the insignificant change in concavity of Experience and a decrease in the concavity of Relative experience. In this light, we see that the impact of Reg. FD is what we would expect, both with regard to Experience (seen in Specifications 5 and 6) and Relative experience.

Panel A of Table 2.4 shows the results of Eq. 2.5, which examines directly how Information asymmetry affects Deviation, controlling for firm, analyst, and regulatory factors. Specifications (1) through (3) look at the effects of Information asymmetry on the full sample of analysts, using the inverse of Analyst coverage, Forecast dispersion, and Relative firm size, respectively. The first proxy (specification 1) is the inverse of Analyst coverage. As predicted in the theoretical section, Information asymmetry is positively related to Deviation. With fewer resources at their disposal, analysts post greater deviations as a means to purchase private information. This supplements the marginally valuable public information available, leading to more accurate forecasts. When information asymmetry is measured by Analyst coverage, a unit increase in the information asymmetry measure leads to a 0.630¢ additional Deviation. When using Forecast dispersion as the proxy, a one-unit increase in the information asymmetry measure leads to a 0.630¢ increase in Deviation. Lastly, using Relative firm size as the proxy for information asymmetry, a one-unit increase in the information asymmetry measure leads to 0.006¢ additional Deviation. From the multivariate analysis, it is evident that Forecast dispersion is the best proxy for Deviation from the consensus. The specifications where information asymmetry is proxied by Forecast dispersion have approximately two times the predictive power of the analyst coverage and size proxy (R-Squared of 0.15 vs. 0.06 and 0.08, respectively).

In specifications (4)–(6) of Table 2.4, we include indicator variables for the three experience terciles (suppressing the constant term in the model) to examine whether information asymmetry affects analysts in a different manner based on experience. We first note that the marginal effect on Information asymmetry remains positive as in the first three specifications. Looking to the marginal effect of novice analysts – the least experienced third of analysts in a given quarter – they find sticking closer to the herd a better alternative (Zhou and Lai 2009). The cost of acquiring precise information for an inexperienced analyst who lacks necessary resources would be prohibitively large, i.e., the analyst would need to post an outrageously optimistic (positively deviated) forecast. The marginal effect on moderately experienced analysts – the middle third in experience in a given quarter – suggests that they post deviations that are more positive than the average analyst in an effort to purchase valuable private information. The marginal effect for the tercile of highly seasoned analysts suggests that they post a deviation with a smaller magnitude. This nonlinear relationship exists regardless of the proxy for information asymmetry. Controlling for the information asymmetry in the estimation environment, these results fall in line with our prediction of nonlinearity for Experience and our conjecture that highly seasoned analysts possess both superior

Table 2.4 Impact of information asymmetry on analyst forecasting

Panel A: All analysts						
Dependent variable = DFC (deviation from consensus)						
	1	2	3	4	5	6
Analyst coverage^{-1}	0.630***			0.627***		
	[0.191]			[0.188]		
Forecast dispersion		0.630***			0.626***	
		[0.004]			[0.004]	
Relative firm size			0.006***			0.007***
			[0.001]			[0.001]
Novice				5.015***	2.756***	2.305**
				[1.079]	[1.045]	[1.156]
Moderately experienced				6.495***	3.920***	3.959***
				[1.080]	[1.046]	[1.157]
Highly seasoned				6.100***	3.320***	3.824***
				[1.082]	[1.049]	[1.159]
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes
Observations	266,708	253,673	217,126	266,708	253,673	217,126

# Brokerage houses	130	130	124	130	130	124
Model R^2	0.06	0.15	0.08	0.12	0.21	0.14
Novice = Moderate?				75***	62***	68***
Moderate = Highly seasoned?				26***	49***	7***
Novice = Moderate = Highly seasoned?				38***	39***	35***

Panel B: Experience subsamples			
Dependent variable = DFC (deviation from consensus)			
Information asymmetry=	Coverage	Forecast dispersion	Relative firm size
	1	2	3
Novice	0.591**	0.555***	0.007***
	[0.264]	[0.006]	[0.002]
Observations	91,993	85,906	73,888
# Brokerage houses	130	130	124
Model R^2	0.05	0.14	0.06
Moderately experienced	1.075***	0.720***	0.014***
	[0.334]	[0.007]	[0.002]
Observations	88,392	83,878	72,656
# Brokerage houses	115	114	109
Model R^2	0.07	0.19	0.09
Highly seasoned	0.456	0.624***	0.001
	[0.432]	[0.007]	[0.001]
Observations	86,323	83,889	70,582
# Brokerage houses	85	85	82
Model R^2	0.08	0.15	0.09

(*continued*)

Table 2.4 (continued)

Panel B: Experience subsamples			
Dependent variable = DFC (deviation from consensus)			
Information asymmetry=	Coverage	Forecast dispersion	Relative firm size
	1	2	3
Novice = Moderate?	2,780***	7,709***	2,421***
Moderate = Highly seasoned?	1,978***	5,862***	3,742***

Table 2.4 shows the results of Eq. 2.5. We create three groups of experience: the bottom third is the novice analysts, the middle third are the moderately seasoned analysts, and the top third are the highly seasoned analysts. Information asymmetry is defined in three ways: Analyst coverage^{-1}, Forecast dispersion, and Relative firm size. Analyst coverage^{-1} is the inverse of the number of analyst covering the firm. Forecast dispersion is the standard deviation of estimates. Relative firm size is equal to the covered firm size minus industry-quarter median firm size (in $billions). Control variables included in the analysis but not shown above include Same quarter, NumRevisions, Reg. FD, Broker reputation, Broker size, Accruals, Intang. assets, and Return std. dev. Same quarter is a dummy variable equal to one if the estimate is in the same quarter as the actual and zero otherwise. Same quarter is orthogonalized (on NumRevisions) to ensure that multicollinearity is not a problem between these two variables. NumRevisions is the total number of revisions the analyst made during the quarter for the quarter-end earnings. Brokerage reputation is a ranking of brokerage reputation, where 0 is the worst and 9 is the best. Brokerage size is the natural log of the number of companies per quarter a brokerage house follows. Brokerage reputation is orthogonalized (on brokerage size) to ensure that multicollinearity is not a problem between these two variables. Accruals are the accrued revenue/liabilities utilized for earnings smoothing. Intang. assets are the covered firm's intangible assets value relative to its total assets. Return standard deviation is the standard deviation of the covered firm's return. I is a vector of one-digit SIC industry dummies. T is a vector of time dummies. A formal fixed effects treatment around analysts is employed. Panel A Specifications 1–3 use the full sample and do not suppress the constant. Panel A Specifications 4–6 suppress the constant so we can include the three experience indicator variables. To test whether the coefficients on the experience indicator variables are statistically different, we use a Wald test for significance. Panel B shows the results of the same analysis using experience-based subsamples. We use a Hausman estimation to test whether there is statistical significance between each experience estimation. Standard errors are reported in brackets

*, **, and *** indicate significance levels of 10 %, 5 %, and 1 %, respectively (two-tailed test)

analytical abilities, which would suggest that perhaps they have to purchase less information to achieve precision, and valuation relationships, which allow them to purchase information at reduced "costs" (i.e., deviation from consensus). To ensure that the marginal effects are significantly different, we perform Wald tests to confirm that the differences between the coefficients are statistically significant. As can be seen at the bottom of Panel A, the differences are statistically significant in every case.

In Panel B, we examine experience-based subsamples to see more intricately how experience affects the impact of Information asymmetry on Deviation. Results confirm that Experience affects the link between Information asymmetry and Deviation. The moderately experienced analyst is the most affected by an increase in information asymmetry, regardless of the definition of information asymmetry. It is interesting to note that the results for Analyst coverage and Forecast dispersion are much stronger than Relative firm size. We use a Hausman test of statistical significance and find that the experience estimations are all statistically different.

These empirical results nicely complement the theoretical predictions of our model. Taken collectively, the results suggest that analysts' rational forecast deviation from the consensus is a result of analysts maximizing their objective function of minimizing error while taking into consideration the information environment in which they operate.

To provide evidence that the deviation from the consensus discussed in the rest of the paper is indeed rationale, we test whether an analyst posting an estimate with the optimal deviation from consensus, empirically represented by the fitted value of either Eqs. 2.4 or 2.5, achieves a lower forecast error (defined as estimate minus actual earnings). Here, a negative association between fitted Deviation from consensus (DFC*) and Forecast error (FE) is indicative of optimal forecasting. Chen and Jiang (2006) argue that a positive (negative) association between observed Deviation from consensus (DFC) and Forecast error (FE) is indicative of an analyst overweighting (underweighting) her own assessment. This is because if analysts follow a Bayesian updating rule and minimize squared forecast error, there should be no link between observed deviation from forecast and forecast error. They find that analysts, on average, overweigh their own belief, and thus the link between observed deviation from consensus and forecast error is positive. In this paper, we extend Chen and Jiang's model to allow for rational biasing of forecasts, similar to Lim (2001). As such, we find that there should be a concave relationship between Experience and Deviation from consensus. If our theory is correct and if our empirical model of deviation from consensus truly captures the essence of rational forecasting that occurs among analysts, the fitted Deviation from consensus should be devoid of analysts' overconfidence about their own information processing. In fact, if fitted Deviation reflects the tendency toward more rational forecasting, we should find a *negative* relationship between fitted Deviation and Forecast error.

Looking to the results, which are found in Table 2.5, Panel A, we see that this is indeed what we find. In all specifications, the fitted Deviation (DFC) is negatively related with Forecast error. All specifications lead to a significant decline in the forecast error of analyst estimation. Specifically, we find that a 1¢ increase in bias

Table 2.5 Optimal deviation from consensus and forecast error

Panel A: All analysts					
Dependent variable = forecast error					
First-stage regressors included	Experience Experience^2 Reg. FD NumRevisions **Controls**	Relative experience Relative experience^2 Reg. FD NumRevisions **Controls**	Forecast dispersion Reg. FD NumRevisions **Controls**	Experience Experience^2 Forecast dispersion Reg. FD NumRevisions **Controls**	Relative experience Relative experience^2 Forecast dispersion Reg. FD NumRevisions **Controls**
	1	2	3	4	5
DFC^*	-1.081^{***}	-1.079^{***}	-1.029^{***}	-1.030^{***}	-1.029^{***}
	[0.009]	[0.009]	[0.006]	[0.006]	[0.006]
Constant	2.350^{***}	2.150^{**}	2.586^{***}	2.845^{***}	2.425^{***}
	[0.851]	[0.851]	[0.850]	[0.850]	[0.850]
Observations	266,708	266,708	253,673	253,673	253,673
# Brokerage houses	130	130	130	130	130
Model R^2	0.05	0.05	0.12	0.12	0.12
DFC^*	-0.456^{***}	-0.444^{***}	-0.427^{***}	-0.433^{***}	-0.427^{***}
	[0.016]	[0.017]	[0.010]	[0.010]	[0.010]
DFC^{*2}	-0.050^{***}	-0.050^{***}	-0.023^{***}	-0.022^{***}	-0.023^{***}
	[0.001]	[0.001]	[0.000]	[0.000]	[0.000]

Constant	1.765**	1.508*	1.487*	1.707**	1.378*
	[0.840]	[0.841]	[0.791]	[0.781]	[0.790]
Observations	266,708	266,708	253,673	253,673	253,673
# Brokerage houses	130	130	130	130	130
Model R^2	0.06	0.06	0.14	0.14	0.14
Panel B: Experience subsamples					
Novice analysts	3.97	3.97	3.27	3.26	3.27
Moderately experienced	5.50	5.49	4.19	4.18	4.18
Highly seasoned	4.71	4.71	3.92	3.94	3.92

Table 2.5 Panel A presents the results of regressing forecast error on the regressors specified in each column. Experience is the natural log of the number of quarters the analyst has followed the firm for which forecast error is calculated. Relative experience is the analyst experience scaled by the average analyst experience (same industry). NumRevisions is the total number of revisions the analyst made during the quarter for the quarter-end earnings. Reg. FD is an indicator variable that takes on a value of 1 if Reg. FD is in effect and 0 otherwise. Controls are vector of firm-specific variables including Same quarter, Broker reputation, Broker size, Accruals, Intang. assets, and Return std dev. Same quarter is a dummy variable equal to one if the estimate is in the same quarter as the actual and zero otherwise. Same quarter is orthogonalized (on NumRevisions) to ensure that multicollinearity is not a problem between these two variables. Brokerage reputation is a ranking of brokerage reputation, where 0 is the worst and 9 is the best. Brokerage size is the natural log of the number of companies per quarter a brokerage house follows. Brokerage reputation is orthogonalized (on brokerage size) to ensure that multicollinearity is not a problem between these two variables. Accruals are the accrued revenue/liabilities utilized for earnings smoothing. Intang. assets are the covered firm's intangible assets value relative to its total assets. Return standard deviation is the standard deviation of the covered firm's return. I is a vector of one-digit SIC industry dummies. T is a vector of time dummies. Panel B presents the fitted values of DFC from each specification for each experience level. A formal fixed effects treatment around analysts is employed. Standard errors are reported in brackets

*, **, *** indicate significance levels of 10 %, 5 %, and 1 %, respectively (two-tailed test)

Table 2.6 Alternate samples

Dependent variable = DFC (deviation from consensus)												
	Without revisions			With only one analyst			Random effects			I/B/E/S data		
	1	2	3	4	5	6	7	8	9	10	11	12
Experience	0.817***			2.822***			2.158***			1.396***		
	[0.141]			[0.669]			[0.155]			[0.117]		
Experience^2	−0.127***			−0.595***			−0.475***			−0.089**		
	[0.042]			[0.186]			[0.039]			[0.036]		
Relative Experience		1.990***			0.157			0.749***			1.938***	
		[0.267]			[1.266]			[0.267]			[0.191]	
Relative Experience^2		−0.829***			−0.365			−0.351***			−0.658***	
		[0.123]			[0.563]			[0.120]			[0.077]	
Forecast dispersion			0.259***			0.616***			0.629***			0.718***
			[0.005]			[0.157]			[0.004]			[0.003]
Same quarter	−2.479***	−2.564***	−2.860***	−1.716***	−2.455***	−5.852	−2.008***	−2.357***	−3.189***	−1.751***	−2.044***	−2.494***
	[0.099]	[0.097]	[0.096]	[0.515]	[0.492]	[3.657]	[0.110]	[0.107]	[0.104]	[0.090]	[0.085]	[0.077]
NumRevisions				1.598***	1.721***	−0.053	1.186***	1.226***	0.773***	1.076***	1.141***	0.700***
				[0.089]	[0.088]	[0.899]	[0.016]	[0.016]	[0.015]	[0.016]	[0.015]	[0.014]
Reg. FD	−4.923***	−5.004***	−4.349***	−4.747***	−5.084***	46.26	−2.223***	−2.412***	−2.560***	−3.243***	−3.490***	−3.268***
	[0.366]	[0.366]	[0.368]	[0.948]	[0.948]	[110.413]	[0.228]	[0.228]	[0.222]	[0.187]	[0.187]	[0.170]
Broker reputation	−0.172	−0.183*	−0.275**	−0.106	−0.052	60.993	−0.368***	−0.400***	−0.520***			
	[0.110]	[0.110]	[0.110]	[0.430]	[0.431]	[65.397]	[0.088]	[0.090]	[0.086]			
Broker size	−1.025***	−1.032***	−1.006***	−0.095	0.132	7.038	−0.514***	−0.404***	−0.347***			
	[0.142]	[0.141]	[0.143]	[0.458]	[0.457]	[21.599]	[0.125]	[0.127]	[0.123]			
Accruals	0.001	−0.01	0.140***	−3.317***	−3.306***	92.249	−0.171***	−0.160***	0.413***	−0.464***	−0.498***	0.244***
	[0.048]	[0.047]	[0.045]	[0.593]	[0.593]	[212.638]	[0.037]	[0.037]	[0.036]	[0.030]	[0.030]	[0.027]
Intangible assets	0.153	0.121	0.870***	−0.38	−0.419	−3.35	1.216***	1.268***	4.191***	0.881***	0.861***	2.196***
	[0.263]	[0.263]	[0.260]	[0.980]	[0.981]	[25.114]	[0.232]	[0.231]	[0.227]	[0.209]	[0.210]	[0.191]

Return std. dev.	10.012***	9.187***	2.975***	10.716***	9.805***	0.384	18.238***	17.662***	5.043***	19.153***	17.969***	−1.710***
	[0.666]	[0.657]	[0.664]	[2.079]	[2.063]	[35.868]	[0.579]	[0.570]	[0.568]	[0.473]	[0.472]	[0.436]
Constant	10.593***	11.004***	9.734***	8.532	10.407*	57.535	3.465***	4.612***	2.021*	−2.682***	−0.828	2.010***
	[1.409]	[1.405]	[1.392]	[5.760]	[5.770]	[165.519]	[1.258]	[1.267]	[1.219]	[0.693]	[0.694]	[0.619]
Firm FE	Yes	Yes	Yes	Yes	Yes	Yes	No	No	No	Yes	Yes	Yes
Industry FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	66,550	66,550	62,014	13,181	13,181	146	266,708	266,708	253,673	295,717	295,717	295,717
# Brokerage houses†	130	130	130	119	119	22	130	130	130	7708	7708	7708
Model R^2	0.03	0.03	0.07	0.06	0.06	0.24	0.06	0.06	0.16	0.07	0.07	0.23

Analyst information comes from First Call (except in Specifications (10)–(12)). Company information comes from Compustat/CRSP. Experience is the natural log of the number of quarters the analyst has followed the firm for which forecast error is calculated. Relative experience is the analyst experience scaled by the average analyst experience (same industry). Same quarter is a dummy variable equal to one if the estimate is in the same quarter as the actual and zero otherwise. Same quarter is orthogonalized (on NumRevisions) to ensure that multicollinearity is not a problem between these two variables. Forecast dispersion is the standard deviation of estimates. X is a vector of firm-specific variables including Broker reputation, Broker size, Accruals, Intang. assets, and Return std dev. Brokerage reputation is a ranking of brokerage reputation, where 0 is the worst and nine is the best. Brokerage size is the natural log of the number of companies per quarter a brokerage house follows. Brokerage reputation is orthogonalized (on brokerage size) to ensure that multicollinearity is not a problem between these two variables. Accruals are the accrued revenue/liabilities utilized for earnings smoothing. Intang. assets are the covered firm's intangible assets value relative to its total assets. Return standard deviation is the standard deviation of the covered firm's return. I is a vector of one-digit SIC industry dummies. T is a vector of time dummies. Specifications (1)–(3) include only original estimates (i.e., no revisions). Specifications (4)–(6) include only firms with one analyst following the firm. Specifications (7)–(9) use random effects. Specifications (10)–(12) use I/B/E/S data. A formal fixed effects treatment around analysts is employed unless otherwise noted. Standard errors are reported in brackets.

*, **, and *** indicate significance levels of 10 %, 5 %, and 1 %, respectively

†In sample derived from I/B/E/S data, this is actually the number of analysts

Table 2.7 Alternate proxies for information asymmetry

	Inf. asymmetry Reg. FD NumRevisions SameQtr **Controls**	Experience^2 Inf. Asymmetry Reg. FD NumRevisions SameQtr **Controls**	Relative experience^2 Inf. asymmetry Reg. FD NumRevisions SameQtr **Controls**	Inf. asymmetry Reg. FD NumRevisions SameQtr **Controls**	Experience^2 Inf. asymmetry Reg. FD NumRevisions SameQtr **Controls**	Relative experience^2 Inf. asymmetry Reg. FD NumRevisions SameQtr **Controls**
	1	2	3	4	5	6
Panel A: Analyst coverage^{-1}						
DFC	−1.079***	−1.081***	−1.079***	−0.446***	−0.456***	−0.444***
	[0.009]	[0.009]	[0.009]	[0.017]	[0.016]	[0.017]
DFC squared				−0.050***	−0.050***	−0.050***
				[0.001]	[0.001]	[0.001]
Constant	2.171**	2.347***	2.150**	1.523*	1.762**	1.508*
	[0.851]	[0.851]	[0.851]	[0.842]	[0.840]	[0.841]
Observations	266,708	266,708	266,708	266,708	266,708	266,708
# Brokerage houses	130	130	130	130	130	130
Panel B: Relative firm size						
DFC	−1.046***	−1.048***	−1.046***	−0.496***	−0.496***	−0.496***
	[0.008]	[0.008]	[0.008]	[0.015]	[0.015]	[0.015]
DFC squared				−0.041***	−0.042***	−0.041***
				[0.001]	[0.001]	[0.001]
Constant	1.505***	1.513***	1.504***	1.100***	1.122***	1.100***
	[0.066]	[0.066]	[0.066]	[0.066]	[0.066]	[0.066]
Observations	217,126	217,126	217,126	217,126	217,126	217,126
# Brokerage houses	124	124	124	124	124	124

*, **, and *** indicate significance levels of 10 %, 5 %, and 1 %, respectively

leads to a decrease in Forecast error from a low of 1.029 (specification 3) to 1.081¢ (specification 1), depending upon the specification. Considering that the sample average Forecast error is −3.79¢, this is a significant decrease, suggesting that posting such deviations is indeed rational.

Given the high marginal effect and considering the nonlinearity of the impact of experience on deviation from consensus, we further test a squared term of the fitted deviation from consensus variable. It would make sense that continually adding a penny to one's estimate would cease to increase the precision of the resulting estimation at some point. In doing so, we see that the linear effect is reduced considerably. The forecast error is decreased to less than half of a penny on average and this reduction is decreasing with every marginal penny of bias.

Finally, we examine the fitted values of Deviation from consensus for the different terciles of analyst experience, shown in Table 2.5, Panel B. For all five specifications, we see evidence of nonlinearity with these fitted values. Once again, the moderately experienced analyst is found to have the highest "optimal" deviation. This confirms nicely both the theoretical and previous empirical results and helps us to come full circle.

2.6 Robustness

To check for consistency in the results, we reexamine specifications (3) and (4) from Table 2.2 and specification (2) from Table 2.4 on subsamples, altering our empirical methodology and using a different data source. We analyze two subsamples: (1) excluding revisions and (2) including only firms covered by one analyst. We alter our empirical methodology by using random effects as opposed to the fixed effect treatment used in the base specification of the paper. Lastly, we alter our data source by using I/B/E/S data. Because the main empirical tests use First Call data, which only looks at brokerage houses, one potential concern is that perhaps we might not be able to control for analysts' fixed effects effectively. Given our specifications and variables we include in the model, we find that there are only 130 unique brokerage houses. Since I/B/E/S reports analysts rather than brokerage houses, we use their data and find results hold when individual analysts fixed effect (7,708 unique analysts) are included. Due to the nature of IBES data, we cannot, however, control for broker reputation and size. Results are qualitatively identical across all robustness tests and are included in Appendix 2.

For brevity, we only use one information asymmetry proxy in Table 2.5. We choose forecast dispersion as our proxy for information asymmetry since it provides the best model fit (i.e., Model R^2) of the three proxies. To show that our results are robust to the other two information asymmetry proxies, we rerun Table 2.5 using these alternate proxies. Our results are qualitatively identical and are shown in Appendix 3.

In results not reported, we examine three final robustness tests. To address any concerns about the power of our empirics based on the number of observations, we examine the subsample of manufacturing firms. While this subsample is less than half of the original, the economic and statistical significance of coefficients remains, providing support for our empirical results. We include alternate industry dummies using Fama-French industry classifications (SIC codes are used for classification in the base specifications). Results once again remain. Finally, to address any concerns about sample selection problems and the dependence of the power of our empirics on the number of brokerage houses, we run the estimations not including Broker reputation (which reduces our sample to 130 houses) and have 315 brokerage houses. Our results remain. These results are available upon request. The congruence of the results from both the included and unreported robustness tests provides further credence to the conclusions of this study.

2.7 Conclusions

To understand the exact role experience and information asymmetry play in forming a rational deviation, we offer an integrated framework in which an optimum forecast is derived from weighing public and private information, where the resulting forecast bias (i.e., deviation from the consensus) improves accuracy through the acquisition of private information. We study how information asymmetry and analyst experience affect the efficacy of the feedback from bias on private information acquisition. Our construct enables a realistic optimization, where rational forecast bias emerges as an outcome of a trade-off between public and private information while minimizing forecast squared error. We show both analytically and empirically that this trade-off and the resulting rational deviation are determined by experience and information asymmetry. While the optimal bias and information asymmetry are monotonically positively related, the optimal bias and experience are concavely linked. Moderately experienced analysts find it optimal to purchase private information with a positively biased deviation from the consensus. The superior analytical skills and access to private information (at a lower cost) of highly seasoned analysts leads them to optimally rely on private information.

We use Reg. FD as an experiment to further illuminate the relationship between experience and forecast bias and to make a distinction between access to private information and analytical expertise. We find that, as we would expect, the biases of experienced analysts, who weigh private information more heavily, are affected more. The extent to which information asymmetry and analyst experience play a role in determining the analyst's bias is directly affected by the dispersion of the common and idiosyncratic signals of all analysts as well as the extent of their correlation with each other.

Our results may help to paint a clearer picture of the types of information produced by analysts. Institutions may be able to use these results to design continuing education programs and to make hiring/firing decisions. Our results can also help to explain why and how changes in the information environment caused by regulatory changes (e.g., Reg. FD), index inclusions, and public disclosures affect the role of analysts as information processors.

Appendix 1: Proofs

The Objective Function

Given that the analyst's forecast is a weighted average of the analyst's unconditional estimate and the consensus, $F = wE + (1 - w)E_c$, the objective function can be expressed as

$$\min_{w|b}\left[w^2(\tau_0+\tau(b))^{-2}+(1-w)^2\tau_c^{-2}+2\rho w(1-w)(\tau_0+\tau(b))^{-1}\tau_c^{-1}\right] \tag{2.6}$$

The first-order condition then is

$$\begin{aligned}&2w(\tau_0+\tau(b))^{-2}-2(1-w)\tau_c^{-2}+2\rho(1-w)(\tau_0+\tau(b))^{-1}\tau_c^{-1}\\&-2\rho w(\tau_0+\tau(b))^{-1}\tau_c^{-1}\equiv 0\end{aligned}$$

By collecting terms, we then have

$$w\left\{(\tau_0+\tau(b))^{-2}+\tau_c^{-2}-2\rho(\tau_0+\tau(b))^{-1}\tau_c^{-1}\right\}=\tau_c^{-2}-\rho(\tau_0+\tau(b))^{-1}\tau_c^{-1}$$

This means that the optimal weight is

$$w=\frac{(\tau_0+\tau(b))^2-\rho\tau_c(\tau_0+\tau(b))}{(\tau_0+\tau(b))^2+\tau_c^2-2\rho\tau_c(\tau_0+\tau(b))} \tag{2.7}$$

Proof of Proposition 1

By taking the derivative of Eq. 2.7 with respect to τ_0, we have

$$\frac{\partial w}{\partial\tau_0}=\frac{2\tau_c^2(\tau_0+\tau(b))-2\rho\tau_c(\tau_0+\tau(b))^2-\rho\tau_c^3}{\left[(\tau_0+\tau(b))^2+\tau_c^2-2\rho\tau_c(\tau_0+\tau(b))\right]^2} \tag{2.8}$$

Clearly, since the denominator of $\partial w/\partial\tau_0$ is positive, then the sign is only a function of the numerator. This implies that the sign changes when the numerator, $2\tau_c(\tau_0 + \tau(b)) - 2\rho(\tau_0 + \tau(b))^2 - \rho\tau_c^2$, is at maximum. To find the maximum, we solve for τ_0 that satisfies the first-order conditions of the numerator. The first-order condition yields $\tau_c - 2\rho(\tau_0 + \tau(b)) \equiv 0$. Thus, at optimal weight $\tau_0 + \tau(b) = 0.5\rho^{-1}\tau_c$.

Proof of Proposition 2

By taking the derivative of Eq. 2.7 with respect to bias, we have

$$\frac{\partial w}{\partial b} = \frac{\left[2\tau_c^2(\tau_0 + \tau(b)) - 2\rho\tau_c(\tau_0 + \tau(b))^2 - \rho\tau_c^3\right]\frac{\partial \tau}{\partial b}}{\left[(\tau_0 + \tau(b))^2 + \tau_c^2 - 2\rho\tau_c(\tau_0 + \tau(b))\right]^2} \tag{2.9}$$

Clearly, since the denominator of $\partial w/\partial b$ is positive, then the sign is only a function of the numerator. This implies (1) that since $\partial \tau/\partial b$ is positive, then the optimal weight would be monotonically increasing with $\partial \tau/\partial b$ or information asymmetry, and (2) that the optimal weight is nonlinearly, concavely related to private information precision. Since the first term in the numerator is a quadratic function of analyst's own precision, the maximum in the function is the point at which the numerator changes sign. This point, however, is exactly the same point at which $\partial w/\partial \tau_0$ maximizes. For biases at which $\tau_0 + \tau(b)$ falls below $0.5\rho^{-1}\tau_c$., then so long as bias increases so does the optimal weight.

Appendix 2: Alternate Samples

See Table 2.6.

Appendix 3: Alternate Proxies for Information Asymmetry for Table 2.5

See Table 2.7.

Appendix 3 presents the results of regressing forecast error on the regressors specified in each column. Inf. asymmetry is analyst coverage in Panel A and relative firm size in Panel B. Same quarter is a dummy variable equal to one if the estimate is in the same quarter as the actual and zero otherwise. Same quarter is orthogonalized (on NumRevisions) to ensure that multicollinearity is not a problem between these two variables. Controls is a vector of firm-specific variables including broker reputation, broker size, accruals, intang. assets, and return std dev. Brokerage reputation is a ranking of brokerage reputation, where 0 is the worst and 9 is the best. Brokerage size is the natural log of the number of companies per quarter a brokerage house follows. Brokerage reputation is orthogonalized (on brokerage size) to ensure that multicollinearity is not a problem between these two variables. Accruals are the accrued revenue/liabilities utilized for earnings smoothing. Intang. assets are the covered firm's intangible assets value relative to its total assets. Return standard deviation is the standard deviation of the covered firm's return. I is a vector of one-digit SIC industry dummies. T is a vector of time dummies.

References

Agrawal, A., & Chen, M. (2012). Analyst conflicts and research quality. *Quarterly Journal of Finance, 2*, 179–218.

Baily, W., Li, H., Mao, C., & Zhong, R. (2003). Regulation fair disclosure and earnings information: Market, analysts, and corporate response. *Journal of Finance, 58*, 2487–2514.

Bannister, J., & Newman, H. (1996). Accrual usage to manage earnings toward financial analysts' forecasts. *Review of Quantitative Finance and Accounting, 7*, 259–278.

Barber, B., Lehavy, R., Trueman, B. (2000). *Are all brokerage houses created equal? Testing for systematic differences in the performance of brokerage house stock recommendations.* University of California at Davis and University of California at Berkeley (unpublished), March.

Bernhardt, D., Campello, M., & Kutsoati, E. (2006). Who herds? *Journal of Financial Economics, 80*, 657–675.

Beyer, A. (2008). Financial analysts' forecast revisions and managers' reporting behavior. *Journal of Accounting and Economics, 46*, 334–348.

Bhattacharya, N. (2001). Investors' trade size and trading responses around earnings announcements: An empirical investigation. *Accounting Review, 76*, 221–244.

Boni, L., & Womack, K. (2006). Analysts, industries, and price momentum. *Journal of Financial Quantitative Analysis, 41*, 85–109.

Brown, L., & Sivakumar, K. (2003). Comparing the quality of two operating income measures. *Review of Accounting Studies, 4*, 561–572.

Brown, S., Hillegeist, S., & Lo, K. (2004). Conference calls and information asymmetry. *Journal of Accounting and Economics, 37*, 343–366.

Carey, M., Post, M., & Sharp, S. (1998). Does corporate lending by banks and finance companies differ? Evidence on specialization in private debt contracting. *Journal of Finance, 53*, 845–878.

Carter, R., & Manaster, S. (1990). Initial public offerings and underwriter reputation. *Journal of Finance, 45*, 1045–1068.

Carter, R., Dark, F., & Singh, A. (1998). Underwriter reputation, initial returns, and the long-run performance of IPO stocks. *Journal of Finance, 53*, 285–311.

Chen, Q., & Jiang, W. (2006). Analysts' weighting of private and public information. *Review of Financial Studies, 19*, 319–355.

Clement, M. (1999). Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting and Economics, 27*, 285–303.

Clement, M., & Tse, S. (2005). Financial analyst characteristics and herding behavior in forecasting. *Journal of Finance, 60*, 307–341.

Conrad, J., Cornell, B., Landsman, W., & Rountree, B. (2006). How do analyst recommendations respond to major news? *Journal of Financial and Quantitative Analysis, 41*, 25–49.

De Jong, P., & Apilado, V. (2008). The changing relationship between earnings expectations and earnings for value and growth stocks during Reg FD. *Journal of Banking and Finance, 33*, 435–442.

Dechow, P., Kothari, S., & Watts, R. (1998). The relation between earnings and cash flows. *Journal of Accounting and Economics, 25*, 133–168.

Doyle, J., Lundholm, R., & Soliman, M. (2003). The predictive value of expenses excluded from pro forma earnings. *Review of Accounting Studies, 8*, 145–174.

Fama, E., & French, K. (1997). Industry costs of equity. *Journal of Financial Economics, 43*, 153–193.

Givoly, D., & Lakonishok, J. (1979). The information content of financial analysts' forecasts of earnings. *Journal of Accounting and Economics, 2*, 165–185.

Gu, Z., & Xue, J. (2007). Do analysts overreact to extreme good news in earnings? *Review of Quantitative Finance and Accounting, 29*, 415–431.

Han, B., Manry, D., & Shaw, W. (2001). Improving the precision of analysts' earnings forecasts by adjusting for predictable bias. *Review of Quantitative Finance and Accounting, 17*, 81–98.

Hirschey, M., & Richardson, V. (2004). Are scientific indicators of patent quality useful to investors? *Journal of Empirical Finance, 11*, 91–107.
Hong, H., Kubik, J., & Salomon, A. (2000). Security analysts' career concerns and herding of earnings forecasts. *RAND Journal of Economics, 31*, 121–144.
Hsu, D., & Chiao, C. (2010). Relative accuracy of analysts' earnings forecasts over time: A Markov chain analysis. *Review of Quantitative Finance and Accounting, 37*, 477–507.
Irani, A. (2004). The effect of regulation fair disclosure on the relevance of conference calls to financial analysts. *Review of Quantitative Finance and Accounting, 22*, 15–28.
Krishnaswami, S., & Subramaniam, V. (1998). Information asymmetry, valuation, and the corporate spin-off decision. *Journal of Financial Economics, 53*, 73–112.
Kwon, S. (2002). Financial analysts' forecast accuracy and dispersion: High-tech versus low-tech stocks. *Review of Quantitative Finance and Accounting, 19*, 65–91.
Leone, A., & Wu, J. (2007). *What does it take to become a superstar? Evidence from institutional investor rankings of financial analysts*. Simon School of Business Working Paper No. FR 02-12. Available at SSRN: http://ssrn.com/abstract=313594 or http://dx.doi.org/10.2139/ssrn.313594.
Lim, T. (2001). Rationality and analysts; forecast deviation. *Journal of Finance, 56*, 369–385.
Lin, B., & Yang, R. (2010). Does regulation fair disclosure affect analysts' forecast performance? The case of restructuring firms. *Review of Quantitative Finance and Accounting, 38*, 495–517.
Loughran, T., & Ritter, J. (2004). Why has IPO underpricing changed over time? *Financial Management, 33*(3), 5–37.
Mest, D., & Plummer, E. (2003). Analysts' rationality and forecast bias: Evidence from analysts' sales forecasts. *Review of Quantitative Finance and Accounting, 21*, 103–122.
Mikhail, M., Walther, D., & Willis, R. (1997). Do security analysts improve their performance with experience? *Journal of Accounting Research, 35*, 131–157.
Nutt, S., Easterwood, J., & Easterwood, C. (1999). New evidence on serial correlation in analyst forecast errors. *Financial Management, 28*, 106–117.
Petersen, M., & Rajan, R. (1994). The benefits of lending relationships: Evidence from small business data. *Journal of Finance, 49*, 3–37.
Ramnath, S. (2002). Investor and analyst reactions to earnings announcements of related firms: An empirical analysis. *Journal of Accounting Research, 40*, 1351–1376.
Stangeland, D., & Zheng, S. (2007). IPO underpricing, firm quality, and analyst forecasts. *Financial Management, 36*, 1–20.
Sufi, A. (2007). Information asymmetry and financing arrangements: Evidence from syndicated loans. *Journal of Finance, 62*, 629–668.
Thomas, S. (2002). Firm diversification and asymmetric information: Evidence from analysts' forecasts and earnings announcements. *Journal of Financial Economics, 64*, 373–396.
Trueman, B. (1990). Theories of earnings-announcement timing. *Journal of Accounting and Economics, 13*, 285–301.
Waymire, G. (1986). Additional evidence on the accuracy of analyst forecasts before and after voluntary managed earnings forecasts. *Accounting Review, 61*, 129–141.
Winkler, R. L. (1972). An introduction to Bayesian inference and decision. New York: Holt, Rinehart and Winston.
Zhang, F. (2006). Information uncertainty and stock returns. *Journal of Finance, 61*, 105–137.
Zhou, T., & Lai, R. (2009). Herding and information based trading. *Journal of Empirical Finance, 16*, 388–393.
Zitzewitz, E. (2002). *Regulation fair disclosure and the private information of analysts*. Available at SSRN: http://ssrn.com/abstract=305219 or http://dx.doi.org/10.2139/ssrn.305219.

An Appraisal of Modeling Dimensions for Performance Appraisal of Global Mutual Funds

3

G.V. Satya Sekhar

Contents

Abstract

A number of studies have been conducted to examine investment performance of mutual funds of the developed capital markets. Grinblatt and Titman (1989, 1994) found that small mutual funds perform better than large ones and that performance is negatively correlated to management fees but not to fund size or expenses. Hendricks, Patel, and Zeckhauser (1993), Goetzmann and Ibbotson (1994), and Brown and Goetzmann (1995) present evidence of persistence in mutual fund performance. Grinblatt and Titman (1992) and Elton, Gruber,

G.V. Satya Sekhar
Department of Finance, GITAM Institute of Management, GITAM University, Visakhapatnam, Andhra Pradesh, India
e-mail: gudimetlavss@yahoo.com

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_3,

and Blake (Journal of Financial Economics 42:397–421, 1996) show that past performance is a good predictor of future performance. Blake, Elton, and Grubber (1993), Detzler (1999), and Philpot, Hearth, Rimbey, and Schulman (1998) find that performance is negatively correlated to fund expense, and that past performance does not predict future performance. However, Philpot, Hearth, and Rimbey (2000) provide evidence of short-term performance persistence in high-yield bond mutual funds. In their studies of money market mutual funds, Domian and Reichenstein (1998) find that the expense ratio is the most important factor in explaining net return differences. Christoffersen (2001) shows that fee waivers matter to performance. Smith and Tito (1969) conducted a study into 38 funds for 1958–1967 and obtained similar results. Treyner (1965) advocated the use of beta coefficient instead of the total risk.

Keywords
Financial modeling • Mutual funds • Performance appraisal • Global investments • Evaluation of funds • Portfolio management • Systematic risk • Unsystematic risk • Risk-adjusted performance • Prediction of price movements

3.1 Introduction

Performance of financial instruments is basically dependent on three important models derived independently by Sharpe, Jensen, and Treynor. All three models are based on the assumptions that (1) all investors are averse to risk and are single-period expected utility of terminal wealth maximizers, (2) all investors have identical decision horizons and homogeneous expectations regarding investment opportunities, (3) all investors are able to choose among portfolios solely on the basis of expected returns and variance of returns, (4) all transactions costs and taxes are zero, and (5) all assets are infinitely divisible.

3.2 Performance Evaluation Methods

The following paragraphs indicate a brief description of the studies on "performance evaluation of mutual funds."

Friend et al. (1962) offered the first empirical analysis of mutual funds performance. Sharpe (1964), Treynor and Mazuy (1966), Jensen (1968), Fama (1972), and Grinblatt and Titman (1989, 1994) are considered to be classical studies in performance evaluation methods. Sharpe (1964) made a significant contribution in the methods of evaluating mutual funds. His measure is based on capital asset prices, market conditions with the help of risk and return probabilities. Sharpe (1966) developed a theoretical measure better known as reward to variability ratio that considers both average return and risk simultaneously in its ambit. It tested efficacy through a sample of 34 open-ended funds considering annual returns and standard deviation of annual return risk surrogate for the period for 1954–1963.

The average reward to variability ratio of 34 funds was considerably smaller than Dow Jones portfolio and considered enough to conclude that average mutual funds performance was distinctly inferior to an investment in Dow Jones Portfolio.

Treynor (1965) advocated the use of beta coefficient instead of the total risk. He argues that using only naïve diversification, the unsystematic variability of returns of the individual assets in a portfolio typically average out of zero. So he considers measuring a portfolio's return relative to its systematic risk more appropriate.

Treynor and Mazuy (1966) devised a test of ability of the investment managers to anticipate market movements. The study used the investment performance outcomes of 57 investment managers to find out evidence of market timing abilities and found no statistical evidence that the investment managers of any of the sample funds had successfully outguessed the market. The study exhibited that the investment managers had no ability to outguess the market as a whole but they could identify under priced securities.

Michael C. Jensen (1967) conducted an empirical study of mutual funds during the period 1954–1964 for 115 mutual funds. His results indicate that these funds are not able to predict security prices well enough to outperform a buy-the-market-and-hold policy. His study ignores the gross management expenses to be free. There was very little evidence that any individual fund was able to do significantly better than which investors expected from mere random chance. Jensen (1968) measured the performance as the return in excess of equilibrium return mandated by capital asset pricing model. Jensen's measure is based on the theory of the pricing of capital assets by Sharpe (1964), Linter (1965), and Treynor.

Smith and Tito (1969) conducted a study into 38 funds for 1958–1967 and published results relating to performance of mutual funds. However, Mc Donald (1974) examined 123 mutual funds for 1960–1969 measures to be closely correlated; more importantly, he found that on an average, mutual funds perform about as well as native "buy and hold" strategy.

Fama (1972) suggested alternative methods for evaluating investment performance with somewhat finer breakdowns of performance on the stock selection, market timing, diversification, and risk bearing. It devised mechanism for segregation part of an observed investment return due to managers' ability to pick up the best securities at a given level of risk from part that is due to the prediction of general market price movements.

Dunn and Theisen (1983) study is about ranking by the annual performance of 201 institutional portfolios for the period 1973 through 1982 without controlling for fund risk. They found no evidence that funds performed within the same quartile over the 10-year period. They also found that ranks of individual managers based on 5-year compound returns revealed no consistency.

Eun et al. (1991) reported similar findings. The benchmarks used in their study were the Standard and Poor's 500 Index, the Morgan Stanley Capital International World Index, and a self-constructed index of US multinational firms. For the period 1977–1986, the majority of international funds outperformed the US market.

However, they mostly failed to outperform the world index. The sample consisted of 19 US-based international funds, and the Sharpe measure was used to assess excess returns.

Barua and Varma (1993b) have examined the relationship between the NAV and the market price on Mastershares. They conclude that market prices are far more volatile than what can be justified by volatility of NAVs. The prices also show a mean reverting behavior, thus perhaps providing an opportunity for discovering a trading rule to make abnormal profits in the market. Such a rule would basically imply buying Mastershares whenever the discount from NAV was quite high and selling Mastershares whenever the discount was low.

Droms and Walker (1994) used a cross-sectional/time-series regression methodology. Four funds were examined over 20 years (1971–1990), and 30 funds were analyzed for a 6-year period (1985–1990). The funds were compared to the Standard and Poor's 500 Index, the Morgan Stanley Europe, Australia, and Far East Index (EAFE) which proxies non-US stock markets, and the World Index. Applying the Jensen, Sharpe, and Treynor indices of performance, they found that international funds have generally underperformed the US market and the international market. Additionally, their results indicated that portfolio turnover, expense ratios, asset size, load status, and fund size are unrelated to fund performance.

Bauman and Miller (1995) studied the persistence of pension and investment fund performance by type of investment organization and investment style. They employed a quartile ranking technique, because they noted that "investors pay particular attention to consultants' and financial periodicals' investment performance rankings of mutual funds and pension funds." They found that portfolios managed by investment advisors showed more consistent performance (measured by quartile rankings) over market cycles and that funds managed by banks and insurance companies showed the least consistency. They suggest that this result may be caused by a higher turnover in the decision-making structure in these less consistent funds. This study controls for the effects of turnover of key decision makers by restricting the sample to those funds with the same manager for the entire period of study.

Volkman and Wohar (1995) extend this analysis to examine factors that impact performance persistence. Their data consists of 322 funds over the period 1980–1989 and shows performance persistence is negatively related to size and negatively related to levels of management fees.

Elton et al. (1996) examined the predictability of stock mutual funds performance based on risk-adjusted future performance. It also demonstrated application of modern portfolio techniques on past data to improve selection, which permitted construction of portfolio funds that significantly outperformed a rule based on the past rank alone. The portfolio so selected was reported to have small, but statistically significant, positive risk-adjusted returns during a period when mutual funds in general had negative risk-adjusted returns.

Jayadeve (1996) paper enlightens performance evaluation based on monthly returns. His paper focuses on performance of two growth-oriented mutual funds

(Mastergain and Magnum Express) on the basis of monthly returns compared to benchmark returns. For this purpose, risk-adjusted performance measures suggested by Jensen and Treynor and Sharpe are employed.

Carhart (1997) shows that expenses and common factors in stock returns such as beta, market capitalization, 1-year return momentum, and whether the portfolio is value or growth oriented "almost completely" explain short-term persistence in risk-adjusted returns. He concludes that his evidence does not "support the existence of skilled or informed mutual fund portfolio managers."

Yuxing Yan (1999) examined performance of 67 US mutual funds and the S&P 500 Index with 10-year daily return data from 1982 to 1992. The S&P index was used as benchmark index. Daily data are transformed into weekly data for computational reasons. In the calculations, it was assumed that the S&P 500 market index is a good one, i.e., it is efficient and its variance is constant.

Redmand et al.'s (2000) study examines the risk-adjusted returns using Sharpe's Index, Treynor's Index, and Jensen's alpha for five portfolios of international mutual funds during 1985–1994. The benchmarks for competition were the US market proxied by the Vanguard Index 500 mutual fund and a portfolio of funds that invest solely in US stocks. The results show that for 1985 through 1994 the portfolio of international mutual funds outperformed the US market and the portfolio of US mutual funds.

Rahul Bhargava et al. (2001) evaluated the performance of 114 international equity managers over the January 1988 to December 1997 period. Performance tests are conducted using Sharpe and Jensen performance methodologies. Three major findings are reported. First, international equity managers, on an average, were unable to outperform the MSCI world market proxy during the sample period. Second, geographic asset allocation and equity style allocation decisions enhanced the performance of international managers during the sample period. Third, separately managed funds were outperformed mutual funds.

Sadhak's (2003) study is an attempt to evaluate the performance of Indian mutual funds with the help of data pertaining to (a) trends in income and expenses, (b) investment yield and risk-associated returns, and (c) returns of Indian mutual funds vis-à-vis returns of other emerging markets.

Bala Ramasamy and Yeung's (2003) survey focused on Malaysia where the mutual fund industry started in the 1950s but only gained importance in the 1980s with the establishment of government-initiated program. The sample size consisting of 56 financial advisors representing various life insurance and mutual fund companies resulted in 864 different profiles of mutual funds. The cojoint analysis was employed to generate the questionnaire and analyze its results. The results of this survey point to three important factors which dominate the choice of mutual funds. These are consistent past performance, size of funds, and costs of transaction.

Chang et al. (2003) identified hedging factor in the equilibrium asset pricing model and used this benchmark to construct a new performance measure. Based on this measure, they are able to evaluate mutual fund managers hedging timing ability in addition to more traditional security selectivity and timing. While security selectivity performance involves forecasts of price movements of selected individual stock, market timing measures the forecasts of next period realizations of the market portfolio. The empirical evidence indicates that the selectivity measure is positive on average and the market timing measure is negative on average.

Obeid (2004) has suggested a new dimension called "modified approach for risk-adjusted performance of mutual funds." This method can be considered as more powerful, because it allows not only for an identification of active resources but also for identification of risk. He observed two interesting results: first, it can be shown that in some cases, a superior security selection effect is largely dependent on taking higher risks. Second, even in the small sample analyzed in the study, significant differences appear between each portfolio manager's styles of selection.

Gupta OP and Amitabh Gupta (2004) published their research on select Indian mutual funds during a 4-year period from 1999 to 2003 using weekly returns based on NAVs for 57 funds. They found that fund managers have not outperformed the relevant benchmark during the study period. The funds earned an average return of 0.041 per week against the average market return of 0.035 %. The average risk-free rate was 0.15 % per week, indicating that the sample funds have not earned even equivalent to risk-free return during the study period.

Subash Chander and Japal Singh (2004) considered selected funds during the period from November 1993 to March 2003 for the purpose of their study. It was found that the Alliance Mutual Fund and Prudential ICICI Mutual Funds have posted better performance for the period of study in that order as compared to other funds. Pioneer ITI, however, has shown average performance and Templeton India mutual fund has staged a poor show.

Amit Singh Sisodiya (2004) makes comparative analysis of performance of different mutual funds. He explains that a fund's performance when viewed on the basis of returns alone would not give a true picture about the risk the fund would have taken. Hence, a comparison of risk-adjusted return is the criteria for analysis.

Bertoni et al. (2005) analyzed the passive role that, implicitly, would place institutional investors in such a context. The study was conducted in Italy using empirical evidence from the Italian stock exchange (Comit Index). This study finds that three factors reduce the freedom of institutional investors to manage their portfolio – the market target size, the fund structure, and the benchmarking.

Sudhakar and Sasi Kumar (2005) made a case study of Franklin Templeton mutual fund. The sample consists of a total of ten growth-oriented mutual funds during the period from April 2004 to March 2005. NIFTY based on NSE Index was used as the proxy for the market index, and each scheme is evaluated with respect to the NSE index to find out whether the schemes were able to beat the market or not. It was found that most of the growth-oriented mutual funds have been able to deliver better returns than the benchmark indicators. In the sample study, all the funds have positive differential returns indicating better performance and diversification of the portfolio, except two funds with negative differential returns, viz., Franklin India Bluechip Fund and Templeton India Income Fund.

Martin Eling (2006) made a remarkable contribution to the theory of "performance evaluation measures." In this study, data envelopment analysis (DEA) is presented as an alternative method for hedge fund performance measurement. As an optimization result, DEA determines an efficiency score, which can be interpreted as a performance measure. An important result of the empirical study is that completely new rankings of hedge funds compared to classic performance measures.

George Comer (2006) examined the stock market timing ability of two samples of hybrid mutual funds. The results indicate that the inclusion of bond indices and a bond timing variable in a multifactor Treynor-Mazuy model framework leads to substantially different conclusion concerning the stock market timing performance of these funds relative to the traditional Treynor-Mazuy model find less stock timing ability over the 1981–1991 time period provide evidence of significant stock timing ability across the second fund sample during the 1999–2000 period.

Yoon K. Choi (2006) proposed an incentive-compatible portfolio performance evaluation measure. In this model, a risk-averse portfolio manager is delegated to manage a fund, and his portfolio construction (and information-gathering) effort is not directly observable to investors, in which managers are to maximize investors' gross returns net of managerial compensation. He considers the effect of organizational elements such as economics of scale on incentive and thus on performance.

Ramesh Chander (2006) study examined the investment performance of managed portfolios with regard to sustainability of such performance in relation to fund characteristics, parameter stationarity, and benchmark consistency. The study under consideration is based on the performance outcome of 80 investment schemes from public as well as private sectors for the 5-year period encompassing January 1998 through December 2002. The sample comprised 33.75 % of small, 26.75 % of medium, 21.25 % of large, and 18.75 % of the giant funds.

Ramesh Chander (2006a) study on market timing abilities enables us to understand how well the manager has been able to achieve investment targets and how well risk has been controlled in the process. The results reported were unable to generate adequate statistical evidence in support of manager's successful market timing. It persisted across measurement criteria, fund characteristics, and the benchmark indices. However, absence of performance is noted for alternative sub-periods signifying the negation of survivorship bias.

Beckmann et al. (2007) found that Italian female professionals do not only assess themselves as more risk averse than their male colleagues, they also prefer a more passive portfolio management compared to the level they are allowed to. Besides, in a competitive tournament scenario near the end of the investment period, female asset managers do not try to become the ultimate top performer when they have outperformed the peer group. However in case of underperformance, the risk of deviating from the benchmark makes female professionals more willing than their male colleagues to seize a chance of catching up.

Gajendra Sidana (2007) made an attempt to classify hundreds of mutual funds employing cluster analysis and using a host of criteria like the 1-year-old return, 2-year annualized return, 3-year annualized return, 5-year annualized return, alpha, and beta. The data is obtained from value research. The author finds inconsistencies between investment style/objective classification and the return obtained by the fund.

Coates and Hubbard (2007) reviewed the structure, performance, and dynamics of the mutual fund industry and showed that they are consistent with competition. It was also found that concentration and barriers to entry are low, actual entry is common and continuous, pricing exhibits no dominant long-term trend, and market shares fluctuate significantly. Their study also focused on "effects of competition on fee" and "pricing anomalies." They suggested legal interventions are necessary in setting fee in mutual funds of United States.

Subha and Bharati's (2007) study is carried out for open-ended mutual fund schemes and 51 schemes are selected by convenient sampling method. NAVs are taken for a period of 1 year from 1 October 2004 to 30 September 2005. Out of the 51 funds, as many as 18 schemes earned higher returns than the market return. The remaining 33 funds however generated lower returns than the market.

Sondhi's (2007) study analyzes the financial performance of 36 diversified equity mutual funds in India, in terms of rates of return, comparison with risk-free return, benchmark comparison, and risk-adjusted returns of diversified equity funds. Fund size, ownership pattern of AMC, and type of fund are the main factors considered in this study. The study reveals that private sector is dominating public sector.

Cheng-Ru Wu et al.'s (2008) study adopts modified Delphi method and the analytical hierarchy process to design an assessment method for evaluating mutual fund performance. The most important criteria for mutual fund performance should be "mutual fund style" followed by "market investment environment." This result indicates investor's focus when they evaluate the mutual fund performance.

Eleni Thanou's (2008) study examines the risk-adjusted overall performance of 17 Greek Equity Mutual Funds between the years 1997 and 2005. The study evaluated performance of each fund based on the CAPM performance methodology, calculating the Treynor and Sharpe Indexes for the 9-year period as well as for three sub-periods displaying different market characteristics. The results indicated that the majority of the funds under examination followed closely the market, achieved overall satisfactory diversification, and some consistently outperformed the market, while the results in market timing are mixed, with most funds displaying negative market timing capabilities.

Kajshmi et al. (2008) studied a sample of schemes in the 8-year period. This study considers performance evaluation and is restricted to the schemes launched in the year 1993 when the industry was thrown open to private sector under the regulated environment by passing the SEBI (Mutual Funds) Regulations 1993. The performance of the sample schemes were in line with that of the market as evident from the positive beta values. All the sample schemes were not well diversified as depicted by the differences in the Jensen alpha and Sharpe's differential return.

Massimo Masa and Lei Zhang (2008) found the importance of organizational structure on Asset Management Company of mutual fund. Their study found that more hierarchical structures invest less in firms located close to them and deliver lower performance. An additional layer in hierarchical structure reduces the average performance by 24 basis points per month. At the same time, more hierarchical structures leads to herd more and to hold less concentrated portfolios.

Manuel Ammann and Michael Verhofen (2008) examined the impact of prior performance on the risk-taking behavior of mutual fund managers. Their sample taken from US funds started in January 2001 and ended in December 2005. The study found that prior performance in the first half of the year has, in general, a positive impact on the choice of the risk level in the second half of the year. Successful fund managers increase the volatility and the beta and assign a higher proportion of their portfolio to value stocks, small firms, and momentum stocks in comparison to unsuccessful fund managers.

Onur et al. (2008) study evaluates the performance of 50 large US-based international equity funds using risk-adjusted returns during 1994–2003. This study provides documentation on the risk-adjusted performance of international mutual funds. The evaluation is based on objective performance measures grounded in modern portfolio theory. Using the methodology developed by Modigliani and Miller in 1997, the study reports the returns that would have accrued to these mutual funds for a 5-year holding period as well as a 10-year holding period. It is evident from the empirical results of this study that the funds with the highest average returns may lose their attractiveness to investors once the degree of risk embedded in the fund has been factored into the analysis.

Qiang Bu and Nelson Lacey (2008) examined the determinants of US mutual fund terminations and provided estimates of mutual fund hazard functions. Their study found that mutual fund termination correlates with a variety of fund-specific variables as well as with market variables such as the S&P 500 Index and the short-term interest rate. This was tested with the underlying assumptions of the semi-parametric Cox model and reject proportionality. They also found that different fund categories exhibit distinct hazard functions depending on the fund's investment objectives.

David M. Smith (2009) discussed the size and market concentration of the mutual fund industry, the market entry and exit of mutual funds, the benefits and costs of mutual fund size changes, the principal benefits and costs of ownership from fund shareholders' perspective, etc. This study is based on data from Morningstar (2009) about US mutual fund industry, which was composed of 607 fund families.

Baker et al. (2010) investigated the relation between the performance and characteristics of 118 domestic actively managed institutional equity mutual funds. The results showed that the large funds tend to perform better, which suggests the presence of significant economies of scale. The evidence indicates a positive relation between cash holding and performance. They also found evidence in a univariate analysis that expense ratio class is an important determinant of performance, and the results are significant in a multivariate setting using Miller's active alpha as a performance metric.

Khurshid et al. (2009) studied the structure of the mutual fund industry in India and analyzed the state of competition among all the mutual funds in private sector and public sector. The levels of competition and their trends have been obtained for the periods March 2003–March 2009. This study found overall mutual fund industry is facing a high competitive environment. An increasing trend of competition was observed within bank institution, private sector foreign, and private sector joint venture mutual funds.

Mohit Gupta and Aggarwal's (2009) study focused on the portfolio creation and industry concentration of 18 ELSS schemes during April 2006 to April 2007. Mutual fund industry concentration was the variable used in classification or cluster creation. This exercise was repeated each month for the period under study. Finally portfolio performance was compared with index fund, portfolio of three randomly picked funds of the previous month, and the return and risk parameters of ELSS category as a whole.

Talat Afza and Ali Rauf's (2009) study aims to provide guidelines to the managers of open-ended Pakistani mutual funds and benefit small investors by pointing out the significant variables influencing the fund performance. An effort has been made to measure the fund performance by using Sharpe ratio with the help of pooled time-series and cross-sectional data and focusing on different fund attributes such as fund size, expenses, age, turnover, loads, and liquidity. The quarterly sample data are collected for all the open-ended mutual funds listed on Mutual Fund Association of Pakistan (MUFAP), for the years 1999–2006. The results indicate that among various funds attributes are: lagged return, liquidity and had significant impact on fund performance.

Amar Ranu and Depali Ranu (2010) critically examined the performance of equity funds and found out the top 10 best performing funds among 256 equity mutual fund schemes in this category. They considered three factors for selection: (a) mutual funds having 5 years of historical performance, (b) fund schemes having a minimum of Rs.400 crore of assets under management, and (c) funds which have average return more than 22.47. They found that HDFC TOP 200 (Growth) option was outperforming among the top 10 best performing equity funds.

Sunil Wahal and Albert Wang (2010) found impact of the entry of new mutual funds on incumbents using the overlap in their portfolio holdings as a measure of competitive intensity. Their study revealed that funds with high overlap also experience quantity competition through lower investor flows, have lower alphas, and higher attrition rates. These effects only appeared after the late 1990s, at which point there appears to be endogenous structural shift in the competitive environment. Their concluding remark is that "the mutual fund market has evolved into one that displays the hallmark features of a competitive market."

Sukhwinder Kaur Dhanda et al.'s (2012) study considered the BSE-30 as a benchmark to study the performance of mutual funds in India. The study period has been taken from 1 April 2009 to 31 March 2011. The findings of the study reveal that only three schemes have performed better than benchmark. In the year 2009, HDFC Capital Builder has the top performer. It was 69.18 returns and 26.37 SD and 0.78 beta. HDFC Capital Builder scheme has given the reward for variability and volatility. HDFC Top 200 Fund and Birla Sun Life Advantage Funds are on second and third position in terms of return. HDFC Top 200 Fund has shown better performance than Birla Sun Life Advantage Fund in terms of SD, beta, Sharpe ratio, and Treynor ratio. Birla Sun Life Advantage Fund has more risk than the benchmark. Kotak Select Focus Fund has the poorer performer in terms of risk and return. Except two schemes all other schemes have performed better than benchmark. Except Kotak Select Focus Fund all other schemes are able to give reward for variability and volatility.

3.3 A Review on Various Models for Performance Evaluation

3.3.1 Jensen Model

Given the additional assumption that the capital market is in equilibrium, all three models yield the following expression for the expected one-period return on any security (or portfolio) *j*:

$$\mathrm{E}\left(R_j\right) = R_F + \beta_J[\mathrm{E}(R_m) - R_F] \tag{3.1}$$

$R_F =$ the one-period risk-free interest rate.

$\beta_J = \mathrm{Cov}(j\,R_J, R_M)/\sigma^2\,R_M =$ the measure of risk (hereafter called systematic risk) which the asset pricing model implies is crucial in determining the prices of risky assets.

$E(R_M)$ = the expected one-period return on the “market portfolio” which consists of an investment in each asset in the market in proportion to its fraction of the total value of all assets in the market. It implies that the expected return on any asset is equal to the risk-free rate plus a risk premium given by the product of the systematic risk of the asset and the risk premium on the market portfolio.

3.3.2 Fama Model

In Fama’s decomposition performance evaluation measure of portfolio, overall performance can be attributed to selectivity and risk. The performance due to selectivity is decomposed into net selectivity and diversification. The difference between actual return and risk-free return indicates overall performance:

$$\mathrm{R_p} - \mathrm{R_f} \tag{3.2}$$

wherein

R_p is actually return on the portfolio, which is monthly average return of fund and

R_f is monthly average return on treasury bills 91 days.

The overall performance further can be bifurcated into performance due to selectivity and risk.

Thus,

$$\mathrm{R}_p - \mathrm{R}_f = \left[R_p - R_p\left(\beta_p + R_p\left(\beta_p - R_f\right)\right]\right. \tag{3.3}$$

In other words, overall performance = selectivity + risk

3.3.3 Treynor and Mazuy Model

Treynor and Mazuy developed a prudent and exclusive model to measure investment managers' market timing abilities. This formulation is obtained by adding squared extra return in the excess return version of the capital asset pricing model as given below:

$$\left(R_{pt} - R_{ft}\right) = \alpha + \beta_p\left(R_{mt} - R_{ft}\right) + yp\left(R_{mt} - R_{ft}\right)^{2+} e_{pt} \tag{3.4}$$

where R_{pt} is monthly return on the fund, R_{ft} is monthly return on 91 days treasury bills, R_{mt} is monthly return on market index, and E_{pt} is error term.

This model involves running a regression with excess investment return as dependent variable and the excess market return and squared excess market return as independent variables. The value of coefficient of squared excess return acts as a measure of market timing abilities that has been tested for significance of using *t*-test. Significant and positive values provide evidence in support of the investment manager's successful market timing abilities.

3.3.4 Statman Model

Statman measured mutual funds using the following equation (Statman 2000):

eSDAR (excess standard deviation and adjusted return)

$$= R_f + \left(R_p - R_f\right)\left(S_m/S_p\right) - R_m \tag{3.5}$$

In this formulae, R_f = monthly return on 3-month treasury bills, R_p = monthly return on fund portfolio, R_m = monthly return on the benchmark index, S_p = standard deviation of portfolio *p*'s return, and S_m = standard deviation of return on the benchmark index.

This model is used for short-term investment analysis. The performance is compared with it benchmark on monthly basis.

3.3.5 Choi Model

Choi provides a theoretical foundation for an alternative portfolio performance measure that is incentive-compatible. In this model, a risk-averse portfolio manager is delegated to manage a fund, and his portfolio construction (and information-gathering) effort is not directly observable to investors. The fund manager is paid on the basis of the portfolio return that is a function of effort, managerial skill, and organizational factors. In this model, the effect of institutional factors is described by the incentive contractual form and disutility (or cost) function of managerial efforts in fund operations. It focuses on the cost function as an organizational factor (simply, scale factor). It was assumed that the disutility function of each fund is determined by the unique nature of its operation (e.g., fund size) and is an increasing function of managerial effort at an increasing rate.

3.3.6 Elango Model

Elango's model also compares the performance of public sector funds vs private sector mutual funds in India. In order to examine the trend in performance of NAV during the study period, growth rate in NAV was computed. The growth rate was computed based on the following formula (Elango 2003):

$$\text{Growth rate: } R_g = (Y_t - Y_0/Y_0) \times 100 \tag{3.6}$$

R_g: growth rate registered during the current year
Y_t: yield in current year
Y_0: yield in previous year

In order to examine whether past is any indicator of future growth in the NAV, six regression analyses were carried out. NAV of base year was considered as the dependent variable and current year as in the independent variable.

$$\text{Equation: } Y = \mathrm{A} + \mathrm{b}\,X \tag{3.7}$$

Dependent variable: $Y =$ NAV of 1999–2000
Independent variable: $X =$ NAV of 2000–2001

In the same way, the second regression equation computed using NAVs of 2000–2001 and 2001–2002, as dependent and independent variables.

3.3.7 Chang, Hung, and Lee Model

The pricing model adopted by Jow-Ran Chang, Nao-Wei Hung, and Cheng-Few Lee is based on competitive equilibrium version of intemporal asset pricing model derived in Campbell. The dynamic asset pricing model incorporates hedging risk as well as market. This model uses a log-linear approximation to the budget constraint to substitute out consumption from a standard intertemporal asset pricing model. Therefore, asset risk premia are determined by the covariances of asset returns with the market return and with news about the discounted value of all future market returns. Formally, the pricing restrictions on asset i imported by the conditional version of the model are

$$E_t r_{i,\,t+1} - rf_{\,,t+1} = -V_{iI}/2 + \gamma V_{im} + (\gamma - 1)V_{ih} \tag{3.8}$$

where
$E_t r_{i,\,t+1}$, log return on asset; $rf,_{t+1}$, log return on riskless asset; V_{ii} denotes $\mathrm{Var}_t\,(r_{i,t+1})$; γ is the agent's coefficient of relative risk aversion; V_{im} denotes $\mathrm{Cov}_t\,(r_{i,\,t+1}, r_{m,t+1})$ and $V_{ih} = \mathrm{Cov}_t\,(r_{i,t+1}, (E_{t+1} - E_t), \_\infty j = 1\ \rho_j\ r_{m,t+1+j})$; the parameter, $\rho = 1 - \exp(c - w)$; and $c - w$ is the mean log consumption to wealth ratio.

This states that the expected excess log return in an asset, adjusted for a Jensen's inequality effect, is a weighted average of two covariances: the covariance with the return from the market portfolio and the covariance with news about future returns on invested wealth. The intuition in this equation that assets are priced using their covariances with the return on invested wealth and future returns on invested wealth.

3.3.8 MM Approach

Leah Modigliani and Franco Modigliani are better known as M^2 in the investment literature. This measure is developed adjusting portfolio return. This adjustment is carried on the uncommitted (cash balances) part of the investment portfolio at the riskless return so as to enable all portfolio holdings to participate in the return generation process. This adjustment is needed to bring out the level playing field for portfolio risk-return and vis-à-vis market return. The effect of this adjustment is reported below (Modigliani and Modigliani 1997):

$$\mathrm{M}^2 = {}^*\mathrm{Rp} - \mathrm{Rm} \tag{3.9}$$

$$^*\mathrm{Rp} = (\mathrm{Rf}^*(1 - \mathrm{Sdm}/\mathrm{Sdp})) + (\mathrm{Rp}^*\mathrm{Sdm}/\mathrm{Sdp}) \tag{3.10}$$

In this formulae * Rp = expected return, Rf = risk-free return, Sdm = standard deviation of market portfolio, and Sdp = standard deviation of managed portfolio.

In case the managed portfolio has twice the standard deviation of the market, then, the portfolio would be half invested in the managed portfolio and the remaining half would be invested at the riskless rate. Likewise, in case the managed portfolio has lower standard deviation than the market portfolio, it would be levered by borrowing money and investing the money in managed portfolio. Positive M^2 value indicates superior portfolio performance, while negative indicates actively managed portfolio manager's inability to beat the benchmark portfolio performance.

3.3.9 Meijun Qian's Stage Pricing Model

Meijun Qian's (2009) study reveals about the staleness, which is measured prices imparts a positive statistical bias and a negative dilution effect on mutual fund performance. First, evaluating performance with non-synchronous data generates

Table 3.1 Overview of different measures

Measures	Description	Interpretation
Sharpe ratio	Sharpe ratio = fund return in excess of risk-free return/standard deviation of fund. Sharpe ratios are ideal for comparing funds that have a mixed asset classes	The higher the Sharpe ratio, the better the fund returns relative to the amount of risk taken
Treynor ratio	Treynor ratio = fund return in excess of risk-free return/beta of fund. Treynor ratio indicates relative measure of market risk	The higher the Treynor ratio shows higher returns and lesser market risk of the fund
Jensen measure	This shows relative ratio between alpha and beta	Jensen measure is based on systematic risk. It is also suitable for evaluating a portfolio's performance in combination with other portfolios
M^2 measure	It matches the risk of the market portfolio and then calculate appropriate return for that portfolio	A high value indicates that the portfolio has outperformed and vice versa
Jensen model	$E(R_j) = R_F + \beta_J[E(R_m) - R_F]$	The expected one-period return on the "market portfolio" which consists of an investment in each asset in the market in proportion to its fraction of the total value of all assets in the market
Fama model	$R_p – R_f = [R_p - R_p(\beta_p + R_p(\beta_p – R_f)]$	Overall performance = selectivity + risk
Treynor and Mazuy model	$(R_{pt} – R_{ft}) = \alpha + \beta_p\ (R_{mt} - R_{ft}) + yp\ (R_{mt} - R_{ft})^{2+} e_{pt}$	This model involves running a regression with excess investment return as dependent variable and the excess market return and squared excess market return as independent variables
Statman model	eSDAR $= R_f + (R_p - R_f)(S_m/S_p) – R_m$	This model used for short-term investment analysis. The performance is compared with it benchmark on monthly basis
Elango model	$R_g = (Y_t - Y_0/Y_0) \times 100$	In order to examine whether past is any indicator of future growth in the NAV, six regression analyses were carried out. NAV of base year was considered as the dependent variable and current year as in the independent variable

a spurious component of alpha. Second, stale prices create arbitrage opportunities for high-frequency traders whose trades dilute the portfolio returns and hence fund performance. This paper introduces a model that evaluates fund performance while controlling directly for these biases. Empirical tests of the model show that alpha net of these biases is on average positive although not significant and about 40 basis points higher than alpha measured without controlling for the impacts of stale pricing. The difference between the net alpha and the measured alpha consists of three components: a statistical bias, the dilution effect of long-term fund flows, and the dilution effect of arbitrage flows. Thus, assuming that information generated in time t is not fully incorporated into prices until one period later, the observed fund return becomes a weighted average of true returns in the current and last periods:

$$r_t = \alpha + \beta r_{mt} + \varepsilon_t, \tag{3.11}$$

$$r_t^* = \eta r_{t-1} + (1 - \eta) r_t, \tag{3.12}$$

where r_t denotes the true excess return of the portfolio with mean μ and variance σ_2 and r_{mt} denotes the excess market return with mean μ_m and variance σ_m. Both r_t and r_{mt} are i.i.d, and the error term ε_t is independent of r_{mt}. R_t* is the observed excess return of the portfolio with zero flows, while η is the weight on the lagged true return. That is, the higher the η, the staler the prices. Assumedly, arbitrage traders can earn the return r_t*, by trading at the fund's reported net assets values (Table 3.1).

3.4 Conclusion

This paper is intended to examine various performance models derived by financial experts across the globe. A number of studies have been conducted to examine investment performance of mutual funds of the developed capital markets. The measure of performance of financial instruments is basically dependent on three important models derived independently by Sharpe, Jensen, and Treynor. All three models are based on the assumption that (1) all investors are averse to risk and are single-period expected utility of terminal wealth maximizers, (2) all investors have identical decision horizons and homogeneous expectations regarding investment opportunities, (3) all investors are able to choose among portfolios solely on the basis of expected returns and variance of returns, (4) all transactions costs and taxes are zero, and (5) all assets are infinitely divisible.

References

Afza, T., & Rauf, A. (2009). Performance Evaluation of Pakistani mutual funds. *Pakistan Economic and Social Review, 47*(2), 199–214.

Ammann, M., & Verhofen, M. (2008). The impact of prior performance on the risk-taking of mutual fund manager. *Annals of Finance*, (5), 69–90.

Arugaslan, O., Edwards, E., & Samant, A. (2008). Risk-adjusted performance of international mutual funds. *Managerial Finance, 34*(1), 5–22.

Baker, M., Litov, L., Wachter, J., & Wurgler, J. (2010). Can mutual fund managers pick stocks? Evidence from their trades prior to earning announcements. *Journal of Financial and Quantitative Analysis, 45*, 1111–1131.

Baker Kent, H., Haslem, J. A., & Smith, D. M. *Performance and Characteristics of Actively Managed Institutional Equity Mutual Funds* [Electronic Copy]. http://ssrn.com/abstract=1124577

Barua, S. K., & Varma, J. R. (1993). *Speculative dynamics: The case of master shares* (Advances in financial planning and forecasting, Vol. 5). Greenwitch: Jai Press.
Bauman, W. S., & Miller, R. E. (1995). Portfolio performance rankings in stock market cycles. *Financial Analysts Journal, 51*, 79–87.
Beckmann, D., Lutje, T., & Rebeggiani, L. (2007). *Italian asset managers' behavior: Evidence on overconfidence, risk taking, and gender* (Discussion Paper No. 358). Hannover: Leibniz Universitat Hannover, Department of Economics, ISSN: 0949-9962.
Bertoni, A., Brtmetti, G., & Cesari, C. (2005). Mutual fund bench marking and market bubbles-a behavioural approach. *Transition Studies Review, 12*(1), 36–43.
Bhargava, R., Gallo, J., & Swason, P. T. (2001). The performance, asset allocation and investment style of international equity manager. *Review of Quantitative Finance and Planning, 17*, 377–395.
Born, E. K. (1983). *International banking in the 19th and 20th centuries*. New York: St. Martin's Press.
Brown, S. J., & Goetzmann, W. N. (1995). Performance persistence. *The Journal of Finance, 50*(2), 679–698.
Bu, Q., & Lacey, N. (2008). On understanding mutual fund terminations. *Journal of Economics and Finance, 33*, 80–99.
Carhart, M. M. (1997). Persistence in mutual fund performance. *Journal of Finance, 52*, 57–82.
Chander, R. (2006). Informational efficiency, parameter stationarity and bench mark consistency of investment performance. *The ICFAI Journal of Applied Finance*. March 2006, p. 29–45.
Chander, R. (2006b). Investment manager's market timing abilities: Empirical evidences. *The ICFAI Journal of Applied Finance, 12*(8), 15–31.
Chander, S., & Singh, J. (2004). Performance of mutual funds in India: An empirical evidence. *The ICFAI Journal of Applied Finance, 10*, 45–63.
Cheng-Ru Wu, Hsin-Yuan Chang, Li-Syuan Wu (2008). A framework of assessable mutual fund performance. *Journal of Modeling in Management*, *3*(2), 125–139.
Choi, Y. K. (2006). Relative portfolio performance evaluation and incentive structure. *Journal of Business, 79*(2), 903–921.
Christofeersen, S. E. (2001). Why do money fund managers voluntarily waive their fee? *Journal of Finance, 74*, 117–140.
Coates, J. C., & Hubbard, G. R. (2007). *Competition in the mutual fund industry: Evidence and implications for policy* (Discussion Paper No.592). Source: http://ssrn.com/abstract=1005426
Comer, G. (2006). Hybrid mutual funds and market timing performance. *Journal of Business, 79*(2), 771–797.
Detzler, L. M. (1999). The performance of global bond mutual funds. *Journal of Banking and Finance, 23*, 1195–1217.
Dhanda, S. K., Batra, G. S., & Anjum, B. (2012). Performance evaluation of selected open ended mutual funds in India. *International Journal of Marketing, Financial Services & Management Research*, *1*(1), January 2012, ISSN 2277 3622. Online Available at http://indianresearchjournals.com
Domain, D. L., & Rechenstein, W. (1998). Performance and persistence in money market mutual fund returns. *Financial Services Review, 6*, 169–183.
Droms, W. G., & Walker, D. A. (1994). Investment performance of international mutual funds. *Journal of Financial Research, 17*, 1–14.
Dunn, P. C., & Theisen, R. D. (1983). How consistently do active managers win? *Journal of Portfolio Management, 9*, 47–51.
Elango, R. (2003). Which fund yields more returns? A comparative analysis on the NAV performance of select public v/s private/foreign open-ended mutual fund schemes in India. *Mutual Funds*. The Management Accountant, Vol. 39(4), (2004), p283–290.
Eling, M. (2006). Performance measurement of hedge funds using data envelopment analysis. *Financial Markets and Portfolio Management, 2o*, 442–471.

Elton, E. J., Gruber, M. J., & Blake, C. R. (1996). Market timing ability and volatility implied in investment newsletters' asset allocation recommendations. *Journal of Financial Economics, 42*, 397–421.

Eun, C. S., Kolodny, R., & Resnick, B. G. (1991). U.S. based international mutual funds: A performance evaluation. *The Journal of Portfolio Management, 17*, 88–94.

Fama, E. F. (1972). Components of investment performance. *Journal of Finance, 27*, 551–567.

Friend, I., Brown, F. E., Herman, E. S., & Vickers, D. (1962). *A study of mutual funds*. Washington, D.C.: U.S. Government Printing Office.

Goetzman, W. N., & Ibbotson, R. G. (1994). Do winners repeat; predicting mutual fund performance. *The Journal of Portfolio Management, 20*, 9–18.

Grinblatt, M., & Titman, S. (1989). Mutual fund performance – An analysis of quarterly holdings. *Journal of Business, 62*, 393–416.

Grinblatt, M., & Titman, S. (1992). The persistence of mutual fund performance. *Journal of Finance, 47*(5), 1977–1984.

Grinblatt, M., & Titman, S. (1994). Momentum investment strategies – Portfolio performance and herding: A study of mutual fund behavior. *American Economic Review, 85*, 1088–1105.

Gupta, M., & Aggarwal, N. (2009). Mutual fund portfolio creation using industry concentration. *The ICFAI Journal of Management Research, Viii*(3), 7–20.

Gupta, O. P., & Gupta, A. (2004). Performance evaluation of select Indian mutual fund schemes: An empirical study. *The ICFAI Journal of Applied Finance, 10*, 81–97.

Hendricks, D., Patel, J., & Zeckhauser, R. (1993). Hot hands in mutual funds: A short run persistence of relative performance. *Journal of Finance, 48*, 93–130.

Jayadeve, M. (1996). Mutual fund performance: An analysis of monthly returns. *Finance India, X*(1), 73–84.

Jensen, M. C. (1967). The performance of mutual funds in the period 1945–1964. *Journal of Finance, 23*, 389–416.

Jensen, M. C. (1968). The performance of mutual funds in the period 1945–1964. *Journal of Finance, 23*, 389–416.

Jow-Ran Chang, Mao-Wei Hung, Cheng-few Lee (2003). An intemporal CAPM approach to evaluate mutual fund performance. *Review of Quantitative Finance and Accounting*, *20*, 425–433.

Kajshmi, N., Malabika, D., & Murugesan, B. (2008). Performance of the Indian mutual funds; A study with special reference to growth schemes. *Asia-Pacific Business Review, Iv*(3), 75–81.

Khurshid, S. M. Z., Rohit, & Sing, G. P. (2009). Level and trends of competition among the mutual funds in India. *Research Journal of Business Management, 3*(2), 47–67.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics, 47*(1), 13–37.

Massa, M., & Lei Zhang. (2008). *The effects of organizational structure on asset management*. http://ssrn.com/abstract=1328189

Meijun Qian (2009). Stale prices and the performance evaluation of mutual funds. *Journal of Financial and Quantitative Analysis*.

Mc Donald, J. G. (1974). Objectives and performance of mutual funds 1960–1967. *Journal of Financial and Quantitative Analysis, 9*, 311–333.

Modigliani, F., & Modigliani, L. (1997). Risk adjusted performance. *Journal of Portfolio Management, 23*, 45–54.

Obeid, A. T. (2004). A modified approach for risk-adjusted performance attribution. *Financial Markets and Portfolio Management, 18*(3), 285–305.

Philpot, J., Hearth, D., Rimbey, J., & Schulman, C. (1998). Active management, fund size and bond mutual fund returns. *The Financial Review, 33*, 115–126.

Ramasamy, B., & Yeung, C. H. M. (2003). Evaluating mutual funds in an emerging market: Factors that matter to financial advisors. *International Journal of Bank Marketing, 21*, 122–136.

Ranu, A., & Ranu, D. (2010, January). The best performing equity funds. *The Analyst*.

Redmand, A. L., Gullett, N. S., & Manakyan, H. (2000). The performance of global and international mutual funds. *Journal of Financial and Strategic Decisions, 13*(1), 75–85.

Sadhak, H. (2003). *Mutual funds in India: Marketing strategies and investment practices*. New Delhi: Response Books.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance, 19*, 225–242.

Sharpe, W. F. (1966). Mutual fund performance. *Journal of Business, 39*, 119–138.

Sidana, G. (2007). Classifying mutual funds in India: Some results from clustering. *Indian Journal of Economics and Business*, *II*(2).

Sisodiya, A. S. (2004, July). Mutual fund industry in India – Coming of age. *Chartered Financial Analyst*, pp. 17–22.

Smith, D. M. (2009). The economics of mutual funds. In J. A. Haslem (Ed.), *A companion to mutual funds* (Chap. 3). New York: Wiley.

Smith, K. & Tito, D. (1969). Risk-return measures of post-portfolio performance. *Journal of Financial and Quantitative Analysis, 4*, 449–471.

Sondhi, H. J. (2007). *Financial performance of equity mutual funds in India*. New Delhi: Deep & Deep.

Statman, M. (2000). Socially responsible mutual funds. *Financial Analysts Journal, 56*, 30–38.

Subha, M. V., & Bharati, J. S. (2007). An empirical study on the performance of select mutual fund schemes in India. *Journal of Contemporary Research in Management*, *I*(1).

Sudhakar, A., & Sasi, K. K. (2005, November). Performance evaluation of mutual funds: A case study. *Southern Economist*, pp. 19–23.

Thanou, E. (2008). Mutual fund evaluation during up and down market conditions: The case of Greek equity mutual funds. *International Research Journal of Finance and Economics*, (13).

Treynor, J. L. (1965). How to rate management of investment funds. *Harvard Business Review, 43*, 63–75.

Treynor, J. L., & Mazuy, K. K. (1066). Can mutual funds outguess the markets. *Harvard Business Review, 44*, 131–136.

Treynor, J. L., & Mazuy, K. K. (1966). Can mutual funds outguess the markets. *Harvard Business Review, 44*, 131–136.

Volkman, D. A., & Wohar, M. E. (1995). Determinants of persistence in relative performance of mutual funds. *Journal of Financial Research, 18*, 415–430.

Wahal, S., & Alber (Yan) Wang (2010). Competition among mutual funds. *Journal of Financial Economics*. Source: http://ssrn.com/abstract=1130822

Yuxing Yan (1999). Measuring the timing ability of mutual fund managers. *Annals of Operations Research, 87*, 233–243.

Simulation as a Research Tool for Market Architects

4

Robert A. Schwartz and Bruce W. Weber

Contents

R.A. Schwartz (✉)
Zicklin School of Business, Baruch College, CUNY, New York, NY, USA
e-mail: robert.Schwartz@baruch.cuny.edu

B.W. Weber
Lerner College of Business and Economics, University of Delaware, Newark, DE, USA
e-mail: bweber@udel.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_4,

Abstract

Financial economists have three primary research tools at their disposal: theoretical modeling, statistical analysis, and computer simulation. In this chapter, we focus on using simulation to gain insights into trading and market structure topics, which are growing in importance for practitioners, policy-makers, and academics. We show how simulation can be used to gather data on trading decision behavior and to analyze performance in securities markets under controlled yet competitive conditions. We find that controlled simulations with participants are a flexible and reliable research tool when it comes to studying issues involving traders and market architecture. The role of the discrete event simulation model we have developed is to create a backdrop, or a controlled stochastic environment, for running market experiments with live subjects. Simulations enable us to gather data on trading participants' decision making and to ascertain the ability of incentives and market structures to influence outcomes. The statistical methods we use include experimental design and careful controls over experimental parameters such as the instructions given to participants. Furthermore, results are assessed both at the *individual* level to understand how participants respond to incentives in a trading setting and also at the *market* level to know whether the predicted outcomes are achieved and how well the market operated.

There are two statistical methods described in the chapter. The first is discrete event simulation and the model of computer-generated trade order flow that we describe in Sect. 4.3. To create a realistic, but not ad hoc, market background, we use draws from a log-normal returns distribution to simulate changes in a stock's fundamental value, or P*. The model uses price-dependent Poisson distributions to generate a realistic flow of computer-generated buy and sell orders whose intensity and supply-demand balance vary over time. The order flow fluctuations depend on the difference between the current market price and the P* value. In Sect. 4.4, we illustrate the second method, which is experimental control to create groupings of participants in our simulations that have the same trading "assignment." The result is the ability to make valid comparisons of traders' performances in the simulations.

Keywords

Trading simulations • Market microstructure • Order flow models • Random walk models • Experimental economics • Experimental control

4.1 Studying Market Structure

Financial economists have three major research methods at their disposal: theoretical modeling, statistical analysis, and simulation. We will focus on using simulation to gain insights into trading and market structure. In so doing, we show how simulation can be used to analyze participant behavior in a security market. We find that controlled simulations with participants are a flexible and reliable research tool when it comes to studying trading behavior and market architecture.

Using good experimental design we can draw statistically valid conclusions from simulations at both the *individual* level to understand how participants respond to incentives in a trading setting and also at the *market* level to know whether the predicted outcomes are achieved and how well the market operated. We begin by considering the interaction between the three tools.

4.2 The Interaction Between Theoretical Modeling, Empirical Analysis, and Simulation

A theoretical formulation based on a limited number of abstract assumptions enables complex reality to be translated into a simplified representation that can be rigorously analyzed. As such, theoretical modeling typically provides the underpinnings for both empirical and simulation analysis. Theory alone, however, can take us only so far. One cannot expect that every detailed aspect of reality can be analyzed from a theoretical vantage point (Clemons and Weber 1997). Moreover, in light of the literature on behavioral economics, it is clear that not all human behavior follows the dictates of rational economic modeling.

We turn to empirical analysis both to provide confirmation of a theoretical model and to describe aspects of reality that theoretical analysis has not been able to explain (Zhang et al. 2011). But necessary data may not be available and empirical analysis, like theoretical modeling, has its limitations. Variability in important variables, such as differences across time or traded instruments, is difficult for empiricists to control (Greene 2011; Kennedy 2008). Moreover, empirical variables may change in correlated ways, yet this should not be mistaken for a causal mechanism. When we seek insights that neither theory nor empirical analysis can provide, simulation has an important role to play (Parker and Weber 2012). An example will help explain.

4.2.1 An Application: Call Auction Trading

Consider an electronic call auction. At a call, submitted orders are batched together for simultaneous execution at a single clearing price. For the batching, buy orders are cumulated from the highest price to the lowest to produce a function that resembles a downward sloping demand curve, and sell orders are cumulated from the lowest price to the highest to produce a function that resembles an upward sloping supply curve. The algorithm generally used for determining the clearing price at the time that the market is called finds the price which maximizes the number of shares that execute. In an abstract, theoretical model, the number of shares that trade is maximized at the price where the downward sloping buy curve crosses the upward sloping sell curve. Consequently, with continuous order functions, the clearing price which maximizes the number of shares that trade in a call auction is uniquely determined by the point where the two curves cross, and this price, in economics parlance, is an equilibrium value.

All told, this call auction procedure has excellent theoretic properties, and one might expect that designing a well-functioning call would be a straightforward task. This is not the case, however. In reality, as the saying goes, "the devil is in the details."

For theoretical modeling, we might for analytic convenience assume a large enough number of participants so that no one individual has any market power. We might further assume that the cumulated buy and sell curves are continuous functions and that the call auction is the only trading facility available. In reality, however, the buy and sell curves are step functions and some players' orders will be large enough to impact a call's clearing price.

To illustrate, assume an exact match of 40,000 shares to buy and 40,000 shares to sell at a price of $50 and that, at the next higher price of $50.10, sell orders totaling 50,000 shares and buy orders totaling 30,000 shares exist. A buyer could move the price up by entering more than 10,000 shares at $50.10 or greater.

Notice also that the real-world buy and sell curves are step functions (neither price nor quantity is a continuous variable) and thus that the cumulated buy orders may not exactly match the cumulated sell orders at the price where the two curves cross. Moreover, many exchange-based call auctions are offered along with continuous trading in a hybrid environment. These realities of the marketplace affect participants' order placement decisions and, consequently, impact market outcomes for both price and the number of shares that trade. In response, participants will enter their orders *strategically* when coming to the market to trade.

These strategic interactions and the decisions that market participants make depend on the call auction's rules of order disclosure (i.e., its transparency) and its rules of order execution which apply when an exact cross is not obtained. What guidance do market architects have in dealing with questions such as these other than their own, hopefully educated, speculation? The questions being raised may be too context specific for theory to address and, if a new market structure is being considered, the data required for empirical analysis will not yet exist. This is when simulation analysis can be used to good advantage.

Regarding transparency, alternative call auction structures include full transparency (i.e., display the complete set of submitted orders), partial transparency (e.g., display an indicated clearing price and any order imbalance at that price), or no transparency at all (i.e., be a dark pool). Regarding the procedure for dealing with an inexact cross, the alternatives for rationing orders on the "heavy" side of the market include pro rata execution, the application of time priority to orders at the clearing price exactly, and the application of time priority to all orders at the clearing price and better (i.e., to higher priced buys or to lower priced sells). These and other decisions have been debated in terms of, for instance, the ability to game or to manipulate an auction, the incentive to enter large orders, and the incentive to submit orders early in the book building period before the auction is called. But definitive answers are difficult to come by. This is when valuable guidance can be (and has been) obtained via the use of simulation analysis.

4.2.2 An Application: The Search for an Equilibrium Price and Quantity

In this section, we consider another application of simulation as a research tool: the search for an equilibrium price and quantity in a competitive marketplace. Determining equilibrium values for a resource's unit price and quantity traded is a keystone of economic analysis. In the standard formulation, equilibrium is determined by the intersection of market demand and supply curves, with scant consideration given to just how the buy and sell orders of participants actually meet in a marketplace. In fact, the "marketplace" is typically taken to be nothing more than a mystical, perfectly frictionless environment, and the actual discovery of equilibrium values for price and quantity is implicitly assumed to be trivial.

Real-world markets are very different. In the non-frictionless environment, a panoply of transaction costs interact to make price and quantity discovery an imperfect process. In this far more complex setting, two further issues need to be analyzed: (1) the trading decisions that market participants make when confronted by the imperfections (frictions) that characterize real-world markets and (2) the determination of actual prices and quantities based on the decisions of the individual participants. Both issues are ideally suited for simulation analysis.

We consider this further in this section of the chapter, with particular reference to one specific market: the secondary market for trading equity shares of already issued stock. Underlying the simulation analysis is an economic model that is based on the following assumptions:

1. The decision maker qua investor is seeking to maximize his/her expected utility of wealth as of the end of a single holding period.
2. The investor is risk averse.
3. There are just two assets, one risk-free (cash) and one risky (equity shares).
4. There are no explicit trading cost (i.e., there are no commissions or borrowing costs, and short selling is unrestricted).
5. Share price and share holdings are continuous variables.
6. A brief trading period (T_0 to T_1) that is followed by a single investment period (T_1 to T_2).
7. The participant's expectation of price at T_2 is exogenous (i.e., independent of the stock's current price).
8. The investor is a perfect decision maker when it comes to knowing his/her demand curve to hold shares of the risky asset.

From this set of assumptions, and following the derivation described in Ho et al. (1985) and Francioni et al. (2010), we can obtain a participant's demand to hold shares of the risky asset. The participant's demand is described by two simple, linear functions:

$$P^0 = a - 2bN \quad (\text{an ordinary demand curve})$$

and

$$P^R = a - bN \quad (\text{a reservation price demand curve})$$

where P^0 denotes price with respect to the ordinary curve, P^R denotes a reservation price, and N is the number of shares held.

The ordinary curve shows that if, for instance, the price of shares is P^0_1, the participant maximizes expected utility by holding N_1 shares; if alternatively price is P^0_2, the participant maximizes expected utility by holding N_2 shares; and so on. Values given by the reservation curve show that at a quantity N_1, the maximum the participant would pay is P^R_1 when the alternative is to hold no shares at all; or, at a quantity N_2, the maximum the participant would pay is P^R_2 when the alternative is to hold no shares at all; and so on. Identifying the reservation price demand curve enables us to obtain easily a monetary measure of the gains from trading.

To facilitate the exposition, assume for the moment that the participant initially holds no shares. Then if, for instance, N_1 shares are acquired, we have

$$\text{Surplus} = N_1\left(P^R - P\right)$$

where "Surplus" denotes the gains from buying N_1 shares and P is the price at which the N_1 shares were bought (note that, for a purchase, we have Surplus > 0 for $P < P^R$). The participant controls P via the price of his/her order but knows neither the price at which the order will execute (if price improvement is possible) nor whether or not the order will, in fact, execute. Because P is not known with certainty, Surplus is not known with certainty and thus the investor seeks to maximize the expected value of Surplus. It follows from the derivations cited above that the maximization of the expected value of Surplus is consistent with the maximization of the expected utility of the end of the holding period wealth (i.e., at time T_2).

With the demand curves to hold shares established, it is straightforward to obtain linear functions that describe the investor's propensity to buy and to sell shares in relation to both ordinary and reservation prices. If the investor's initial position is zero shares, the buy curve is the same as the downward sloping demand curve. By extending the demand curve up and through the price intercept (a) into the negative quadrant, we see that at prices higher than the intercept, a, the investor will want to hold a negative number of shares (i.e., establish a short position by selling shares that he/she does not currently own). By flipping the portion of the demand to hold curve that is in the negative quadrant into the positive quadrant (and viewing a negative shareholdings adjustment as a *positive* sell), we obtain a positively inclined (vertical) mirror image that is the sell function.

The procedure just described can easily be replicated for any initial share holdings (either a long or a short position). The individual buy and sell curves can be aggregated across investors and, following standard economic theory, the intersection of the aggregate buy curve with the aggregate sell curve establishes the equilibrium values of share price and the number of shares that trade. This equilibrium would be achieved if all participants were, simultaneously, to submit their complete demand functions to the market, as they presumably would in a perfect, frictionless environment.

Continue to assume that investors know their continuous, negatively inclined demand curves to hold shares of the risky asset, but let us now consider how they might operate in a non-frictionless marketplace where they cannot all simultaneously submit their complete and continuous, downward sloping buy functions and upward sloping sell functions. What specific orders will they send to the market, and how will these orders interact so as to be turned into trades? Will equilibrium values for price and quantity be achieved? This highly complex issue might best be approached by observing participant behavior and, to this end, simulation can be used as the research tool. Here is how one might go about it.

Let participants compete with each other in a networked, simulated environment. Each participant is given a demand curve to hold shares of a risky stock, and each is asked to implement that demand curve by submitting buy or sell orders to the market. The participants are motivated to place their orders strategically given their demand curves, the architectural structure of the marketplace, and the objective against which their performance is assessed – namely, the maximization of expected surplus.

The simulation can be structured as follows. Give all of the participants in a simulation run the same demand curve,

$$P = 20 - 0.5N_0$$

where N_0 represents the number of shares initially held. Divide the participants into two equal groups, A and B, according to the number of shares they are initially holding, with $N_{0A} = 4$ for group A players and $N_{0B} = 8$ for group B players. Accordingly, the buy curve for each individual in group A is

$$P = 18 - 0.5Q$$

and the sell curve for each individual in group B is

$$P = 16 + 0.5Q$$

where Q is the number of shares bought or sold. The associated reservation curves are

$$P^R = 18 - 0.25Q \quad (\text{for the buyers})$$

and

$$P^R = 16 + 0.25Q \quad (\text{for the sellers})$$

From the ordinary (as opposed to the reservation) buy and sell curves, and recalling that the groups A and B are of equal size, we obtain an equilibrium price of 17 and an equilibrium quantity traded of 2 (per participant). From the reservation buy and sell curves, we see that if each participant bought or sold two shares, the surplus for each would be

$$2(17.50 - 17.00) = \$1.00 \ \ \text{(for the buyers)}$$

and

$$2(17.00 - 16.50) = \$1.00 \ \ \text{(for the sellers)}$$

Participants, however, do not know the distribution of initial shareholdings across the other investors, and thus they know neither the buy and sell functions of all the other traders, nor the equilibrium price and quantity for the market. It is up to each of them individually to submit their orders wisely, and it is up to all of them collectively to find the equilibrium price along with the quantity to trade at that price. How do they operate? How well do they do? How quickly and successfully can they collectively find equilibrium values for P and Q? And how are participant decisions and market outcomes affected by different market structures? With regard to each of these issues, important insights can be obtained through the use of simulation as a research tool.

4.2.3 The Realism of Computer-Generated Data Versus Canned Data

A live equity trading simulation can either depend exclusively on person-to-person interaction or add computer-driven order flow; we focus on the latter. Most experimental economics research is based on the former. Computer-driven simulations can be based on either canned data or on data that the computer itself generates; we focus on the latter. Canned data have two major limitations. First, participants in the simulation cannot affect the stream of prices that they are trading against and, consequently, the dynamic, two-way interaction between live participants and the marketplace is absent. Second, canned data cannot be used to analyze new market structure because, quite simply, it is exclusively the product of the actual market within which it was generated.

On the other hand, a simulation model based on computer-generated data faces a formidable challenge: capturing the dynamics of a real-world market. Canned data does not face this problem – it is, after all, generated in a real marketplace. We discuss the challenge of realism in this section of the chapter with specific reference to an equity market.

It has been well established in the financial economics literature that, in an equity market which is fully efficient, security prices follow random walks. “Efficiency” in this financial markets context is generally understood as referring to “informational efficiency,” by which we mean that market prices reflect all existing information. In a nutshell, the raison d’être of a stock’s price following a random walk in a fully informationally efficient market can be understood as follows. If all (and we mean *all*) information about a security is reflected in a stock's market price, then only totally new, totally unanticipated information can cause the stock's price to change. But totally new and thus totally unanticipated information can be either bullish or bearish with equal probability and thus, with equal probability, can lead to either positive or negative price changes (returns). Thus, the argument goes, in an informationally efficient market, returns are not predictable and stock prices follow random walks.

It would be relatively straightforward to structure an equity market simulation based on machine-driven prices that follow a random walk. One would start a simulation run with an arbitrarily selected seed price and have that price evolve as the simulation progresses according to random draws from a (log-normal) returns distribution with arbitrary variance and a zero mean. Real-world prices do not evolve in this fashion, however. In a world characterized by trading costs, imperfect information, and divergent (i.e., nonhomogeneous) expectations based on publicly available information, prices do not follow simple random walks. Rather, price changes (returns) in relatively brief intervals of time (e.g., intraday) evolve in dynamic ways that encompass complex patterns of first-order and higher-order correlations. Structuring a computer simulation to produce prices that capture this dynamic property of real-world markets is the objective. In the next session of this chapter, we set forth the major properties of a machine-driven trading simulation, TraderEx, which we have formulated so as to achieve this goal.

4.3 An Equity Market Trading Simulation Model

In this section, we focus on a key conceptual foundation of the TraderEx simulation model: how the machine-generated order flow is structured. Above all else, it is this modeling that captures the dynamic property of trades and prices and, in so doing, that enables our software to compete with canned data in terms of realism. First, we present a brief overview of the functions the computer performs.

Looking under the hood, the TraderEx software can be thought of, first and foremost, as a package of statistical distributions. The prices and sizes of the machine-driven orders that power the TraderEx simulation are determined by draws from these distributions. The software also maintains and displays an order book or set of dealer quotes, drives a ticker tape that displays trade prices and sizes, computes summary statistics for each simulation run (e.g., a stock's volume-weighted average price and participant performance measures such as profit or loss), and provides post trade analytics (both statistics and graphs). The simulation game can be played either individually (i.e., one person interacting with the machine-driven order flow in a solitaire environment) or as a group (i.e., in a networked environment that also includes machine-driven order flow).

While our machine-driven order flow, as we have said, is powered by random draws from distributions, we are able to tell meaningful economic stories about these statistical draws. These stories involve exogenous information change and three different economic agents: informed traders, liquidity traders, and noise traders. Interestingly, while the simulation requires this tripartite division, it has also been established in the academic microstructure literature that, for an equity market not to fail, informed traders must interact with liquidity traders, and noise traders must also be part of the mix.

4.3.1 Informed Orders

In TraderEx, orders submitted by informed traders are specified with reference to an exogenous variable we call P* that can be viewed as an equilibrium price. That is, at the value P*, the aggregate flow of buy and sell orders is in balance (much as, in economic analysis, buy and sell orders are in balance at the price where a demand curve crosses a supply curve). More specifically, the TraderEx market is in equilibrium when the lowest posted offer is greater than P* and the highest posted bid is less than P*. When this equilibrium is achieved, no informed orders are submitted to the market, and an incoming (liquidity) order can be from a buyer or a seller with equal probability. On the other hand, if P* is greater than the lowest posted offer, informed orders kick in and the probability of an incoming order being from a buyer is raised to 0.6 (a parameter that can, of course, be adjusted). Equivalently, if P* is lower than the highest posted bid, informed orders again kick in and the probability of an incoming order being from a seller is raised to 0.6. This asymmetry between the buy and sell orders that exists when P* is not within the quotes keeps the quotes loosely linked to P*.

P* evolves as the simulation progresses according to a Poisson arrival process. Each jump in P* symbolizes informational change. The size of the change in P* at each new arrival is determined by a draw from a log-normal returns distribution with a zero mean and a variance that is a controllable parameter.

4.3.2 Liquidity Orders

The second component of the order flow, liquidity orders, is also modeled as a Poisson arrival process, but with one important difference: at any point of time, the probability of the newly arriving liquidity order being a buy equals the probability of its being sell equals 0.5. All liquidity orders are priced, with the price determined by a draw from a double triangular distribution that is located with reference to the best posted bid and offer quotes.

A new liquidity order is entered on the book as a limit order if it is a buy with a price lower than the best posted offer or if it is a sell with a price higher than the best posted bid. A new liquidity order with a price equal to or more aggressive than the best posted offer (for a buy) or the best posted bid (for a sell) is executed immediately as a market order. Liquidity orders can (randomly) cause the market's bid and ask quotes to drift away from the equilibrium value, P*. When this occurs, informed orders that are entered as market orders pull market prices back towards P*.

4.3.3 Momentum Orders

The third component of the order flow is orders entered by noise traders. TraderEx activity includes just one kind of noise trader – a momentum player – and it operates as follows: whenever three or more buy orders (or sell orders) arrive sequentially, the conditional probability is increased that the next arriving order will also be a buy (or a sell).

As in the microstructure literature, noise traders are needed in the simulation model to keep participants from too easily identifying price movements that have been caused by informed orders responding to a change in P*. This is what our momentum orders achieve. For instance, assume that P* jumps several ticks above the best posted offer. An accelerated arrival of informed buy orders would be triggered and prices on the TraderEx book would rise over a sequence of trades, causing a pattern of positively autocorrelated price movements that can, with relative ease, be detected by a live participant. But, to obscure this, the momentum orders create faux price trends that mimic, and therefore obfuscate, the information-induced trends.

Momentum orders play a further role in the TraderEx simulations. They systematically cause transaction prices to overshoot P*. Then, as informed orders kick in, prices in the simulation mean revert back to P*. This mean reversion and its associated accentuated short-run volatility encourage the placement of limit orders. This is because overshooting causes limit orders to execute, and limit order placers profit when price then mean reverts. To see this, assume that the stock is currently trading at the $23.00 level and that P* jumps from $23.00 to $24.00. As price starts to tick up to $24.00, momentum orders join the march and carry price beyond $24.00–$24.20. Assume a limit order to sell is on the book at $24.20 and that it executes. The limit order placer then benefits from having sold at $24.20 when the momentum move ends and when a P* of $24.00 exerts its influence and price mean reverts back towards $24.00.

4.4 Simulation in Action

Since the earliest version of TraderEx was developed in 1995, we have run hundreds of "live market" simulations with students in our trading and market microstructure electives and with executive education participants. In addition, we have developed training modules on trading for new hires at a number of global banks. We have also run controlled experimental economics studies of trading decision making and alternative market structures (Schwartz and Weber 1997).

To illustrate the potential for research from using market simulation, we will examine the data generated by simulation participants' behavior in a number of settings. In a January 2011 class session of the "Trading and Financial Market Structure" elective at London Business School, a simulation was run which covered

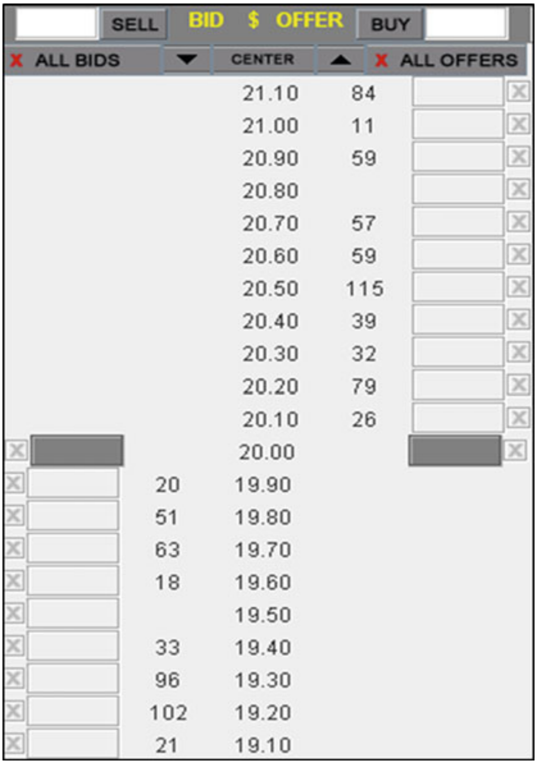


Fig. 4.1 Initial order book at start of simulated trading day. Limit orders to buy are on the *left* and limit orders to sell on the *right*. Participants enter buy and sell market orders in the *white boxes* at the *top* of the book. Limit orders are entered by clicking on the *gray rectangles* at the price level the user selects

1 day of trading in an order-driven market structure. In it, 42 graduate students were divided into 21 teams of two. The order book maintained price and time priority over limit orders. The price increment was 10 cents, and the order book at the open looked similar to Fig. 4.1 below.

4.4.1 Trading Instructions: Simulation A

Eight teams were each given the instruction to sell 1,500 units, and seven teams were each asked to buy 1,300. Five other teams had the role of either day traders or proprietary trading desks. Three of these five teams were instructed to buy 900 then sell 900 and to have a closing position of 0. Two teams were asked to sell 800, then buy 800, and to close with a flat position. A trial simulation was run, and performance metrics were discussed before the simulation began. The teams with a sell instruction were told they would be assessed on the basis of the *highest* average selling price, while buying teams competed on the basis of the *lowest* average buying price. The “prop” teams were told that maximizing closing P&L was their objective but that they should finish flat and have no “overnight risk.” A screen similar to the one the participants saw is shown in Fig. 4.2.

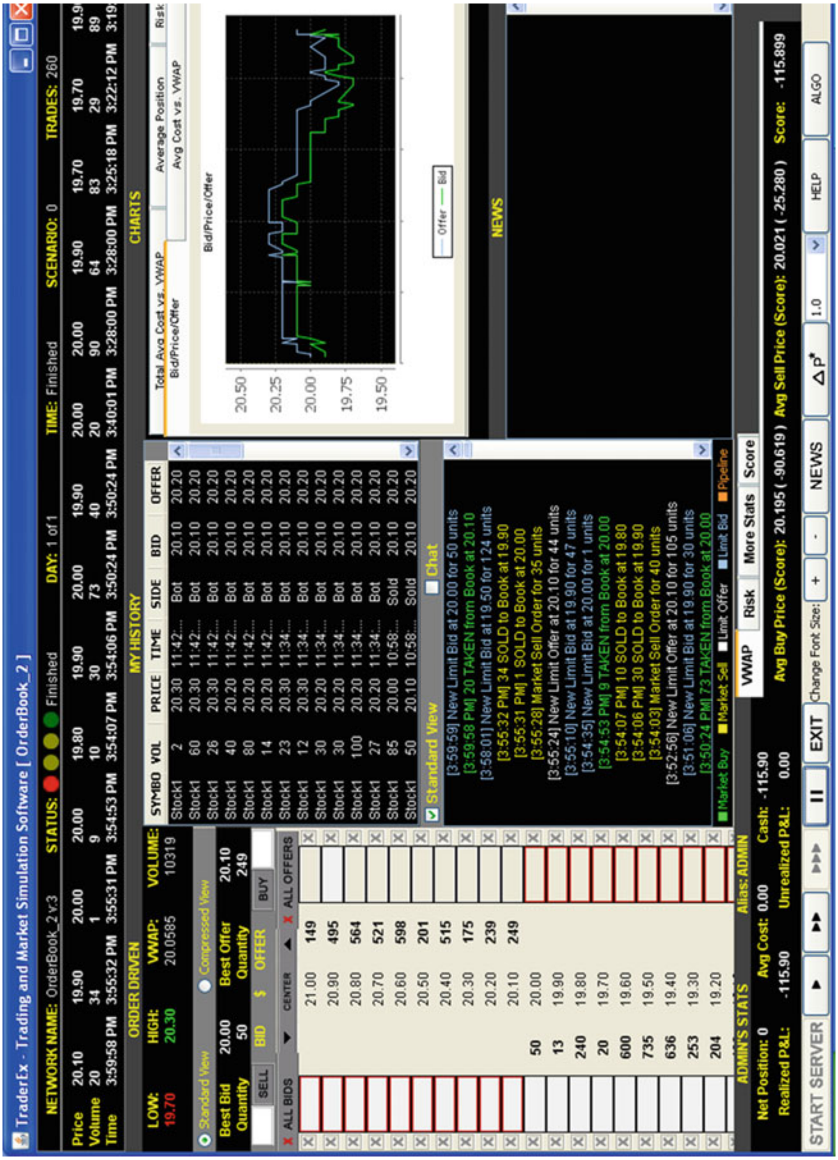


Fig. 4.2 End of a trading day in a networked order book simulation from the screen used by the instructor or simulation administrator

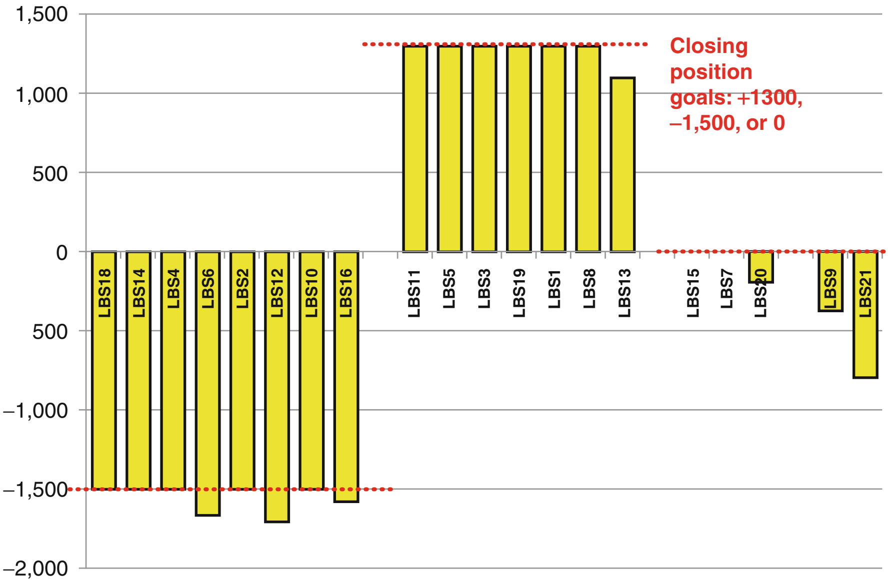


Fig. 4.3 Final positions of 21 trading teams. Teams were supposed to end with holdings of either −1,500 (sellers), +1,300 (buyers), or flat (0, prop-day traders)

One of the first lessons of behavioral studies of trading done via simulation is that following instructions is *not simple* for participants. As Fig. 4.3 shows, seven of the 21 teams did not end the simulation with the position they were instructed to have. Three of the selling teams sold more than instructed and three of the proprietary trading teams had short, nonzero positions at the end of the day. Of course, the noncompliant teams had excuses – "*the end of the day came too fast*," or "*there were not enough willing buyers in the afternoon*," or "*we didn't want to pay more than the VWAP price on the screen*." These complaints could, of course, also apply in a real equity market.

In the simulation, the market opened at £20.00. The day's high was £23.60, the low £18.80, and the last trade of the day was at £21.50. The £4.80 high-low range (24 %) reflects a volatile day. The day's volume-weighted average price (VWAP) was £20.03. Trading volume was 42,224 units and 784 trades took place. Although the teams were in the same market, and had the same buying or selling instructions with the same opportunities, there were vast differences in performance. As Fig. 4.4 shows, the best buying team paid £19.68, or £0.35 less than both the worst team and VWAP, adding nearly 2 % to the investment return. The best selling team received £0.95 more per share than VWAP and £1.23 per share more than the worst selling team. This outcome would add almost 5 % to one selling investor's return relative to the average. The conclusion is clear: trading performance has a substantial impact on investors' returns.

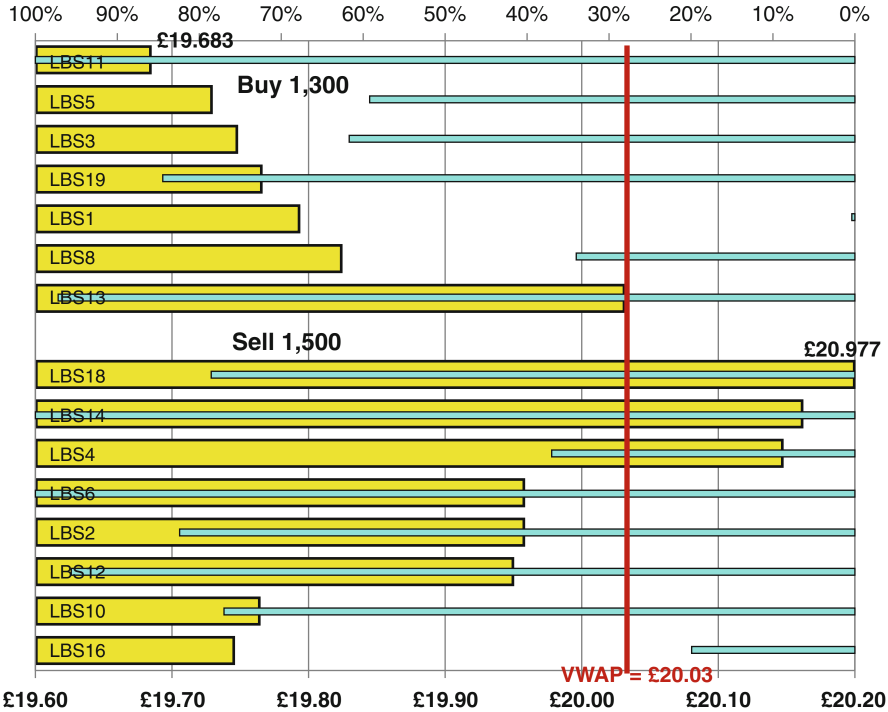


Fig. 4.4 Performance of trading teams as measured by average buying or selling price (scale shown on *bottom* of figure reading *left* to *right*) and the percentage of teams' trading done via limit orders (scale shown on *top* of figure reading *right* to *left*)

Also shown in Fig. 4.4 is the team's use of limit orders. Again, there is substantial variation, with the best buying team trading exclusively with limit orders and the best selling team completing 79 % of its trading with limit orders. Note that, as the chart shows, a higher use of limit orders did not assure good trading outcomes.

Note in Fig. 4.4 that the buying teams all matched or improved on VWAP, while only three of the selling teams were able to for more than VWAP. Teams' trading was completed with a varied combination of limit orders and market orders, with the best team, for instance, selling its 1,500 with 77 % limit orders and 23 % market orders. No significant correlation existed between order choice and performance.

As Fig. 4.5 shows, among the five proprietary trading teams, the largest loss was generated by the team (LBS21) that also had the largest risk as measured by average absolute value of their inventory position during the trading day. The greatest profit came from LBS15, a team that only took a moderate level of risk. Again, the simulation reveals substantial variation and behavioral differences across traders.

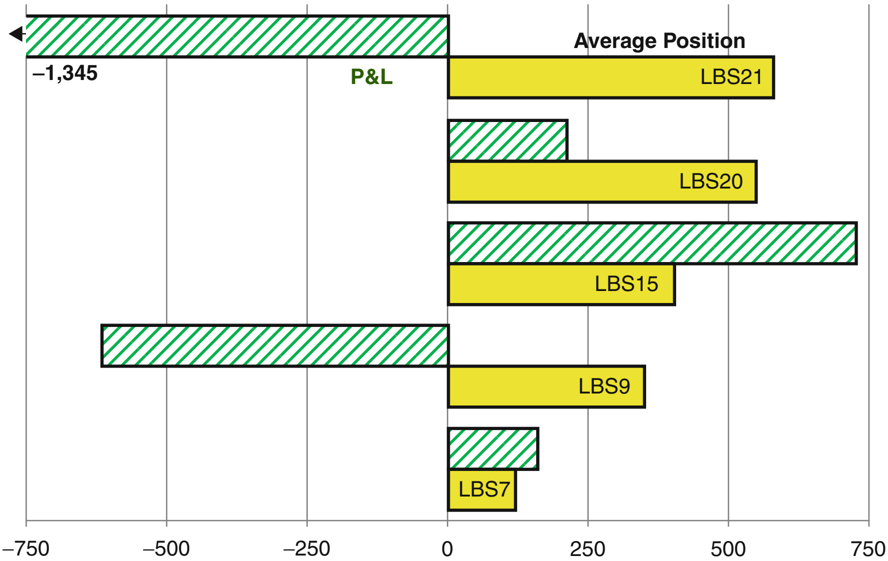


Fig. 4.5 Performance of 5-day trading teams. Collectively the proprietary traders lost money, and only two teams returned to a zero position

4.4.2 Trading Instructions: Simulations B-1 and B-2

A quote-driven market structure was used in two other simulations with the same graduate students. The underlying order flow and P* model simulated by the computer is the same as in the order-driven market. The students were put into 22 teams, and two markets (i.e., two separate networks) with 11 teams each were run. In each of the networks, seven teams were market makers, and the other teams were given large orders to fill over the course of the day.

In B-1, the market opened at 20.00, and the day's high and low prices were 21.00 and 18.70, respectively. The last trade was 19.50, and the VWAP was 20.00, with 706 trades generating a volume of 39,875 units. Participants were told that a unit represented 1,000 shares and that those with buy or sell instructions were handling institutional-sized orders that could affect prices.

In their trading, four of the market makers generated a positive profit; as a group they earned 1.4 pence per share, or seven basis points, in their trading. As Fig. 4.6 shows, only three of the dealers (LBS1, LBS3, LBS2) ended the day with a flat (either zero or no more than ±35 units) closing position. The chart shows the seven dealers' average inventory position, closing profit or loss, and their closing inventory position. LBS4, for instance, had an average position over the day of +797, a loss of 674, and a closing position of +1,431.

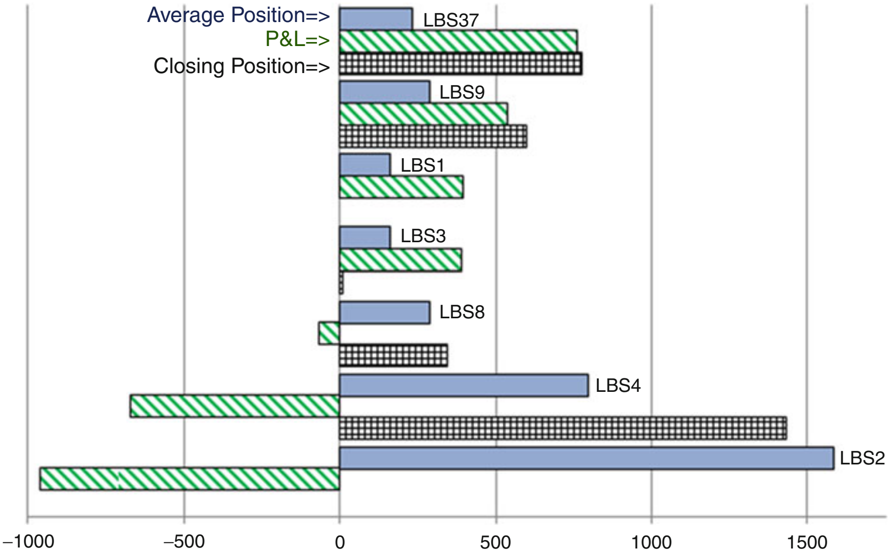


Fig. 4.6 Performance of seven market maker trading teams. The teams that controlled risk by keeping their average absolute value of their positions below 350 were able to make profits. Teams LBS4 and LBS2 incurred large losses and held large risk positions

In the B-2 market, the simulation started at 20.00, and the day's high and low prices were 22.30 and 19.00, respectively. The last trade was 22.20, and the VWAP was 20.22, with 579 trades generating volume of 36,161 units. Reflecting higher volatility, only two market makers generated a positive profit and, as a group, they lost 2.7 pence per share, or −13 basis points, in their trading. Figure 4.7 shows that three of the six dealers (LBS31, LBS25, LBS3) had a flat position at the end of the day. The other dealers had short positions reflecting the P* (equilibrium price) increases and buying pressure that drove the prices up during the simulation.

4.4.3 Trading Instructions: Simulation C

An order-driven market structure was used to study trading when participants are given either an *upward sloping supply curve* or a *downward sloping demand curve*. Rather than being instructed to buy a fixed quantity of shares, buyers and sellers are asked to maximize their surplus or profit, which is the number of shares bought or sold times the amount their *reservation price* differs from the average price they paid. The reservation prices for buyers are decreasing in the quantity they hold in their position, a demand curve. The reservation prices for sellers are increasing in the quantity they have sold, a supply curve. Our interest is in whether a group of trading participants will trade the optimal quantity in the market or whether in the absence of explicit order quantities they overtrade, or trade too little.

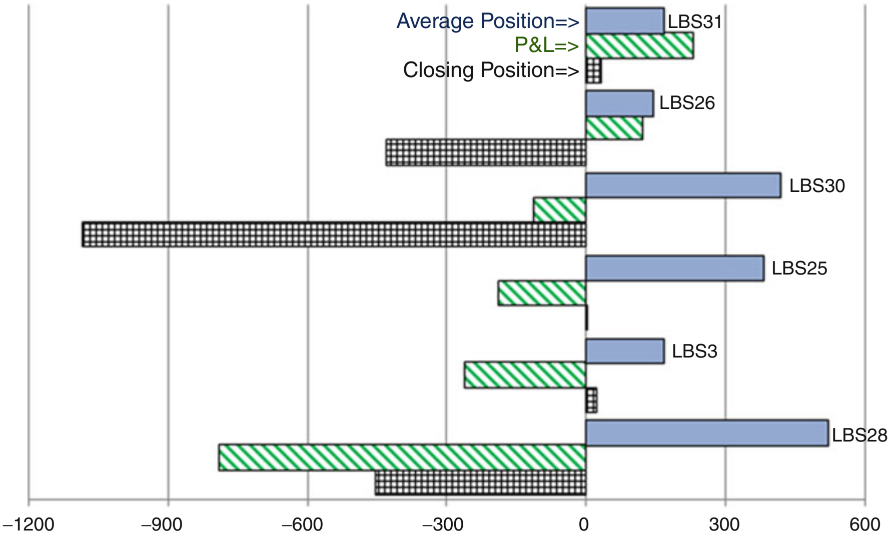


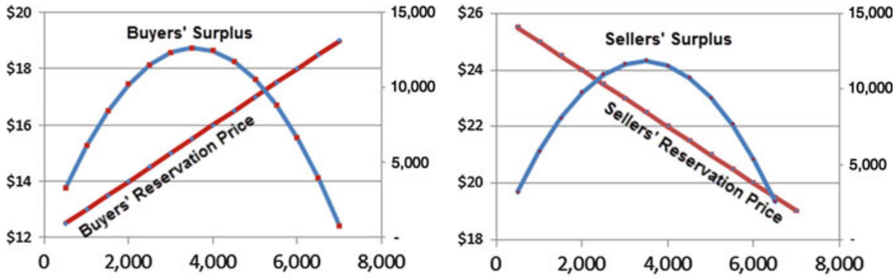
Fig. 4.7 Performance of six market maker trading teams. Only two teams, LBS #31 and #26, were able to make profits, and the teams that allowed their average position to exceed 300 had the largest losses

In the example below, we provide the buyers' reservation price function, which starts at $26.00 but decreases by one for each 1,000 shares bought. If a participant buys 6,000 shares over the course of the simulation, for instance, the reservation price is $20.00. If they paid on average $18, then the surplus generated is 12,000.

<u>Reservation BUY curve:</u>	P^R = \$26 – (X units/1,000)	
# shares bought	=	__________
P^R = 26 – (# shares bought/1,000)	=	__________
Average buy price	=	__________
Surplus = (# shares) × (P^R – Avg buy price)	=	__________

The sellers' reservation price function is below. It starts at $12.00 but increases by one for each 1,000 shares bought. If a participant sells 2,000 shares, the reservation price is $14.00. If the sellers' average selling price is $18, for instance, the surplus is 16,000.

<u>Reservation SELL curve:</u>	P^R = \$12 + (X units/1,000)	
# shares sold	=	__________
P^R = 12 + (# shares sold/1,000)	=	__________
Average sell price	=	__________
Surplus = (# shares) × (Avg sell price – P^R)	=	__________



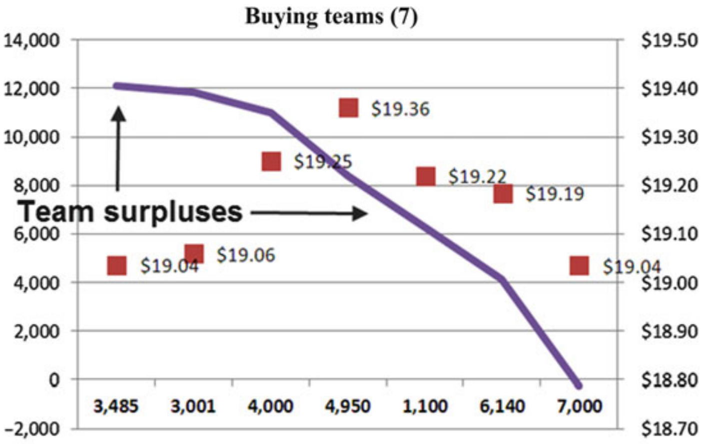
Avg Price=	$ 19.11	$12.00	$26.00	BUY	SELL
	500	$ 12.50	$ 25.50	3,195	3,305
	1,000	$ 13.00	$ 25.00	5,890	6,110
	1,500	$ 13.50	$ 24.50	8,085	8,415
	2,000	$ 14.00	$ 24.00	9,780	10,220
	2,500	$ 14.50	$ 23.50	10,975	11,525
Optimal	3,000	$ 15.00	$ 23.00	11,670	12,330
Quantity:	**3,500**	$ 15.50	$ 22.50	**11,865**	**12,635**
	4,000	$ 16.00	$ 22.00	11,560	12,440
	4,500	$ 16.50	$ 21.50	10,755	11,745
	5,000	$ 17.00	$ 21.00	9,450	10,550
	5,500	$ 17.50	$ 20.50	7,645	8,855
	6,000	$ 18.00	$ 20.00	5,340	6,660
	6,500	$ 18.50	$ 19.50	2,535	3,965
	7,000	$ 19.00	$ 19.00	-	770

Fig. 4.8 Trading experiment with supply and demand functions and surplus calculations based on a participant trading at the market average price over the day of $19.11. The optimal trading strategy was to end with a position of 3,500 long (buyer) or short (seller)

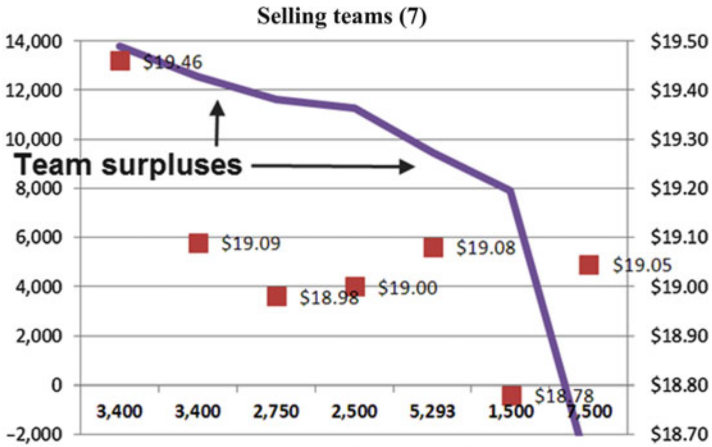
We provided the surplus functions above to a group of 28 participants. The group was split into teams of two, with seven buying teams and seven selling teams. Although the teams were not aware of the other teams' curves, the equilibrium price was $19, and the maximum surplus was achieved by participants that built a position of 3,500 and bought at the lowest possible prices or sold at the highest prices (see Fig. 4.8). In our experiment, average per share trade price was $19.11.

In the experiment, participants in the 14 teams built positions ranging from 1,100 to 7,500 (see Fig. 4.9). There was a substantial dispersion of performance, yet the market's trade prices converged to within $0.11 of the $19.00 equilibrium price. The average ending position was 4,001, so participants were within 15 % of the optimal position implicit in the reservation price functions. The teams with the greatest surpluses were able to come close to the optimal position of 3,500 and to buy below the average price and sell at greater than the average price.

As these three simulation exercises show, a number of insights can be gained into the effect of market structure on participant behavior and market outcomes (the quality of price and quantity discovery) by running "live market" simulations. First, participants' performance varied widely despite being in the same environment with the same trading instructions. Second, markets are complicated and not



BUYERS	Name	Qty	Avg Price	P r	Surplus
	Mustafa	3,485	$ 19.04	$ 22.52	**12,124**
	Thilo	3,001	$ 19.06	$ 23.00	11,821
	George	4,000	$ 19.25	$ 22.00	10,996
	Frank	4,950	$ 19.36	$ 21.05	8,366
	Pietro	1,100	$ 19.22	$ 24.90	6,248
	Oliver	6,140	$ 19.19	$ 19.86	4,145
	Jung	7,000	$ 19.04	$ 19.00	-245



SELLERS	Name	Qty	Avg Price	P r	Surplus
	Laurianne	3,400	$ 19.46	$ 15.40	**13,804**
	Maja	3,400	$ 19.09	$ 15.40	12,546
	Fredrick	2,750	$ 18.98	$ 14.75	11,633
	Jim	2,500	$ 19.00	$ 14.50	11,250
	Marcus	5,293	$ 19.08	$ 17.29	9,459
	Harald	1,500	$ 18.78	$ 13.50	7,920
	Ulrich	7,500	$ 19.05	$ 19.50	-3,405

Fig. 4.9 Even when provided with reservation price functions, the live participants in the markets traded near the equilibrium and showed little inclination to "overtrade." The average position was only 501 units away from the optimal closing position of 3,500

perfectly liquid; filling large orders, while competing to outperform price benchmarks, is challenging. Participants often did not complete their instructions even though completion was part of their assessment. Third, trading can add (or subtract) value for investment managers. Even in fairly simple simulation games, performance varied widely across participant teams. Finally, risk was not well controlled by the (admittedly inexperienced) trading participants. Although given specific position limits as guidelines, many participants had large average positions and suffered substantial losses from adverse price movements.

Market architects today combine different trading systems designs (Zhang et al. 2011). Beyond studies of trading decision making, the live simulations provide a method for comparing alternative market structures. For instance, implicit trading costs incurred in handling a large order over a trading day can be contrasted in markets with and without dealers, and with and without a dark liquidity pool. Live simulations, by extending what can be learned with analytical modeling and empirical data analytics, provide a laboratory for examining a broad set of questions about trading behavior and market structures.

4.5 Conclusion

Simulation is a powerful research tool that can be used in conjunction with theoretical modeling and empirical research. While simulation can enable a researcher to delve into issues that are too detailed and specific for theory to handle, a simulation structure must itself be based on a theoretical model. We have illustrated this with reference to TraderEx, the simulation that we have developed and used to analyze details of market structure and trading in an equity market.

One further research methodology incorporates simulation: experimental economics. This application applies the simulated market in a laboratory where multiple players are networked together. With a well-defined performance measure and carefully crafted alternative market structure and/or information environments, a simulation-based experimental application can yield valuable insights into the determinants of market outcomes when market structure affects individual behavior and when behavioral economics along with theoretically rational decision making characterizes participant actions.

Appendix

Modeling Securities Trading

The simulation methodology was chosen for its ability to accommodate critical institutional features of the market mechanism and off-exchange dealers' operations. While higher level abstractions and simplifications could yield an analytically tractable model, it is not consistent with the goals of this chapter. These complexities included here are the specialist's role, the use of large and small order sizes, and the limit and market orders.

Earlier market structure research provides useful insights, but missing institutional details prevent it from determining the effect of third markets. Garman's (1976) model of an auction market identifies a stochastic order arrival process and a market structure consistent with negative serial autocorrelations or the tendency of price changes to reverse themselves. Garman's model has no specialist role, and an analytic solution is obtained only for the case with one possible price. He notes "the practical difficulties of finding analytic solutions in the general case are considerable, and numerical techniques such as Monte Carlo methods suggest themselves." Mendelson (1987) derives analytic results that provide a comparison of market consolidation or fragmentation on market performance. The work provides numerous insights into market design trade-offs. As a simplification, however, all orders from traders are for one unit of the security.

Simulation has yielded useful results in other microstructure research. Garbade (1978) investigated the implications of interdealer brokerage (IDB) operations in a competing dealer market with a simulation model and concluded that there are benefits to investors from IDBs through reduced dispersion of quotes and transaction prices closer to the best available in the market. Cohen et al. (1985) used simulation to analyze market quality under different sets of trading priority rules. They showed that systems that consolidate orders and that maintain trading priorities by an order's time of arrival in the market increase the quality of the market. Hakannson et al. (1985) studied the market effects of alternative price-setting and own-inventory trading policies for an NYSE-style specialist dealer using simulation. They found that pricing rules "independent of the specialist's inventories break down."

Further Details on Simulation Model and Environment

Our simulation model has been used in experiments with subject-traders to test hypotheses concerning market structures. The simulation model is dynamic, with informational changes occurring in a way that creates the possibility of realizing trading profits from close attention to price changes and careful order handling. In our human-machine interactive environment, the computer generates orders from an unlimited number of "machine-resident" traders and investors. This enables us easily to satisfy the conditions for an active, competitive market.

To be useful and valid, simulation models must reflect real-world dynamics without being burdened by unnecessary real-world detail. A simulation model also requires a strong theoretical foundation. The advantage of simulation over theoretical modeling is that "theorizing" requires abstracting away from some of the very details of market structure that exchange officials and regulators wish to study. Consequently, theoretical modeling can give only limited insight into the effects of market design changes on the behavior of market participants. The advantage of simulation vis-à-vis empirically testing of new market structures is that the simulated experiments can be run at much lower cost and across a broader range of alternatives.

The objective of our computer simulation is to provide a backdrop for assessing the decisions of live participants. Trading in the model is in a single security and is the result of "machine-generated" order flow interacting with the order placement decisions of the live participants. Assumptions are made about the arrival process of investors' orders, changes to a price, p*, which is an "equilibrium value" at which the expected arrival rate of buy orders equals the expected arrival rate of sell orders.

P* follows a random walk jump process. In other words, the equilibrium value jumps randomly from one level at interarrival times based on sampling an exponential distribution. After a shift in p*, the orders of the informed traders pull the quotes and transaction prices up or down, causing them to trend towards the new p* level. Occasionally, market prices can also trend away from p* because of the orders of momentum traders or the chance arrival of a string of buy or sell liquidity orders. However, movements away from p* are unsustainable; eventually the imbalanced order flow causes a price reversal and market prices gravitate back towards p*.

Little work in experimental economics has used computers to create background order flow into which participants individually enter orders (Smith and Williams 1992; Friedman 1993). Our test environment does. We use discrete event computer simulation to do the following:

- Generate a background public order flow that can (1) be placed on a public limit order book for later execution, or (2) execute as market orders immediately against the public limit order book.
- Give the live participants the opportunity to trade a quantity of stock. Depending on the market structure, the live participants can (i) place them in a public limit order book and wait for a trade to occur or (ii) execute them against the public limit order book immediately. Variants of the simulation model include market makers, or call auctions, or dark liquidity pools to facilitate transactions.
- Maintain the screen which displays (i) orders on the public limit order book, (ii) a time-stamped record of all transaction sizes and prices for each trading session, and (iii) the users' position, risk, and profit performance data.
- Capture information concerning (i) the live participants' decisions and (ii) market quality measures such as bid-ask spreads.

In summary, with this realistic and theory-based model and the ability in the simulation to control the level of transparency provided, we have a rigorous environment to assess trading decisions and the effects of different market rules.

Components of the Market Model

In the simulation model, assumptions are made about the arrival process of investors' orders, elasticity of supply and demand, and order placement strategies, price volatility, and the proportions of market and limit orders.

Order Arrival. Orders arrive according to a price-dependent Poisson function. Using time-stamped transactions data on six stocks traded on the London Stock Exchange, a Kolmogorov-Smirnov goodness-of-fit test fails to reject the null hypothesis of exponential interarrival in 17 out of 22 sample periods at the 0.10 level of significance. We would expect to reject in just over two cases due to random realizations. The fit is not perfect in part because transactions tend to

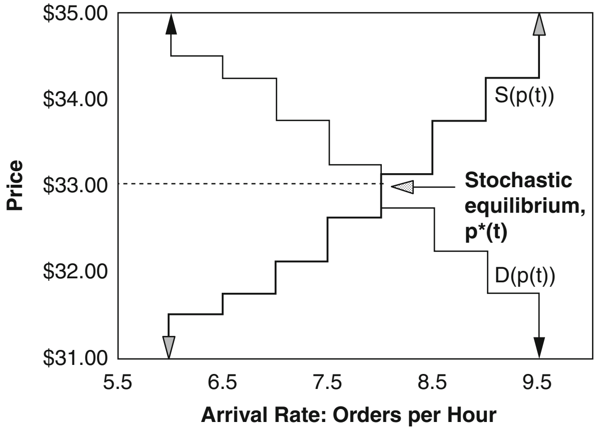


Fig. 4.10 Buy and sell order arrival rates. At market prices greater than p*, sell orders will arrive with greater intensity than buy orders. At market prices less than p*, buy orders will arrive with greater intensity than sell orders

cluster somewhat more than predicted by the theoretical model (Weber 1991). Given the shortcoming of using empirical distributions in simulations (Law and Kelton 1989), the Poisson assumption appears sufficiently justified for capturing the typical behavior of the order arrival process.

The Poisson interarrival time, T, is exponentially distributed with β_t equal to the mean interarrival time at time *t*. The mean interarrival time is set at the beginning of each experiment and assumed to hold constant. A realization at time t is thus $T_t = \Psi(\beta)$. The supply and demand structure follows closely those previously developed in the market microstructure literature (Garbade and Silber 1979), in which buy and sell arrival rates are step functions of the difference between the quoted price and the equilibrium value of the security Fig. 4.10.

Garman (1976) termed the intersection of the supply and demand functions a "stochastic equilibrium."

Demand/buy orders, D(p):

$$\begin{aligned}
&\lambda_B(p_i, p^*) = a_1 \quad \text{for} \quad {p_t}^* < p_i \\
&\lambda_B(p_i, p^*) = a_1 + a_2({p_t}^* - p_i) \quad \text{for} \quad p_i = {p_t}^* + \alpha \\
&\quad \text{with} \quad \alpha = \text{tick size}, 2(\text{tick size}), 3(\text{tick size}), \ldots \delta \\
&\lambda_B(p_i, p^*) = a_1 + a_2\delta \quad \text{for} \quad {p_t}^* - p_i > \delta
\end{aligned}$$

Supply/sell orders, S(p):

$$\begin{aligned}
&\lambda_S\left(p_j, p^*\right) = a_1 \quad \text{for} \quad {p_t}^* > p_i \\
&\lambda_S\left(p_j, p^*\right) = a_1 + a_2\left(p_j - {p_t}^*\right) \quad \text{for} \quad p_j = {p_t}^* - \alpha \\
&\quad \text{with} \quad \alpha = \text{tick size}, 2(\text{tick size}), 3(\text{tick size}), \ldots \delta \\
&\lambda_S\left(p_j, p^*\right) = a_1 + a_2\delta \quad \text{for} \quad p_j - {p_t}^* > \delta
\end{aligned}$$

The constant a_1 reflects the proportion of arrivals that are market orders. The coefficient a_2 determines the arrival rate of limit orders with reservation prices. Limit order traders are sensitive to discrepancies between available prices and the equilibrium value. The parameter, δ, is the range around the equilibrium value from which limit prices for limit orders are generated. At a price p_i lower than the equilibrium value at the time, p_t^*, the arrival rate of buy orders will exceed the rate of sell order arrivals. The resulting market buy orders and limit order bids will exceed the quantity of sell orders for below-equilibrium values. The arrival rate discrepancy will cause prices to rise since in expectation, orders will trade against the lowest offer quotes, and add new, higher priced bid quotes.

Order Size. Orders are between one and 250 units of the security. This reflects a convenient normalization that is consistent with the empirically observable range of order sizes. A unit may represent, for instance, three round lots, or 300 shares. Beyond 250 units, we assume the trade would be handled as a block trade, and negotiated outside of the standard market design, or arrive in the market in smaller broken-up pieces. Large orders can have "market impact," and can move prices up for large buyers, and force them down for larger sellers. The functioning of the market for large orders is consistent with observed trade discounts for large sell orders and premiums for large buy orders.

Order Placement Strategies. The machine-generated order flow consists of liquidity, informed, and momentum trading orders. The liquidity orders are either limited price orders or market (immediately executable) orders. Market orders execute on arrival but are "priced" to reflect a maximum acceptable premium or discount to the current bid or offer. If the market order is large enough, its price impact (the need to hit successive lower priced bids or lift higher priced offers) will exceed the acceptable discount or premium, and the remaining order quantity will become a limit order after partially executing against the limit order book.

Information Generation. Idiosyncratic information events occur that change the share value, p*, at which buying and selling order arrival rates are balanced. Information event occurs according to a Poisson arrival process. When an information innovation occurs, the price will have a random walk jump.

Price Random Walk. Idiosyncratic information events occur that change the share value, p*, at which the arrival rate of buy orders equals the arrival rate of sell orders. The time between information change is assumed to be exponentially distributed with mean, 12 h. Empirical validation is difficult, because informatio-n affects the share values in unobservable ways. When there is a change in information that will shift the "balance price," p* evolves according to a random walk without return drift. To assure nonnegative prices, the natural log of price is used, yielding a log-normal distribution for the equilibrium price. The white noise term, et, is normally distributed with variance linear in the time since the last observation. This is consistent with the price diffusion models used in the financial economics literature (Cox and Rubinstein 1985):

$\ln p_t * = \ln p_t - T * + e_t$ where, $et \sim N(0, T\sigma^2)$
where $p_t * \sim LN(\ln pt - T *, T\sigma^2)$

The natural logarithm of the current price is an unbiased estimator of the natural logarithm of any subsequent price:

$$E(\ln\ p_t + T^* \mid pt^*) = \ln \quad p_t *$$

Empirical validation for the random walk model comes from numerous tests, whose results "are remarkably consistent in their general finding of randomness . . . serial correlations are found to be small" (Malkiel 1987).

Information Effects. If the bid and offer quotes straddle p*, there is no informed order flow and buying and selling order arrival rates will be equal. When p* is outside of the bid-offer range, additional one-sided market orders will be generated according to a Poisson process.

References

Clemons, E., & Weber, B. (1997). Information technology and screen-based securities trading: Pricing the stock and pricing the trade. *Management Science, 43*(12), 1693–1708.

Cohen, K., Conroy, R., & Maier, S. (1985). Order flow and the quality of the market. In Y. Amihud, T. Ho, & R. Schwartz (Eds.), *Market making and the changing structure of the securities industry* (pp. 93–109). Lanham, MD: Lexington Books.

Cox, J., & Rubinstein, M. (1985). *Options markets*. Englewood Cliffs: Prentice-Hall.

Francioni, R., Hazarika, S., Schwartz, R., & Reck, M. (2010). Security market microstructure: The analysis of a non-frictionless market. In C. F. Lee & A. C Lee (Eds.), *Handbook of quantitative finance and risk management* (pp. 333–353). Germany: Springer New York & Heidelberg.

Friedman, D. (1993). How trading institutions affect financial market performance: Some laboratory evidence. *Economic Inquiry, 31*, 410–435.

Garbade, K. (1978). The effect of interdealer brokerage on the transactional characteristics of dealer markets. *Journal of Business, 51*(3), 477–498.

Garbade, K., & Silber, W. (1979). Structural organization of secondary markets: clearing frequency, dealer activity and liquidity risk. *The Journal of Finance, 34*, 577–593.

Garman, M. (1976). Market microstructure. *Journal of Financial Economics, 3*, 257–275.

Greene, W. H. (2011). *Econometric analysis* (7th ed.). Upper Saddle River, NJ: Prentice Hall.

Hakansson, N., Beja, A., & Kale, J. (1985). On the feasibility of automated market making by a programmed specialist. *Journal of Finance, 40*, 1–20.

Ho, T., Schwartz, R., & Whitcomb, D. (1985). The trading decision and market clearing under transaction price uncertainty. *Journal of Finance, 40*, 21–42.

Kennedy, P. (2008). *A guide to econometrics* (6th ed.). Hoboken, New Jersey: Wiley-Blackwell.

Law, A., & Kelton, W. (1989) *Simulation modeling and analysis*, McGraw-Hill, New York, NY.

Malkiel, B. (1987). Efficient market hypothesis. In B. Malkiel (Ed.), *The New Palgrave: A dictionary of economics* (pp. 127–134). London: Macmillan Press.

Mendelson, H. (1987) Consolidation, Fragmentation, and Market Performance, *Journal of Financial and Quantitative Analysis, 22*, 189–207.

Parker, C., & Weber, B. (2012). *How IT can disrupt markets: A simulation analysis* (Working Paper). London: London Business School.

Schwartz, R., & Weber, B. (1997). Next-generation securities market systems: An experimental investigation of quote-driven and order-driven trading. *Journal of Management Information Systems, 14*, 57–79.

Smith, V., & Williams, A. (1992, December). Experimental market economics. *Scientific American*, pp. 116–121.

Weber, B. (1991) *Information technology and securities markets: Feasibility and desirability of alternative electronic trading systems*, Unpublished dissertation, University of Pennsylvania.

Zhang, S. S., Storkenmaier, A., Wagener, M., & Weinhardt, C. (2011). The quality of electronic markets. In *Forty Fourth Hawaii International Conference on System Sciences, Kauai*, Hawaii, pp. 1–10.

Zhang, X. F. (2010). *High frequency trading, stock volatility, and price discovery* (Working Paper). New Haven, CT: Yale University.

Motivations for Issuing Putable Debt: An Empirical Analysis

5

Ivan E. Brick, Oded Palmon, and Dilip K. Patro

Contents

Abstract

This paper examines the motivations for issuing putable bonds in which the embedded put option is not contingent upon a company-related event. We find

We would like to thank Sanjay Deshmukh and conference participants at the Seattle FMA meetings, October 2000, for helpful comments on an earlier version of the paper circulated under a different title. The views expressed in this paper are strictly that of the authors and not of the OCC or the Comptroller of the Currency.

I.E. Brick (✉)
Department of Finance and Economics, Rutgers, The State University of New Jersey, Newark/New Brunswick, NJ, USA
e-mail: ibrick@business.rutgers.edu; ibrick@andromeda.rutgers.edu

O. Palmon
Department of Finance and Economics, Rutgers Business School Newark and New Brunswick, Piscataway, NJ, USA
e-mail: palmon@business.rutgers.edu

D.K. Patro
RAD, Office of the Comptroller of the Currency, Washington, DC, USA
e-mail: dilip.patro@occ.treas.gov

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_5,

that the market favorably views the issue announcement of these bonds that we refer to as bonds with European put options or European putable bonds. This response is in contrast to the response documented by the literature to other bond issues (straight, convertible, and most studies examining poison puts) and to the response documented in the current paper to the issue announcements of poison put bonds. Our results suggest that the market views issuing European putable bonds as helping mitigate security mispricing. Our study is an application of important statistical methods in corporate finance, namely, *event studies* and the use of general method of moments for cross-sectional regressions.

Keywords

Agency costs • Asymmetric information • Corporate finance • Capital structure • Event study methodology • European put • General method of moments • Management myopia • Management entrenchment • Poison put

5.1 Introduction

This paper examines the motivations for issuing putable bonds in which the embedded option is not contingent upon company-related events. The option entitles bondholders to sell the bond back to the firm on the exercise date (usually 3–10 years after the bond is issued) at a predetermined price (usually at par). We refer to these bonds as bonds with European put options or European putable bonds. Unlike the poison put bonds (i.e., bonds with event risk covenants) that have been studied by the literature,[1] the exercise of the option in a putable bond is **not** contingent upon a company-related event. This distinction is important because a poison put protects the bondholder from a specific event (e.g., a takeover) and may be designed to help prevent that event. In contrast, putable debt provides protection to the bondholder against any deterioration in the value of her claim. This distinction is important also because the two types of embedded put options may serve different purposes. Crabbe and Nikoulis (1997) provide a good overview of the putable bond market.

Corporate managers determine which contingent claims the company issues to finance its activities. This choice includes the debt-equity mix and the specific design of the debt. The design of the debt includes its maturity, seniority, collateral, and the type of embedded options included in the bond contract. The theoretical corporate finance literature indicates that including convertible, callable, and/or putable bonds in the capital structure may help mitigate agency costs and reduce asymmetric information. Haugen and Senbet (1981) theoretically demonstrate that the optimal combination of embedded call and put options should eliminate the

[1]See for example, Crabbe (1991), Bae et al. (1994, 1997), Cook and Easterwood (1994), and Roth and McDonald (1999), Nash et al. (2003). Billett et al. (2007) find that 5 % of corporate bonds of their sample have a non-poison putable option.

asset substitution problem. Bodie and Taggart (1978) and Barnea et al. (1980) show that bonds with call options can mitigate the underinvestment problem. Robbins and Schatzberg (1986) demonstrate that the increased interest cost of callable bond can be used to convey the true value of the firm to the market.[2]

The literature has not empirically examined the motivation and equity valuation impact of issuing debt with European put options. We fill this gap in the literature in several ways. First, we examine the stock price reaction around the announcement dates of the two types of putable bond issues: bonds with European put options and bonds with poison puts. Further, we consider four alternative motivations for incorporating either a European or a poison put option in a bond contract and test the hypotheses that putable debt is issued to reduce security mispricing, agency costs of debt, management entrenchment, and myopia. The first possible motivation for issuing putable debt is asymmetric information. Consider a company that is undervalued by the market. The market should also undervalue its risky straight-debt issue. However, if this firm were to issue putable bonds, the put option would be overvalued. (Recall that put options are negatively related to the value of the underlying asset.) Consequently, the market value of bonds with a put option is less sensitive to asymmetric information than straight bonds, minimizing market mispricing of these debt securities. Additionally, if the market overestimates the risk of the bond, it may overvalue the embedded put option at issuance, thereby increasing bond proceeds and benefiting the shareholders.[3] The second possible motivation is mitigating agency costs of debt. For example, the existence of a put option mitigates the advantage to stockholders (and the loss to bondholders) from risk shifting, thereby reducing the incentive to shift risk. Risk shifting hurts putable bondholders less than holders of straight bonds because the value of the put option (which is held by bondholders) is an increasing function of the firm's risk. The third possible motivation is the relatively low coupon rate (a myopic view that ignores the potential liability to the firm due to the put option). The fourth possible motivation is that the put option may serve to entrench management. This fourth motivation is most relevant for firms issuing poison puts since these put options are exercisable contingent upon company-related events that are usually related to a change in ownership.

We find that the market reacts favorably to the issue announcement of European put bonds. We also examine the relationship between the abnormal returns around the put issue announcement date and firm characteristics that proxy for asymmetric information problems, potential agency costs (i.e., risk-shifting ability and the level of free cash flow), and management myopia. The empirical evidence is consistent with the view that the market considers issuing putable bonds as mitigating security mispricing caused by asymmetric information. The results do not support the idea

[2]Other benefits posited by the literature include minimizing tax liabilities. See for example, Brick and Wallingford (1985), and Brick and Palmon (1993).

[3]Brennan and Schwartz (1988) offer a similar argument to explain the benefits of issuing convertible bonds.

that putable bonds are issued to obtain lower coupon rates (i.e., management myopia). Our empirical findings are robust to a number of alternate specifications.

In contrast to European putable bonds, and consistent with the management entrenchment hypothesis (see Cook and Easterwood (1994) and Roth and McDonald (1999)), we find that the market reacts unfavorably to the issue announcement of poison put bonds. However, consistent with Bae et al. (1994) who argue that poison put bonds are useful in mitigating agency cost problems, we find that the abnormal returns around the issue announcement of poison put bonds are positively related to the protection level of the event risk covenant. Thus, our results are consistent with the view that European put bonds are effective in mitigating security mispricing problems, but, in contrast, poison put bonds are related to management entrenchment or mitigating agency cost problems.

The paper's organization is as follows. In the next section we summarize the previous literature. Section 5.3 develops the empirical hypotheses. We describe the data and empirical methodology in Sect. 5.4. The empirical results are summarized in Sect. 5.5. We offer concluding remarks in Sect. 5.6.

5.2 Literature Review

The theoretical corporate finance literature concludes that the firm's financing decision may affect its equity value for several reasons. First, issuing debt increases firm value because it decreases its tax liabilities.[4] Second, issuing debt may be, in part, a signaling mechanism that informs the market of private information. For example, Ross (1977) and Ravid and Sarig (1991) demonstrate that the manager of a firm with better prospects than the market perceives has an incentive to signal her firm's quality by issuing a greater amount of debt than issued in a symmetric information environment. Third, prudent level of debt can reduce the agency costs arising from the conflict of interest between managers and shareholders as demonstrated by Jensen and Meckling (1976). For example, Jensen (1986) demonstrates that leverage can minimize the deleterious effect of free cash flow on the firm.[5] However, leverage is also shown to generate other

[4]See, for example, Modigliani and Miller (1963), Scott (1976) and Kim (1978). In contrast, Miller (1977) suggests that the tax benefit of interest is marginal. However, Mackie-Mason (1990) empirically demonstrates the significant impact of corporate taxes upon the observed finance choices of firms.

[5]Hence, empirically, we would expect that as firms announce increased levels of debt, the stock price should increase. However, studies by Dann and Mikkelson (1984), Mikkelson and Partch (1986), Eckbo (1986), and Shyam-Sunder (1991) indicate that there is no systematic relationship between the announcement of firm's debt financing and its stock price or that this relationship is weakly negative. One potential explanation for this result is that the market can predict future debt offerings as argued by Hansen and Chaplinsky (1993). Another potential explanation, as suggested by Miller and Rock (1985) and documented by Hansen and Crutchley (1990), is that raising external capital may indicate a cash shortfall.

agency problems because of the conflict of interest between stockholders and bondholders. These agency problems include underinvestment and risk shifting (or asset substitution).[6]

Other studies indicate that these agency problems can be mitigated or eliminated by optimal security design. For example, Haugen and Senbet (1981) demonstrate that the optimal combination of embedded call and put options completely eliminates the asset substitution problem. Bodie and Taggart (1978) and Barnea et al. (1980) show that the call options of bonds can mitigate the underinvestment problem. Similarly, Green (1984) demonstrates that including convertible bonds in the capital structure may also mitigate the underinvestment problem.

The literature has also demonstrated that an appropriate debt security design may alleviate asymmetric information problems. For example, Robbins and Schatzberg (1986) demonstrate that the increased interest cost of callable bond can be used to signal the value of the firm. Moreover, as similarly stated for the case of convertible bonds by Brennan and Schwartz (1988), the inclusion of a put option can minimize the mispricing of debt securities if the market overestimates the risk of default. Chatfield and Moyer (1986) find that putable bonds may be issued by financial institutions for asset-liability management in a period of volatile interest rates. Tewari and Ramanlal (2010) find that callable-putable bonds provide protection to bondholders and improved returns to stockholders.

The literature examines the equity valuation impact of a special type of putable bond known as the poison put bond. In poison put bonds, the put option is exercisable contingent upon a company-related event such as a leveraged restructuring, takeover, or downgrading the debt credit rating to speculative grade. David (2001) shows that puts may have higher strategic value than intrinsic value. Crabbe (1991) finds that such event-related covenant put bonds reduce the cost of borrowing to the firm. Bae et al. (1994) conclude that stockholders benefit from the inclusion of event-related (risk) covenants. Furthermore, Bae et al. (1997) empirically document that the likelihood of a firm to include event risk covenants is positively related to the firm's agency costs of debt. In contrast, Cook and Easterwood (1994) and Roth and McDonald (1999) find that the inclusion of poison put bonds benefits both management and bondholders at the expense of stockholders.

In contrast to these studies of poison put options, our study examines debt issues in which the embedded put option is equivalent to a European put with a fixed exercise date that is usually 3–10 years after the issue date. That is, bondholders may exercise the put option only at the exercise date, and their ability to do so is not contingent upon any particular company-related event. Thus, we believe that issuing bonds with an embedded put option that is not contingent upon a company-related event (such as a change in ownership) is not likely to be motivated by

[6]Myers (1977) demonstrates that shareholders may avoid some profitable net present value projects because the benefits accrue to the bondholders. Jensen and Meckling (1976) demonstrate that leverage increases the incentives for managers, acting as agents of the stockholders, to increase the risk of the firm.

management entrenchment. Consequently, we consider the security mispricing motivation and two other possible motivations: mitigating agency costs and the myopic behavior on the part of the management.

5.3 Hypotheses

In this section, we describe four alternative motivations for incorporating a put option in a bond contract and outline their empirical implications.

The first motivation is **reduction in the level of security mispricing due to asymmetric information**. Consider a company that is undervalued by the market.[7] The market should also undervalue its risky straight-debt issue. However, if this firm were to issue putable bonds, the put option would be overvalued. Consequently, the market value of bonds with put option is less sensitive to asymmetric information than the market value of straight bonds, minimizing market mispricing of these debt securities. Additionally, the overvaluation of the implicit put option increases the debt proceeds at the issuance thereby benefiting the shareholders. This possible motivation has the following empirical implications. *First*, issuing bonds with put option should be associated with an increase in equity value. *Second*, this increase in firm value should be negatively related to the accuracy with which the firm's value has been estimated prior to the bond issue. *Third*, because the value of the put option is directly related to the magnitude of firm undervaluation, the increase in equity value should be directly related to the value of the put option.

The second possible motivation is **mitigating the bondholder-stockholder agency costs**.[8] The value of a put option is an increasing function of the firm's risk. Thus, its existence mitigates the gains to stockholders from, and hence their incentives for, risk shifting. This possible motivation has the following empirical implications. *First*, the benefit to a firm from incorporating a put option in the bond contract should be directly related to the firm's ability to shift risk in a way that increases the value of stockholders claims at the expense of bondholders. *Second*, the inclusion of a put option should help restrain management from taking on negative net present value projects in the presence of Jensen's (1986) free cash flow problem.[9] In the absence of a put option, undertaking negative net present value projects should reduce security prices for both stockholders and bondholders.

[7]We assume that managers of companies that are overvalued have no incentive to resolve security mispricing. Consequently, managers of undervalued firms could send a credible signal to the market of the firm's undervaluation with the inclusion of a put feature. Overvalued firms should not be able to mimic since the put option represents a potential liability that is greater to these firms than of the undervalued firms.

[8]The bondholders-stockholders agency conflict has been found to be important in security design by Bodie and Taggart (1978), Barnea et al. (1980), Haugen and Senbet (1981), Jung et al. (1996), and Lewis et al. (1998).

[9]Equivalently, the inclusion of a put option should restrain management from consuming a suboptimal amount of perquisites or nonpecuniary benefits.

In contrast, giving bondholders the option to put the bond back to the firm at face value shifts more of the negative price impact of undertaking negative net present value projects to the stockholders. This negative stock price impact may reduce management compensation that is tied to equity performance and/or induce tender offers that will ultimately help replace the current management. These increased costs to stockholders and management should induce management to refrain from undertaking negative net present value projects. Thus, we hypothesize that the benefit to the stockholders from incorporating a put option in the bond contract should be directly related to the level of free cash flow. *Third*, the gain in firm value should be related to the magnitude of the agency cost problem that the firm faces. The valuation of the put option by the market (assuming market efficiency and symmetric information) should be positively related to the magnitude of the agency cost problem faced by the company. Thus, the benefit to the firm should also be directly related to the aggregate value of the implied put option of the issue (scaled by firm size). *Fourth*, for our sample of poison put bonds, the benefit to shareholders should increase with the strength of the event risk covenant, since the greater the strength of the event risk covenant, the less likely management will engage in value decreasing activities. Bae et al. (1994) tested a similar set of hypotheses for a sample of poison put bonds.[10]

The third possible motivation is the low coupon rate (compared to the coupon rates of straight or callable debt issues). This motivation reflects **management myopia** as it ignores the potential liability to the firm due to the put option.[11] That is, myopic management may not fully comprehend the increased risk to the firm's viability that is posed by the put option. In particular, bondholders would have the right to force the firm to (prematurely) retire its debt at a time that is most inconvenient to the firm, which in turn can precipitate a financial crisis. Further, if the cost of financial distress is significant, given rational markets and myopic management, issuing putable debt may negatively impact the value of equity. This possible motivation has the following empirical implications. *First*, because management myopia implies that management pursues suboptimal policies, the issuance of putable debt should be associated with a decline in equity value.[12] *Second*, the decline should be more severe the larger is the aggregate value of the implied put option. *Third*, because expected costs of financial distress are

[10]Although bonds become due following formal default, hence all bonds in a sense become putable, bondholders usually recover substantially less than face value following formal default because equity value has already vanished. In contrast, a formal inclusion of put option may allow bondholders to recover the face value at the expense of shareholders when financial distress is not imminent but credit deterioration has occurred.

[11]Other studies that have examined managerial myopia include Stein (1988), Meulbroek et al. (1990), Spiegel and Wilkie (1996), and Wahal and McConnell (2000).

[12]Haugen and Senbet (1978, 1988) theoretically demonstrate that the organizational costs of bankruptcy are economically insignificant. However, if the putable bonds increase the potential of technical insolvency, then it will increase potential agency costs that are not necessarily insignificant.

negatively related to the financial stability of the company, the decline should be more severe the lower the credit rating of the bond.

The fourth possible motivation is that the use of put options enhances **management entrenchment**. This hypothesis is relevant for firms issuing poison but not European put bonds. In particular, many bonds with event risk covenants (i.e., poison puts) place restrictions on the merger and acquisition activity of the issuing firms, thereby strengthening the hands of management to resist hostile takeover bids. This possible motivation has the following empirical implications. *First*, issuing bonds with put options should be followed by a decrease in equity value. *Second*, in contrast to the **mitigating agency costs hypothesis**, the abnormal returns of equity around the issue announcement date should be inversely related to the level of event-related covenant protection offered to the bondholder.

5.4 Data and Methodology

Our sample of bonds with put options is taken from the *Warga Fixed Income Database*. The database includes pricing information on bonds included in the Lehman Brothers Bond Indices from January 1973. Using this database we select all bonds with fixed exercise date put option (i.e., not event contingent), issued between January 1973 and December 1996, excluding issues by public utilities, multilateral agencies (such as the World Bank), and sovereign issues. In our sample we keep only those issues for which we find announcement dates in the *Dow Jones News Retrieval Service*. This resulted in a sample of 158 bonds of which the earliest bond is issued in 1979. Our final sample of bonds is selected according to the following criteria:

(a) The issuing company's stock returns are available on the CRSP tapes. For 20 companies CRSP data are not available, resulting in 138 issues. To reduce confounding effects, all repeat issues by the same company within a year of a sample issue are eliminated. We also eliminated observations for which we find other contemporaneous corporate events. This further reduced our sample to 104 issues.
(b) Furthermore, we eliminate from the sample companies that do not have sufficient accounting data, credit rating, issue size, and industry code in either the *Compustat* tapes, Moody's *Industrial Manuals*, or the Warga's *Fixed Income* database. This reduces our sample size by 13 firms.
(c) We eliminate one more company for which we found no I/B/E/S data, thus yielding a final sample of 90 firms.

The list of these firms and the characteristics of these bonds are reported in Appendix 1. The maturity of these put bonds ranges from 5 to 100 years, with an average initial maturity of 24 years. The put exercise dates ranges from 2 to 30 years, with an average put expiration period of 7 years.[13]

[13]Four bonds have multiple put exercise dates.

For the sake of comparison, we also construct a sample poison put bonds.[14] The issue announcements of poison put bonds are taken from *Dow Jones Interactive*. We also searched *LexisNexis* but did not find any new announcements.[15] This search resulted in 67 observations. Our final sample of bonds with poison put feature is further refined using criteria (a), (b), and (c) described above. Criterion (a) reduces the sample size to 57 observations, and applying criterion (b) further reduces it to 47 observations. The list of these firms and the characteristics of these bonds are reported in the Appendix 2.

For these issues, we collect *CRSP* daily returns for our event study.[16] We calculate the abnormal returns (*AR*) on equity for each firm around the date of the issue announcement, assuming that the stochastic process of returns is generated by the market model. We define the announcement date to be the earlier of the date on which the bond is issued and the date on which the issue information appears in the news wires, or published in the *Wall Street Journal*, as depicted by the *Dow Jones Interactive*. Our final sample includes only bonds whose issue announcement explicitly mentions the inclusion of a put option. We estimate the market model coefficients using the time period that begins 200 trading days before and ends 31 trading days before the event, employing the CRSP value-weighted market index as the benchmark portfolio. We use these coefficients to estimate abnormal returns for days −30 to +30. We calculate the t-statistics for the significance of the abnormal and cumulative abnormal returns using the methodology employed by Mikkelson and Partch (1986). See Appendix 3 for details.

We examine the abnormal returns and their determinants for the sample of issuers of bonds with European put options. First, we test whether issuing these bonds significantly affects equity values. The stock price should, on average, react positively to the issue announcement if either reducing debt security mispricing due to asymmetric information or mitigating agency costs is a major motivation for issuing putable bonds. On the other hand, the stock price should, on average, react negatively to the issue announcement if management myopia (i.e., the relatively low coupon rate compared to straight or callable debt issues) is a major motivation of management in issuing putable bonds and if costs of financial distress are significant.

Second, we estimate the following cross-sectional regression equation that relates the cumulative abnormal return of the issuing firm's equity to firm

[14]Unlike the European put bond sample which we were able to obtain from *Warga's Fixed Income Database*, we obtained our sample of poison puts from *Dow Jones Interactive* by using keywords such as poison put, event risk, and covenant ranking. Warga's database does not include an identifiable sample of poison put bonds.

[15]We were only able to find a sample of poison put bonds with issue announcements using *Dow Jones Interactive* and *LexisNexis* for the period between 1986 until 1991. We did use these news services for the periods between 1979 and 2000.

[16]The event study methodology was pioneered by Fama et al. (1969).

characteristics that proxy for agency costs, asymmetric information problems, and managerial myopia:

$$\begin{aligned} CAR3_i = \beta_0 + \beta_1 FCF_i + \beta_2 RISK_i + \beta_3 SIZE_i + \beta_4 INTSAVED_i \\ + \beta_5 ANALYSTS_i + \beta_6 FINS_i + \varepsilon_i, \end{aligned} \tag{5.1}$$

where $CAR3_i$ is the 3-day (i.e., $t = -1, 1$) cumulative abnormal return for firm I. The variables, *FCF*, *RISK*, *SIZE*, *INTSAVED*, and *ANALYSTS* proxy for agency costs, information asymmetry, or management myopia. *FINS* is a dummy variable that indicates whether the company is a financial institution.

FCF is the level of free cash flow of the firm for the fiscal year prior to the issue announcement of putable bonds. We construct two alternative measures of *FCF* which are similar to the definition employed by Lehn and Poulsen (1989) and Bae et al. (1994). *FCF1* is defined as [Earnings before Interest and Taxes – Taxes – Interest Expense – Preferred Dividend Payments – Common Stock Dividend Payments]/Total Assets. *FCF2* is defined as [Earnings before Interest and Taxes – Taxes – Interest Expense]/Total Assets.[17] The inclusion of a put option in the bond contract should help restrain management from misusing its cash flow resources. We therefore expect β_1 to be positive if reducing agency cost is a major motivation for issuing putable bonds.

RISK is a dummy variable that equals 1 if the putable bond is rated by **Standard and Poor's** as BBB+ or below and zero otherwise.[18] The impact of issuing putable bonds on equity value may be associated with the variable *RISK* because of two alternative hypotheses. First, the level of agency costs due to the asset substitution problem is positively related to the probability of financial distress. We use this dummy variable as a proxy for the firm's ability and incentive to shift risk. According to the agency cost motivation, we expect β_2 to be positive. Second, the greater the probability of financial distress, the more likely are bondholders to exercise their put option and force the firm to prematurely retire the debt at a time that is most inconvenient to the firm. Thus, according to the management myopia hypothesis, we expect β_2 to be negative.

SIZE is the natural logarithm of the total asset level of the issuing firm at the end of the fiscal year preceding the issue announcement date. *SIZE* may have two contrasting impacts on *CAR3*. First, *SIZE* may be interpreted as a proxy for the level of asymmetric information. We expect that the degree of asymmetric information is inversely related to *SIZE*. Hence, according to the debt security mispricing motivation, we expect the putable bond issue announcement to have a larger positive impact on small firms than on large firms. Consequently, we expect

[17]Note that *FCF1* is the actual free cash flow of the firm while *FCF2* is the (maximum) potential free cash flow if dividends are not paid. We do not subtract capital expenditures from the free cash flow variables because they maybe discretionary and maybe allocated suboptimally to satisfy management's interests.

[18]Had we assigned the value of 1 to only junk bonds, the number of observations with the risk variable equal to one would be too small for statistical inference.

β_3 to be negative. On the other hand, we also expect that the risk of default is inversely related to firm size. The higher the probability of default, the lower the probability that the firm will be in existence and be able to pay its obligations (including the par value of the bond if the put option is exercised) at the exercise date of these European put options. Thus, size may be an indirect proxy of the aggregate value of the put option. In this case, we expect β_3 to be positive.

INTSAVED is the scaled (per $100 of assets for the fiscal year prior to the issue announcement date) annual reduction in interest expense due to the incorporation of a European put option in the bond contract. The annual interest expense reduction is calculated as the product of the dollar amount of the putable bond issue and the yield difference between straight and putable bonds. These yield differences are calculated by subtracting the yields to maturity of the putable bond from the yield to maturity of an equivalent non-putable bond, also taken from Warga's *Fixed Income Database*. The equivalent bond has similar maturity, credit rating, and call feature as the putable bond. The equivalent non-putable bond is selected from the issuing firm if, at the issue date of the putable debt, the issuing firm has an appropriate (in terms of maturity, credit rating, and call feature) outstanding non-putable bond. Otherwise, we choose an appropriate equivalent bond from a firm of the same industry code as given by the Warga's *Fixed Income* database. We posit that value of the embedded put option is directly related to the (present) value of *INTSAVED*. If the benefits of mitigating either agency costs or security mispricing are a major reason for issuing European put bonds, and if these benefits are directly related to the value of the embedded option, then β_4 should be positive. In contrast, if managerial myopia is a major reason for issuing European put bonds, and if the expected reduction in equity value due to the cost of financial distress is related to the value of the option, then β_4 should be negative. The β_4 coefficient may be negative also because it proxies for a loss of interest tax shield due issuing putable rather than straight bonds.

The variable *ANALYSTS*, defined as the natural logarithm of one plus the number of analysts who follow the issuing firm for the quarter prior to the putable bond issue announcement date, is another proxy for the degree of asymmetric information.[19] We obtain the number of analysts following each firm, for the year prior to the bond announcement, from the *I/B/E/S* tapes. We hypothesize that the degree of asymmetric information is negatively related to *ANALYSTS*. Thus, if asymmetric information motivation is a major motivation for issuing putable debt, we expect β_5 to be negative.

We note that the European put and the poison put samples vary in their proportion of financial service companies. The European put sample contains 25 (out of 90) financial service companies, while the poison put contains only one. To control for the potential sector impact, we introduce the dummy variable *FINS* which equals one if the company is a financial institution and zero otherwise

[19]An alternative measure of the degree of asymmetric information is the number of analysts. The empirical results are robust to this alternative measure.

Table 5.1 A summary of expected signs of abnormal returns and their determinants for issuers of bonds with European put options

	Agency cost	Security mispricing	Managerial myopia
Abnormal returns	***Positive***	***Positive***	*Negative*
FCF	*Positive*	*No prediction*	*No prediction*
RISK	*Positive*	*No prediction*	*Negative*
SIZE	*No prediction*	*Ambiguous*	*No prediction*
INTSAVED	***Positive***	***Positive***	*Negative*
ANALYSTS	*No prediction*	***Negative***	*No prediction*

The determinants of abnormal returns are: *RISK* is a dummy variable equal to one if the bond issue has an S&P bond rating of BBB+ or below and zero otherwise. *FCF* is defined in two separate ways. *FCF1* and *FCF2* are the two free cash flow measures as described in the text. In particular, *FCF1* is defined as [Earnings before Interest and Taxes – Taxes – Interest Expense – Preferred Dividend Payments – Common Stock Dividend Payments]/Total Assets. *FCF2* is defined as [Earnings before Interest and Taxes – Taxes – Interest Expense]/Total Assets. *INTSAVED* measures the relative amount of aggregate interest expense saved per $1,000 of total assets of the issuing firm. *ANALYSTS* is the natural logarithm of one plus the number of analysts who follow the issuing firm for the year prior to the putable bond issue announcement date. Those predictions that are confirmed by our empirical study are in bold letters, and those that are significantly different from zero are also underlined

for our regression analysis of the European put sample.[20] Table 5.1 provides a summary of the expected signs of the regression coefficients for the sample of European put bonds.

We estimate our regressions using the general method of moments (GMM) procedure. See Appendix 4. The procedure yields unbiased White t-statistics estimates that are robust to heteroscedasticity.[21] Note that because our instruments are the regressors themselves, the parameter estimates from OLS and GMM are identical. This is also discussed in Appendix 3.

We test the robustness and appropriateness of our specification by estimating alternative specifications that include additional variables. First, we include a dummy variable for the existence of a call feature to take into account the effectiveness of the call feature to mitigate agency costs. Second, we include a dummy variable to indicate that the bond's put option expires within 5 years because effectiveness of the put bond in mitigating agency costs may be related to

[20] In the next section, we report the regression results when we exclude financial service companies from our European put sample. Essentially, the basic results of our paper are not affected by the inclusion or exclusion of financial services company.

[21] The violation of the homoscedastic assumption for OLS does not lead to biased regression coefficients estimators but potentially biases the computed t-statistic. The GMM procedure provides an asymptotically unbiased estimation of the t-statistics without specifying the heteroscedastic structure of the regression equation. The t-statistics obtained using GMM are identical to those obtained using ordinary least squares (OLS) in the absence of heteroscedasticity.

time to expiration of the put option. Third, we also include interaction variables between these dummy variables and *FCF*, *SPRISK*, *INTSAVE*, and *ANALYSTS*.

In alternative specifications, we include interest rate volatility measures to take into account the potential sensitivity of the value of the put option to interest rate volatility. It should be noted that the value of the put option depends on the volatility of the corporate bond yield. This volatility may be due to factors specific to the company (such as corporate mismanagement and asymmetric information about firm's value) and macroeconomic factors relating to the stability and level of market-wide default-free interest rates. Our independent variables in Eqs. 5.1 and 5.2 control for firm-specific yield volatility. To incorporate the possible impact of market-wide interest rate volatility, we include several alternative variables. We measure interest rate volatility as the standard deviation of the monthly 5-year Fama-Bliss discount rate from CRSP. The standard deviation is alternatively measured during a period of 60 months and 24 months immediately prior and after the announcement date. Alternatively, we use the level of the Fama-Bliss discount rate as a proxy for interest rate volatility.

We repeat the analysis for the sample of issuers of poison put bonds. As with the sample of bonds with an embedded European put option, we first examine the abnormal returns and their determinants for the sample of issuers of poison putable bonds. In essence, we test whether issuing these bonds significantly affects equity value. The stock price should, on average, react positively to the issue announcement if either mitigating agency costs or reducing debt security mispricing due to asymmetric information is a major motivation for issuing poison bonds. On the other hand, the stock price should, on average, react negatively to the issue announcement if management myopia (i.e., the relatively low coupon rate compared to straight or callable debt issues) or management entrenchment is a major motivation of management in issuing putable bonds.

We use two alternative specifications for the cross-sectional study that relates the abnormal returns of the poison putable bond sample to firm characteristics. The first is described by Eq. 5.1 The second includes a variable, COVRANK, that equals the S&P event risk ranking on a scale of 1–5. S&P event risk ranking of one implies that the embedded put option provides the most protection to bondholders against credit downgrade events. Event risk ranking of five offers the least protection to bondholders against credit downgrade events. Thus, the second specification is:

$$CAR3_i = \beta_0 + \beta_1 FCF_i + \beta_2 RISK_i + \beta_3 SIZE_i + \beta_4 INTSAVED_i + \beta_5 ANALYSTS_i + \beta_6 COVRANK_i + \varepsilon_{i.} \tag{5.2}$$

If mitigating agency costs was a major motivation for issuing poison put bonds as suggested by Bae et al. (1994), then we would expect regression coefficients β_1, β_2, and β_4 to be positive. We also expect β_6 to be negative because COVRANK is negatively related to extent of bondholder protection, and this protection should deter management from asset substitution activities.

In contrast, if, as Cook and Easterwood (1994) and Roth and McDonald (1999) argue, management entrenchment is a major motivation behind for issuing putable bonds, then we expect β_1 to be negative since we expect the potential loss of value due to management entrenchment to be positively related to free cash flow and β_6 to be positive.[22]

5.5 Empirical Results

This section presents the empirical results for tests of our hypotheses discussed above. Panel A of Table 5.2 provides summary statistics of issuers of bonds with a European put option. The average annual sales, long-term debt, and total assets for the fiscal year prior to the issue announcement date are \$14.66 billion, \$4.21 billion, and \$30.79 billion, respectively. The mean leverage ratio, defined as the long-term debt to total assets, is 18.6 %. The average issue size of the putable bond is \$190 million and represents, on average, 2.7 % of the firm's assets. As measured by *RISK*, 24 % of putable bonds are rated below A-. The average free cash flow of the firm as a percentage of the firm's total assets is below 4 %. The average amount of interest saved (*INTSAVED*) due to the inclusion of a put option feature in the bond issue is approximately 0.03 % of the total assets of the issuing firm. The maximum interest expense saved is as high as \$1.23 per \$100 of total assets.[23] However, *INTSAVED* is negative for ten observations, which may be due to the lack of a closely matched straight bond. To guard against this error-in-variable problem, we estimate our regression Eq. 5.1 using two alternative samples: the entire sample and a sample comprising of positive *INTSAVED* observations. The number of analysts following a company ranges from 1 to 41. Panel B of Table 5.2 provides the corresponding summary statistics for the sample of issuers of poison put bonds. We note that these firms are smaller than the issuers of bonds with European put options. Additionally, issuers of poison puts have a smaller number of following analysts, and the bonds tend to be somewhat riskier than the issuers of bonds with European put options.

Panel A of Table 5.3 presents the average daily abnormal performance of the equity of issuers of bonds with European put options in our sample for $t = -30$ to $t = 30$. These abnormal returns are obtained from a market model. Please note that the t-statistics presented in Table 5.3 are based on standardized abnormal returns. In an efficient market, we expect the market to impound the economic informational impact of the new bond issue on the day of the announcement ($t = 0$). The average abnormal return at $t = 0$ is almost 0.33 % and is significantly positive at the 1 % level.

[22]If a major motivation for issuing putable bonds is to enhance management entrenchment, then we expect that the value of entrenchment is directly related to the level of free cash flow which management can misappropriate.

[23]Crabbe (1991) demonstrates that bonds with poison puts reduce the cost of borrowing for firms.

Table 5.2 Summary statistics of putable bond issuing firms

Variable	Mean	Std dev	Minimum	Maximum
Panel A: Sample of 90 issuers of bonds with European put options				
Sales	14,656.250	23,238.860	89.792	124,993.900
Long-term debt	4,206.870	7,568.520	51.522	50,218.300
Total assets	30,789.070	41,623.450	182.281	174,429.410
Leverage ratio	0.186	0.122	0.012	0.457
Issue size	190.162	106.701	14.300	500.000
Ebit	1,818.390	2,647.810	10.457	13,275.700
ISSUE	0.027	0.056	0.001	0.439
RISK	0.244	0.432	0.000	1.000
FCF1	0.020	0.093	−0.082	0.823
FCF2	0.034	0.095	−0.080	0.824
Number of analysts	21.922	8.388	1.000	41.000
INTSAVED	0.029	0.130	−0.028	1.226
Panel B: Sample of 47 issuers of bonds with poison put options				
Sales	5,033.849	5,921.869	406.360	34,922.000
Long-term debt	1,203.060	2,148.865	25.707	13,966.000
Total assets	5,080.781	7,785.288	356.391	51,038.000
Leverage ratio	0.208	0.112	0.031	0.476
Issue size	158.114	80.176	50.000	350.000
Ebit	481.121	639.842	−30.800	3,825.000
ISSUE	0.062	0.048	0.006	0.251
RISK	0.314	0.469	0.000	1.000
FCF1	0.034	0.032	−0.075	0.111
FCF2	0.051	0.036	−0.050	0.132
Number of analysts	18.803	7.792	2.000	39.000
INTSAVED	0.011	0.044	−0.070	0.216

Sales, long-term debt, total assets, and EBIT are in millions of dollars and are for the fiscal year prior to the putable bond issue announcement. Leverage ratio is the ratio of the long-term debt to total assets. Issue size is the dollar amount of putable bond issue in millions of dollars. ISSUE is the ratio of issue size to the total assets of the issuing firm as of the fiscal year prior to the issue announcement date. *RISK* is a dummy variable equal to one if the bond issue has an S&P bond rating of BBB+ or below and zero otherwise. *FCF1* and *FCF2* are the two free cash flow measures as described in the text. In particular, *FCF1* is defined as [Earnings before Interest and Taxes – Taxes – Interest Expense – Preferred Dividend Payments – Common Stock Dividend Payments]/Total Assets. *FCF2* is defined as [Earnings before Interest and Taxes – Taxes – Interest Expense]/Total Assets. *Number of analysts* is the number of analysts who follow the issuing firms for the year prior to the putable bond issue. *INTSAVED* measures the relative amount of aggregate interest expense saved per $100 of total assets of the issuing firm

This suggests that the market views favorably the announcement of putable bonds. This may be due to mitigating agency costs or resolving asymmetric information. However, for the 3-day window ($t = -1, 1$), the abnormal return is 0.19 % but is not statistically significant from zero.

Table 5.3 The average daily abnormal returns for firms issuing putable bonds from 30 days prior to the putable bond issuance announcement to 30 days after the announcement. The t-statistics are based on standardized abnormal returns. CAR is the cumulative abnormal return

Event day	Abnormal return	t-statistic	CAR
Panel A: Sample of 90 issuers of bonds with European put options			
−30	0.0006	0.1988	0.0006
−29	0.0004	0.2305	0.0010
−28	0.0011	0.8169	0.0021
−27	−0.0003	−0.1248	0.0017
−26	0.0002	−0.2028	0.0019
−25	0.0003	0.0263	0.0022
−24	0.0003	0.2174	0.0025
−23	−0.0002	−0.1306	0.0023
22	0.0007	0.2573	0.0030
−21	−0.0003	−0.0367	0.0026
−20	−0.0024	−1.6700	0.0003
−19	0.0000	0.1348	0.0002
−18	−0.0025	−1.6932	−0.0022
−17	0.0029	1.7141	0.0007
−16	−0.0005	0.0154	0.0002
−15	0.0022	1.4378	0.0024
−14	0.0022	1.4701	0.0046
−13	−0.0002	−0.1052	0.0043
−12	0.0007	0.4783	0.0050
−11	−0.0014	−1.2803	0.0036
−10	0.0003	0.1904	0.0039
−9	−0.0014	−0.5100	0.0025
−8	−0.0006	−0.6052	0.0019
−7	0.0001	−0.4020	0.0021
−6	0.0005	0.1268	0.0026
−5	0.0027	1.3209	0.0053
−4	−0.0001	−0.0050	0.0052
−3	0.0020	1.4407	0.0071
−2	−0.0028	−1.8590	0.0043
−1	−0.0021	−1.4323	0.0022
0	0.0033	2.6396	0.0055
1	0.0007	0.3534	0.0061
2	0.0007	0.3177	0.0068
3	0.0002	0.2628	0.0070
4	−0.0010	−0.7364	0.0060
5	0.0000	−0.2821	0.0059
6	−0.0036	−2.4342	0.0023
7	−0.0002	−0.4355	0.0022
8	−0.0004	−0.4253	0.0018

(*continued*)

Table 5.3 (continued)

Event day	Abnormal return	t-statistic	CAR
9	−0.0007	−0.5236	0.0011
10	0.0010	1.0509	0.0021
11	0.0039	2.8180	0.0060
12	0.0010	0.7853	0.0070
13	−0.0002	−0.0036	0.0068
14	−0.0013	−0.7663	0.0055
15	0.0006	−0.0920	0.0061
16	0.0014	1.2984	0.0075
17	0.0002	0.0847	0.0077
18	−0.0001	−0.2386	0.0076
19	0.0002	−0.3055	0.0078
20	0.0005	0.3692	0.0083
21	−0.0006	−0.0998	0.0077
22	−0.0016	−1.1099	0.0061
23	−0.0014	−0.8924	0.0047
24	0.0015	1.1956	0.0061
25	0.0001	−0.3050	0.0062
26	−0.0001	0.0257	0.0061
27	0.0004	0.4372	0.0065
28	−0.0011	−0.4242	0.0054
29	0.0004	0.8785	0.0058
30	−0.0010	−0.5929	0.0048
Panel B: Sample of 47 issuers of bonds with poison put options			
−30	−0.0002	−0.2385	−0.0002
−29	−0.0029	−1.7698	−0.0032
−28	−0.0005	−0.1247	−0.0036
−27	0.0015	0.2635	−0.0021
−26	−0.0010	−0.4019	−0.0031
−25	0.0013	0.6437	−0.0018
−24	0.0003	−0.1287	−0.0015
−23	−0.0007	−0.3447	−0.0022
−22	−0.0021	−1.3803	−0.0044
−21	0.0023	0.5731	−0.0020
−20	−0.0028	−1.2905	−0.0048
−19	−0.0009	0.1682	−0.0058
−18	−0.0024	−1.1787	−0.0081
−17	0.0034	1.4394	−0.0048
−16	0.0001	−0.0490	−0.0047
−15	−0.0002	−0.2326	−0.0049
−14	0.0035	1.8138	−0.0014
−13	0.0010	0.7301	−0.0005
−12	−0.0026	−1.9908	−0.0031

(*continued*)

Table 5.3 (continued)

Event day	Abnormal return	t-statistic	CAR
−11	0.0002	−0.0182	−0.0029
−10	0.0026	1.3253	−0.0003
−9	−0.0034	−1.5412	−0.0038
−8	−0.0002	0.2221	−0.0040
−7	0.0003	0.0028	−0.0037
−6	0.0019	1.0805	−0.0018
−5	−0.0009	−0.7397	−0.0026
−4	−0.0006	−0.5318	−0.0032
−3	−0.0001	−0.2118	−0.0033
−2	−0.0003	−0.0870	−0.0036
−1	−0.0028	−1.5751	−0.0064
0	−0.0028	−1.3777	−0.0092
1	0.0004	−0.5639	−0.0088
2	0.0019	0.8618	−0.0069
3	0.0027	1.4826	−0.0041
4	0.0022	1.4571	−0.0019
5	0.0034	1.2458	0.0014
6	−0.0006	−0.5491	0.0009
7	−0.0002	−0.1866	0.0006
8	−0.0036	−2.2457	−0.0030
9	0.0026	1.4036	−0.0003
10	−0.0017	−1.0479	−0.0020
11	−0.0016	−1.1626	−0.0036
12	−0.0008	−0.7991	−0.0043
13	−0.0020	−1.0808	−0.0064
14	−0.0004	−0.4107	−0.0068
15	0.0002	−0.2023	−0.0067
16	−0.0005	−0.2077	−0.0072
17	−0.0003	−0.3234	−0.0075
18	0.0005	0.5278	−0.0070
19	−0.0018	−0.9564	−0.0087
20	0.0014	0.7079	−0.0073
21	−0.0022	−1.0614	−0.0095
22	−0.0015	−0.4152	−0.0110
23	−0.0004	−0.2990	−0.0114
24	−0.0015	−0.7324	−0.0129
25	−0.0006	−0.2189	−0.0134
26	0.0004	0.4731	−0.0130
27	0.0015	0.2512	−0.0115
28	0.0014	0.2703	−0.0102
29	−0.0022	−1.1175	−0.0124
30	−0.0002	0.2668	−0.0126

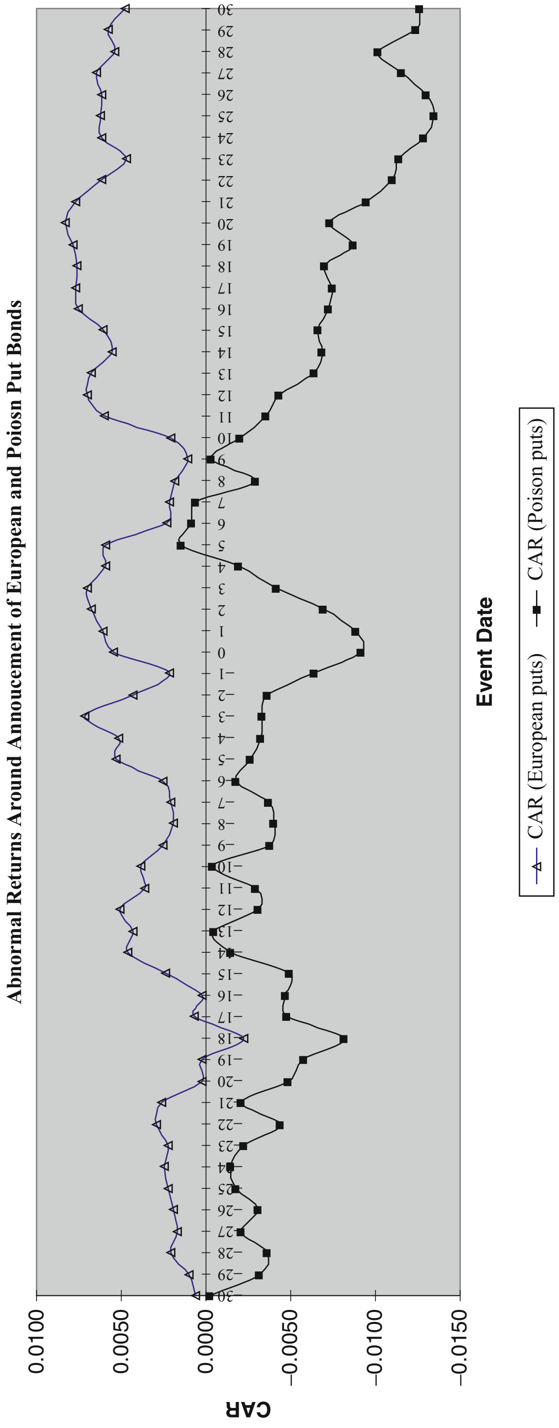


Fig. 5.1 Cumulative abnormal returns from $t = -30$ to $t = +30$ around the issue announcement of bonds with European or poison put features

Table 5.4 Cross-sectional regression results for the European put sample: base case

	Full sample of 90 firms			Sample of 80 bonds with positive INTSAVED		
	Coef.	t-ratio	p-value	Coef.	t-ratio	p-value
Intercept	−0.0388	−1.69	0.095	−0.03657	−1.58	0.1193
FCF1	−0.0016	−0.09	0.9289	−0.00392	−0.22	0.8277
RISK	−0.0016	−0.25	0.8026	−0.00245	−0.37	0.7117
SIZE	0.0097	2.93	0.0044	0.010637	3.03	0.0034
INTSAVED	0.0269	2.59	0.0115	0.023819	2.27	0.0259
ANALYSTS	−0.0153	−2.48	0.0152	−0.01848	−2.82	0.0062
FINS	−0.0203	−2.79	0.0065	−0.02062	−2.75	0.0075
		χ^2	p-value		χ^2	p-value
H_0:β_1=β_2=β_3=β_4=β_5=β_6=0		26.38	0.0002		29.19	0.000
Adj. R^2		0.1185			0.1314	
	Coef.	t-ratio	p-value	Coef.	t-ratio	p-value
Intercept	−0.0390	−1.71	0.0915	−0.03642	−1.58	0.1191
FCF2	−0.0004	−0.02	0.9808	−0.00392	−0.21	0.8309
RISK	−0.0016	−0.24	0.8075	−0.00247	−0.37	0.7094
SIZE	0.0097	2.93	0.0043	0.010633	3.03	0.0034
INTSAVED	0.0270	2.59	0.0112	0.023751	2.26	0.0266
ANALYSTS	−0.0152	−2.47	0.0157	−0.01849	−2.8	0.0065
FINS	−0.0202	−2.73	0.0078	−0.02066	−2.7	0.0085
		χ^2	p-value		χ^2	p-value
H_0:β_1=β_2=β_3=β_4=β_5=β_6=0		26.73	0.0002		29.25	0.000
Adj. R^2		0.1184			0.1314	

This table reports the regression coefficients and their t-statistics of the following regression equation:

$CAR3_i = \beta_0 + \beta_1 FCF_i + \beta_2 RISK_i + \beta_3 SIZE_i + \beta_4 INTSAVED_i + \beta_5 ANALYSTS_i + \beta_6 FINS + \varepsilon_i$

CAR3 is the cumulative abnormal return measured from the day before the announced issue to the day after the announced issue; *FCF1* and *FCF2* are the two free cash flow measures as described in the text; *RISK* is a dummy variable equal to one if the bond issue has an S&P bond rating of BBB+ or below and zero otherwise; SIZE is the natural logarithm of the total assets of the issuing firm at the end of the fiscal year prior to the issue announcement; *INTSAVED* measures the relative amount of aggregate interest expense saved per $100 of total assets of the issuing firm; *ANALYSTS* is the natural logarithm of one plus the number of analysts following the firm; and *FINS* is equal to one if the parent company is a financial institution and is equal to zero otherwise. The p-values assume a two-tail test. All the t-ratios are heteroscedasticity consistent

Panel B of Table 5.3 presents the corresponding results for the issuers of poison put bonds. Note that the abnormal return at $t = 0$ is −0.28 % with a t-statistic of −1.38. Furthermore, during the 3-day (i.e., $t = -1, 1$) window, the abnormal return is negative (−0.52 %) and is significantly different than zero (t-statistic of −2.03). This negative abnormal return is consistent with the view that poison put bonds help entrench the current management. The cumulative abnormal returns of the two sample types of embedded put option are also depicted in Fig. 5.1. The patterns of CARs around the issue announcement dates clearly show the positive trend of CARs for bonds with European puts versus the negative trend of CARs for bonds with poison puts.

Next, to test our hypotheses, we examine what determines the cross-sectional variations in the CARs for the sample of European and poison put bonds. Table 5.4 reports the parameter and t-statistic estimates for regression Eq. 5.1 for the sample of issuers of bonds with European put options. The four regressions reported in Table 5.4 differ by the sample that is used and by the specification used for the free cash flow variable. The regression in the top left panel uses the entire sample of 90 observations and the variable *FCF1*. The regression in the top right panel reports the results using only the 80 observations in which the variable *INTSAVED* is positive. The regressions reported in the bottom panels use *FCF2* instead of *FCF1*.

First note that our estimates are robust to these alternative specifications. The Wald test of the hypothesis that all five independent variables are jointly zero is rejected at the 1 % significance level for all four regression specifications. The regression coefficients on the *ANALYSTS* variable, β_5, are negative and significantly different from zero for all four regressions. These estimates are consistent with the reduction in security mispricing motivation for issuing putable bonds. In contrast, the regression coefficients for *FCF* and *RISK* are not significantly different from zero, indicating a lack of empirical support for the mitigating agency cost hypothesis. The regression coefficients for the *INTSAVED* variables are positive and are significantly different from zero in all four regressions. This result is again consistent with the security mispricing motivation and inconsistent with the management myopia motivation. Additionally, the regression coefficient β_3 for the size variable is also positive and significantly different from zero, a result more consistent with size being related to the probability of survivorship of the firm. In summary, announcement of bonds with European puts is associated with positive abnormal returns which are related to our proxies for potential benefits from mitigating security mispricing.

Panel A of Table 5.5 presents the corresponding estimates for the sample of issuers of poison put bonds, while panel B of Table 5.5 presents the estimates of Eq. 5.2 for the same sample. The estimates reported in Table 5.5 are different from those in Table 5.4. The coefficients of the *ANALYSTS* variable, β_5, are significantly positive for the full sample. The coefficients for *Size* are significantly negative for the full sample. All the other variables are not significantly different than zero. The estimates of Eq. 5.2 indicate that the coefficients of the variable *COVRANK* are negative and significantly different from 0.

Our results so far indicate that the equity abnormal returns around the issue announcement dates of poison put bonds are negative, consistent with the managerial entrenchment evidence in Cook and Easterwood (1994) and Roth and McDonald (1999). Additionally, the positive and significant coefficient of the *ANALYSTS* variable may also be consistent with the management entrenchment hypothesis because it appears that the market negative response to the issuance of poison putable bonds arises from the less followed firms, where the management

Table 5.5 Cross-sectional regression results for the poison put sample

	Full sample of 47 firms			Sample of 28 bonds with positive INTSAVED		
	Coef.	t-ratio	p-value	Coef.	t-ratio	p-value
Panel A: Sample of issuers of bonds with poison put options						
Intercept	0.015	0.530	0.602	0.030	0.680	0.506
FCF1	−0.170	−1.330	0.191	−0.082	−0.510	0.617
RISK	0.010	1.550	0.129	0.001	0.090	0.932
SIZE	−0.007	−2.250	0.030	−0.006	−1.090	0.288
INTSAVED	0.022	0.370	0.714	−0.024	−0.280	0.782
ANALYSTS	0.014	2.100	0.042	0.006	0.370	0.717
		χ^2	p-value		χ^2	p-value
H_0:β_1=β_2=β_3=β_4=β_5=0		11.927	0.036		1.906	0.861
Adj. R^2		0.015			−0.181	
Intercept	0.018	0.730	0.472	0.026	0.640	0.532
FCF2	−0.153	−1.280	0.208	−0.045	−0.270	0.788
RISK	0.009	1.400	0.168	0.000	0.030	0.978
SIZE	−0.007	−2.480	0.017	−0.005	−1.020	0.318
INTSAVED	0.027	0.460	0.649	−0.016	−0.180	0.858
ANALYSTS	0.014	1.980	0.054	0.006	0.330	0.744
		χ^2	p-value		χ^2	p-value
H_0:β_1=β_2=β_3=β_4=β_5=0		11.538	0.042		1.857	0.868
Adj. R^2		0.007			−0.192	
	Full sample of 40 firms			Sample of 23 bonds with positive INTSAVED		
	Coef.	t-ratio	p-value	Coef.	t-ratio	p-value
Panel B: Sample of issuers of bonds with poison put options with COVRANK						
Intercept	0.052	1.200	0.239	0.136	2.550	0.021
FCF1	−0.158	−1.100	0.281	−0.271	−1.330	0.201
RISK	0.011	1.140	0.264	0.006	0.370	0.720
SIZE	−0.008	−1.800	0.081	−0.014	−2.370	0.031
INTSAVED	0.029	0.430	0.670	−0.071	−0.720	0.484
COVRANK	−0.008	−3.900	0.000	−0.012	−3.520	0.003
ANALYSTS	0.013	2.090	0.045	0.009	0.700	0.497
		χ^2	p-value		χ^2	p-value
H_0:β_1=β_2=β_3=β_4=β_5=β_6=0		48.723	0.000		20.513	0.002
Adj. R^2		0.114			0.058	
Intercept	0.053	1.440	0.160	0.117	3.030	0.008
FCF2	−0.132	−1.000	0.327	−0.159	−0.890	0.386
RISK	0.010	1.110	0.275	0.000	0.010	0.995
SIZE	−0.009	−2.030	0.051	−0.013	−2.590	0.020
INTSAVED	0.032	0.510	0.616	−0.030	−0.240	0.810
COVRANK	−0.008	−4.100	0.000	−0.010	−3.610	0.002

(*continued*)

Table 5.5 (continued)

	Full sample of 40 firms			Sample of 23 bonds with positive INTSAVED		
	Coef.	t-ratio	p-value	Coef.	t-ratio	p-value
Panel B: Sample of issuers of bonds with poison put options with COVRANK						
ANALYSTS	0.014	1.970	0.057	0.010	0.710	0.488
		χ^2	p-value		χ^2	p-value
$H_0{:}\beta_1=\beta_2=\beta_3=\beta_4=\beta_5=\beta_6=0$		53.590	0.000		31.077	0.000
Adj. R^2		0.102			–0.006	

Panels A reports the regression coefficients and their t-statistics of the following regression equation:
$CAR3_i = \beta_0 + \beta_1 FCF_i + \beta_2 RISK_i + \beta_3 SIZE_i + \beta_4 INTSAVED_i + \beta_5 ANALYSTS_i + \varepsilon_i$
Panel B reports the regression coefficients the regression coefficients and their t-statistics of the following regression equation:
$CAR3_i = \beta_0 + \beta_1 FCF_i + \beta_2 RISK_i + \beta_3 SIZE_i + \beta_4 INTSAVED_i + \beta_5 ANALYSTS_i + \beta_7 COVRANK_i + e_i$
CAR3 is the cumulative abnormal return measured from the day before the announced issue to the day after the announced issue; *FCF1* and *FCF2* are the two free cash flow measures as described in the text; *RISK* is a dummy variable equal to one if the bond issue has a S&P bond rating of BBB+ or below, and zero otherwise; SIZE is the natural logarithm of the total assets of the issuing firm at the end of the fiscal year prior to the issue announcement; *INTSAVED* measures the relative amount of aggregate interest expense saved per $100 of total assets of the issuing firm; *ANALYSTS* is the natural logarithm of one plus the number of analysts following the firm; and *COVRANK* equals the S&P Event Risk Ranking on a scale of 1–5. The p-values assume a two-tail test. All the t-ratios are heteroscedasticity consistent

strategy may not be as well known prior to the bond issuance. However, these returns are negatively related to the event risk covenant ranking consistent with the agency cost evidence in Bae et al. (1994). The negative abnormal returns experienced by issuers of poison puts and the different coefficients on *SIZE*, *INTSAVED*, and *ANALYSTS* indicate that the European put bonds are viewed differently than poison put bonds.

In summary, the results reported in Tables 5.2–5.5 are consistent with the view that European put bonds, where the put option does not depend on a specific company event, are more effective in mitigating problems that are associated with security mispricing. Furthermore, there is no empirical support for the hypothesis that European put bonds are used as a vehicle for management entrenchment. In contrast, the evidence for poison put bonds is consistent with both management entrenchment and mitigating agency costs.

We now discuss several robustness checks of our basic results. Note that, as reported in Table 5.4, the regression coefficient of the financial industry dummy variable, β_6, is negative and significantly different from zero. To verify that the differing results between the European put bond sample (where 25 out of 90 firms

Table 5.6 Cross-sectional regression results for the European put sample excluding financial service companies

	Full sample of 65 firms			Sample of 61 bonds with positive INTSAVED		
	Coef.	t-ratio	p-value	Coef.	t-ratio	p-value
Intercept	−0.0410	−1.27	0.2087	−0.0416	−1.25	0.2182
FCF1	0.0005	0.04	0.9715	−0.0016	−0.10	0.9200
RISK	−0.0028	−0.41	0.6821	−0.0033	−0.49	0.6279
SIZE	0.0108	2.79	0.0071	0.0114	2.80	0.0071
INTSAVED	0.0248	1.85	0.0694	0.0242	1.79	0.0785
ANALYSTS	−0.0178	−2.54	0.0138	−0.0191	−2.70	0.0093
		χ^2	p-value		χ^2	p-value
H_0:β_1=β_2=β_3=β_4=β_5=0		32.66	0.0000		32.10	0.0000
Adj. R^2		0.1450			0.1524	
Intercept	−0.0409	−1.28	0.2073	−0.0414	−1.25	0.2181
FCF2	0.0000	0.00	0.9992	−0.0031	−0.17	0.8676
RISK	−0.0028	−0.41	0.6805	−0.0034	−0.50	0.6206
SIZE	0.0108	2.78	0.0072	0.0114	2.79	0.0071
INTSAVED	0.0248	1.85	0.0686	0.0241	1.79	0.0784
ANALYSTS	−0.0178	−2.54	0.0137	−0.0191	−2.70	0.0091
		χ^2	p-value		χ^2	p-value
H_0:β_1=β_2=β_3=β_4=β_5=0		32.66	0.0000		32.10	0.0000
Adj. R^2		0.1450			0.1525	

This table reports the regression coefficients and their t-statistics of the following regression equation:

$CAR3_i = \beta_0 + \beta_1 FCF_i + \beta_2 RISK_i + \beta_3 SIZE_i + \beta_4 INTSAVED_i + \beta_5 ANALYSTS_i + \varepsilon_i$

CAR3 is the cumulative abnormal return measured from the day before the announced issue to the day after the announced issue; *FCF1* and *FCF2* are the two free cash flow measures as described in the text; *RISK* is a dummy variable equal to one if the bond issue has a S&P bond rating of BBB+ or below, and zero otherwise; SIZE is the natural logarithm of the total assets of the issuing firm at the end of the fiscal year prior to the issue announcement; *INTSAVED* measures the relative amount of aggregate interest expense saved per $100 of total assets of the issuing firm; *ANALYSTS* is the natural logarithm of one plus the number of analysts following the firm. The p-values assume a two-tail test. All the t-ratios are heteroscedasticity consistent

are from the financial service sector) and the poison put sample (where only one firm is from the financial service sector), we replicate the regressions of Tables 5.4 and 5.5 excluding all financial service companies. The estimates for the subsample of European put bonds are reported in Table 5.6. Note that the coefficients and significance levels are very similar to those reported in Table 5.4. Because only one poison put company is from the financial sector and because the estimates of the corresponding subsample of poison put bonds are very similar to those reported in Table 5.5, we do not report these estimates. Thus, we conclude that the differences between poison and European put bonds are not due to the different sector composition of our samples.

Table 5.7 The cross-sectional regression results for the European put sample: impact of time to expiration

	Full sample of 90 firms			Sample of 80 bonds with positive INTSAVED		
	Coef.	t-ratio	p-value	Coef.	t-ratio	p-value
Intercept	−0.0288	−1.27	0.2063	−0.0282	−1.22	0.2279
FCF1	0.0037	0.24	0.8147	0.0007	0.05	0.9627
RISK	−0.0010	−0.16	0.8703	−0.0021	−0.32	0.7527
SIZE	0.0088	2.74	0.0075	0.0098	2.88	0.0052
INTSAVED	0.0143	1.35	0.1814	0.0131	1.22	0.2249
ANALYSTS	−0.0165	−2.68	0.0090	−0.0193	−2.95	0.0043
FINS	−0.0185	−2.60	0.0112	−0.0194	−2.61	0.0109
EXPLT5	0.0097	1.91	0.0601	0.0083	1.37	0.1739
		χ^2	p-value		χ^2	p-value
H_0:β_1=β_2=β_3=β_4=β_5=β_6= β_7=0		34.82	0.0000		37.88	0.000
Adj. R^2		0.1413			0.1420	
Intercept	−0.0291	−1.3	0.1981	−0.0283	−1.22	0.2248
FCF2	0.0051	0.33	0.7397	0.0010	0.06	0.9506
RISK	−0.0010	−0.15	0.8794	−0.0021	−0.31	0.7542
SIZE	0.0088	2.74	0.0075	0.0098	2.88	0.0052
INTSAVED	0.0144	1.36	0.1784	0.0132	1.22	0.2257
ANALYSTS	−0.0164	−2.67	0.0092	−0.0193	−2.93	0.0045
FINS	−0.0184	−2.53	0.0133	−0.0194	−2.57	0.0122
EXPLT5	0.0097	1.91	0.0593	0.0083	1.38	0.1733
		χ^2	p-value		χ^2	p-value
H_0:β_1=β_2=β_3=β_4=β_5=β_6=β_7=0		34.85	0.0000		37.82	0.000
Adj. R^2		0.1415			0.1420	

This table reports the regression coefficients and their t-statistics of the following regression equation: $CAR3_i = \beta_0 + \beta_1 FCF_i + \beta_2 RISK_i + \beta_3 SIZE_i + \beta_4 INTSAVED_i + \beta_5 ANALYSTS_i + \beta_6 FINS + \beta_7 EXPLT5 + \varepsilon_i$ *CAR3* is the cumulative abnormal return measured from the day before the announced issue to the day after the announced issue; *FCF1* and *FCF2* are the two free cash flow measures as described in the text; *RISK* is a dummy variable equal to one if the bond issue has a S&P bond rating of BBB+ or below and zero otherwise; SIZE is the natural logarithm of the total assets of the issuing firm at the end of the fiscal year prior to the issue announcement; *INTSAVED* measures the relative amount of aggregate interest expense saved per $100 of total assets of the issuing firm; *ANALYSTS* is the natural logarithm of one plus the number of analysts following the firm; *FINS* is equal to one if the parent company is a financial institution and is equal to zero otherwise; and *EXPLT5* is equal to one if the expiration of the embedded option is less than 5 years from the issue date and is equal to zero otherwise. The p-values assume a two-tail test. All the t-ratios are heteroscedasticity consistent

Next, we examine whether the term to expiration of European put bonds affects our estimates.[24] Table 5.7 repeats the regressions of Table 5.4 when we introduce an additional explanatory dummy variable, *EXPLT5*, that is equal to one if the

[24]Recall that by definition the option of poison put bonds does not have an expiration date.

expiration date of the embedded European put option is less than 5 years from the issue date and zero otherwise. We find that *INTSAVE* is still positive but is no longer significant and the regression coefficient for *EXPLT5* is positive and is significantly different from zero at the 10 % level for the full sample. The estimates and the significance of the other coefficients are largely unaffected. We conclude that these results still confirm the security mispricing hypothesis because the regression coefficients for *ANALYSTS* are significantly negative.

Our results are robust to alternative specifications we described at the end of the previous section. In particular, including a dummy variable for the existence of a call feature, interaction variables and alternative measures of interest rate volatility do not affect our results. These results are not reported here but are available upon request.

Finally, we test for multicollinearity by examining the eigenvalues of the correlation matrix for (non-dummy) independent variables. For orthogonal data the eigenvalue, λ, for each variable should equal 1, and $\Sigma 1/\lambda$ = the number of regressors (i.e., five for our study). For our sample of bonds with an embedded European put option, this sum is 7.14 when FCF1 is used, and this sum = 7.19 when FCF2 is used, indicating a lack of significant multicollinearity. Similar results are obtained for the poison put sample. Therefore, our findings are robust to alternate specifications and are not driven by multicollinearity.

5.6 Concluding Remarks

This paper examines the motivations and equity valuation impact of issuing European putable bonds, a bond containing an embedded European put option held by bondholders. The option entitles them to sell the bond back to the firm on the exercise date at a predetermined price. Unlike a poison put bond which has been studied by the literature, the exercise of the put option in a European putable bond is not contingent upon a company-related event.

We find that the market reacts favorably to the issue announcement of such putable bonds. We consider three alternative motivations for incorporating a European put option in a bond contract: reducing the security mispricing impact of asymmetric information, mitigating agency costs, and the relatively low coupon rate (a myopic view that ignores the potential liability to the firm due to the put option). We test these hypotheses by conducting a cross-sectional empirical study of the impact of putable debt issue announcements on the equity value of the issuing companies. Our results indicate that the market favorably views putable bonds as a means to reduce security mispricing. We also find that the market reaction to poison put announcements differs from the market reaction to issue announcements for bonds with European put options.

Appendix 1: Sample of Firms Issuing Putable Bonds and the Put Bond Characteristics

Name	Coupon	Amount ($000's)	Issue date	Maturity (years)	Put expiration (years)	Callability	YTM (%)	S&P credit rating
Air Products and Chemicals Inc	7.34	100,000	06/15/1996	30	12	N	7.20	7
American Express	8.50	150,000	06/09/1989	10	5	N	8.56	5
Anadarko Petroleum Corp	7.25	100,000	03/17/1995	30	5	N	6.90	9
Anadarko Petroleum Corp	7.73	100,000	09/19/1996	100	30	N	7.21	9
Bankamerica Corp	7.65	150,000	04/26/1984	10	2	C	12.23	10
Bausch & Lomb Inc	6.56	100,000	08/12/1996	30	5	N	6.71	8
Baxter International Inc	8.88	100,000	06/14/1988	30	5	C	8.96	9
Burlington Northern Santa Fe	7.29	200,000	05/31/1996	40	12	N	7.39	10
Champion International Corp	15.63	100,000	07/22/1982	12	3	C	15.68	9
Champion International Corp	6.40	200,000	02/15/1996	30	10	N	6.69	10
Chase Manhattan Corp – old	7.55	250,000	06/12/1985	12	5	C	8.87	5
Chrysler Corp	12.75	200,000	11/01/1984	15	5	N	12.75	10
Chrysler Corp	9.65	300,000	07/26/1988	20	5	C	9.65	10
Chrysler Corp	9.63	200,000	09/06/1988	20	2	C	9.65	11
Circus Circus Enterprise Inc	6.70	150,000	11/15/1996	100	7	N	6.66	9
Citicorp	9.40	250,000	12/06/1983	12	2	C	11.05	4
Citicorp	10.25	300,000	12/19/1984	10	2	C	10.31	4

(*continued*)

Name	Coupon	Amount ($000's)	Issue date	Maturity (years)	Put expiration (years)	Callability	YTM (%)	S&P credit rating
Citicorp	8.75	250,000	12/12/1985	15	3	C	8.84	4
Coca-Cola Enterprises	7.00	300,000	09/27/1996	30	10	N	7.00	5
Commercial Credit	8.50	100,000	02/05/1988	10	5	N	8.84	8
Commercial Credit	8.70	150,000	06/09/1989	20	10	N	9.06	8
Commercial Credit	8.70	100,000	06/11/1990	20	3	N	7.54	7
Commercial Credit	7.88	200,000	02/01/1995	30	10	N	7.30	6
Conagra Inc	7.13	400,000	10/02/1996	30	10	N	6.95	9
Corning Inc	7.63	100,000	07/31/1994	30	10	N	7.81	6
Deere & Co	8.95	199,000	06/08/1989	30	10	C	8.99	7
Diamond Shamrock Inc	7.65	100,000	06/25/1996	30	10	N	7.37	10
Dow Chemical	8.48	150,000	08/22/1985	30	14	C	9.35	7
Eastman Kodak Co	7.25	125,000	04/08/1987	10	5	N	8.48	8
Eaton Corp	8.00	100,000	08/18/1986	20	10	N	7.90	7
Eaton Corp	8.88	38,000	06/14/1989	30	15	N	9.00	7
Eaton Corp	6.50	150,000	06/09/1995	30	10	N	6.64	7
Enron Corp	9.65	100,000	05/17/1989	12	7	N	9.50	10
First Chicago Corp	8.50	98,932	05/13/1986	12	7	N	8.70	7
First Interstate Bancorp	7.35	150,000	08/23/1984	15	6	C	13.70	4
First Interstate Bancorp	9.70	100,000	07/08/1985	15	5	C	9.91	4
First Union Corp (NC)	7.50	250,000	04/25/1995	40	10	N	7.96	8
First Union Corp (NC)	6.82	300,000	08/01/1996	30	10	N	7.43	8
Ford Motor Co	7.50	250,000	10/30/1985	15	3	N	9.63	6

(*continued*)

Name	Coupon	Amount ($000's)	Issue date	Maturity (years)	Put expiration (years)	Callability	YTM (%)	S&P credit rating
Ford Motor Co	9.95	300,000	02/10/1992	40	3	N	8.56	6
General Electric Co	6.75	250,000	11/06/1986	25	5	C	6.80	2
General Electric Co	8.25	500,000	04/26/1988	30	3	C	8.41	2
General Motors Corp	8.38	200,000	04/30/1987	10	5	N	8.38	5
General Motors Corp	8.63	400,000	06/09/1989	10	5	N	8.59	5
General Motors Corp	8.88	500,000	05/31/1990	20	5	N	8.80	5
Harris Corp	6.65	100,000	08/01/1996	10	5	N	6.90	8
Ingersoll-Rand Co	6.48	150,000	06/01/1995	30	10	N	6.64	7
Intl Business Machines Corp	13.75	125,000	03/09/1982	12	3	C	14.05	2
ITT Industries Inc	8.50	100,000	01/20/1988	10	5	N	8.50	7
ITT Industries Inc	8.55	100,000	06/12/1989	20	9	N	9.20	7
ITT Industries Inc	3.98	100,000	02/15/1990	15	3	N	8.60	7
Johnson Controls Inc	7.70	125,000	02/28/1995	20	10	N	7.83	7
K N Energy Inc	7.35	125,000	07/25/1996	30	10	N	7.47	9
Litton Industries Inc	6.98	100,000	03/15/1996	40	10	N	7.01	9
Lockheed Martin Corp	7.20	300,000	05/01/1996	40	12	N	7.29	9
Marriott Corp	9.38	250,000	06/11/1987	20	10	N	9.74	8
Merrill Lynch & Co	11.13	250,000	03/26/1984	15	5	C	13.36	4
Merrill Lynch & Co	9.38	125,000	06/04/1985	12	6	C	9.48	4
Merrill Lynch & Co	8.40	200,000	10/25/1989	30	5	N	8.67	6
Motorola Inc	8.40	200,000	08/15/1991	40	10	N	8.36	4
Motorola Inc	6.50	400,000	08/31/1995	30	10	N	6.55	4

(*continued*)

Name	Coupon	Amount ($000's)	Issue date	Maturity (years)	Put expiration (years)	Callability	YTM (%)	S&P credit rating
Occidental Petroleum Corp	9.25	300,000	08/03/1989	30	15	N	9.37	10
Penney (JC) Co	6.90	200,000	08/16/1996	30	7	N	7.07	7
Philip Morris Cos Inc	7.00	150,000	07/15/1986	5	2	N	7.30	7
Philip Morris Cos Inc	9.00	350,000	05/09/1988	10	6	N	9.39	7
Philip Morris Cos Inc	6.95	500,000	06/01/1996	10	5	N	6.91	7
Pitney Bowes Inc	8.63	100,000	02/10/1988	20	10	N	8.70	4
Pitney Bowes Inc	8.55	150,000	09/15/1989	20	10	Y	8.64	4
Regions Financial Corp	7.75	100,000	09/15/1994	30	10	N	8.07	7
Ryder System Inc	9.50	125,000	07/01/1985	15	2	C	9.69	7
Seagram Co Ltd	4.42	250,000	08/03/1988	30	15	N	9.90	7
Security Pacific Corp	12.50	150,000	09/27/1984	12	6	C	12.62	3
Security Pacific Corp	7.50	150,000	04/03/1986	15	3	C	7.54	3
Service Corp International	7.00	300,000	05/26/1995	20	7	N	6.80	9
Southtrust Corp	1.62	100,000	05/09/1995	30	10	N	7.28	8
State Street Corp	7.35	150,000	06/15/1996	30	10	N	7.23	5
Suntrust Banks Inc	6.00	200,000	02/15/1996	30	10	N	6.33	7
Triad Systems Corp	14.00	71,500	08/09/1989	8	3	C	13.99	16
TRW Inc	9.35	100,000	05/31/1990	30	10	N	9.28	7
Union Carbide Corp	6.79	250,000	06/01/1995	30	10	N	6.83	10
United Dominion Realty Trust	8.50	150,000	09/22/1994	30	10	N	8.60	9
Westinghouse Electric	11.88	100,000	03/19/1984	12	3	C	13.38	6

(*continued*)

Name	Coupon	Amount ($000's)	Issue date	Maturity (years)	Put expiration (years)	Callability	YTM (%)	S&P credit rating
Westinghouse Electric	8.88	150,000	05/31/1990	24	4	N	8.84	13
Whitman Corp	7.29	100,000	09/19/1996	30	8	N	7.26	9
WMX Technology	8.75	250,000	04/29/1988	30	5	C	8.79	4
WMX Technology	7.65	150,000	03/15/1991	20	3	N	7.71	6
WMX Technology	6.22	150,000	05/09/1994	10	3	N	7.41	6
WMX Technology	6.65	200,000	05/16/1995	10	5	N	6.43	6
WMX Technology	7.10	450,000	07/31/1996	30	7	N	7.16	7
Xerox Corp	11.25	100,000	08/25/1983	15	3	C	11.44	5

Appendix 2: Sample of Firms Issuing Poison Put Bonds and the Put Bond Characteristics

Name	Coupon	Amount ($million)	Issue date	Years to maturity	Callability	YTM (%)	S&P credit rating
Aar Corp	9.500	65	10/27/89	12	N	9.425	10
AMR	9.750	200	03/15/90	10	N	9.85	7
Anheuser-Busch	8.750	250	12/01/89	10	N	8.804	5
Armstrong World	9.750	125	08/18/89	19	N	9.5	5
Ashland Oil	11.125	200	10/08/87	30	C	10.896	7
Becton Dickinson	9.950	100	03/13/89	10	N	10	6
Bowater Inc	9.000	300	08/02/89	20	N	9.331	7
Chrysler Financial	10.300	300	06/15/90	2	N	10.43	11
Coastal Corporation	10.250	200	12/06/89	15	N	10.02	11
Consolidated Freightways	9.125	150	08/17/89	10	N	9.202	6
Corning Inc	8.750	100	07/13/89	10	N	8.655	7
CPC Intl	7.780	200	12/15/89	15	N	7.780	8

(*continued*)

Name	Coupon	Amount ($million)	Issue date	Years to maturity	Callability	YTM (%)	S&P credit rating
Cummins Engine	9.750	100	03/21/86	30	C	10.276	9
Cyprus Minerals	10.125	150	04/11/90	12	N	10.172	10
Dresser Industries	9.373	68.8	03/10/89	11	C	9.650	8
Eaton Corp	9.000	100	03/18/86	30	C	9.323	7
Federal Express	9.200	150	11/15/89	5	N	9.206	9
General American Transportation	10.125	115	03/22/90	12	N	10.269	8
Georgia Pacific	10.000	300	06/13/90	7	N	10.053	9
Grumman Corp	10.375	200	01/05/89	10	C	10.375	9
Harris Corp	10.375	150	11/29/88	30	C	10.321	8
Harsco	8.750	100	05/15/91	5	N	8.924	8
International Paper	9.700	150	03/21/90	10	N	9.823	8
Kerr-Mcgee	9.750	100	04/01/86	30	C	9.459	8
Knight-Rydder	9.875	200	04/21/89	20	N	10.05	5
Lockhee Corp	9.375	300	10/15/89	10	N	9.329	7
Maytag	8.875	175	07/10/89	10	N	9.1	8
Monsanto	8.875	100	12/15/89	20	N	8.956	7
Morton International	9.250	200	06/01/90	30	N	9.358	5
Parker-Hannifin	9.750	100	02/11/91	30	C	9.837	7
Penn Central	9.750	200	08/03/89	10	N	9.358	11
Penn Central	10.875	150	05/01/91	20	N	11.016	11
Potlatch	9.125	100	12/01/89	20	N	9.206	8
Questar	9.875	50	06/11/90	30	C	9.930	6
Ralston Purina	9.250	200	10/15/89	20	N	9.45	8
Rite-Aid	9.625	65	09/25/89	27	C	9.99	6
Rohm And Haas	9.373	100	11/15/89	30	C	9.618	7
Safety-Kleen	9.250	100	09/11/89	10	N	9.678	9
Sequa Corp	9.625	150	10/15/89	10	N	9.574	11
Stanley Works	8.250	75	04/02/86	10	C	8.174	7
Strawbridge And Clothier	8.750	50	10/24/89	7	C	9.374	8
Union Camp	10.000	100	04/28/89	30	C	10.185	6
Unisys	10.300	300	05/29/90	7	N	10.794	10

(*continued*)

Name	Coupon	Amount ($million)	Issue date	Years to maturity	Callability	YTM (%)	S&P credit rating
United Airlines	12.500	150	06/03/88	7	C	12.500	15
United Technologies	8.875	300	11/13/89	30	N	9.052	5
VF Corp	9.500	100	10/15/89	10	C	9.500	7
Weyerhaeuser	9.250	200	11/15/90	5	N	9.073	6

The S&P ratings are based on a scale from 1 (AAA+) to 24 (Not rated). In the Callability column, C denotes callable bond and N denotes non-callable bond

Appendix 3: Estimating the Standard Abnormal Returns and the White t-Statistic

Fama et al. (1969) introduced the event study methodology when they analyzed the impact of stock dividend announcements upon stock prices. Essentially, they used the Market Model to estimate the stochastic relationship between stock returns and the market portfolio. In particular, we estimate the following regression:

$$R_{i,t} = \alpha_i + \beta_i R_{m,t} + \varepsilon_{i,t} \tag{5.3}$$

We estimate the market model coefficients using the time period that begins 200 trading days before and ends 31 trading days before the event, employing the CRSP value-weighted market index as the benchmark portfolio. We use these coefficients to estimate abnormal returns for days −30 to +30. The abnormal return is defined as

$$AR_{i,t} = R_{i,t} - \alpha_i - \beta_i R_{m,t}. \tag{5.4}$$

The mean abnormal return for the sample of *i* firms, AR_t, at time t is found by

$$AR_t = \sum_{i=1}^{n} AR_{i,t}/n. \tag{5.5}$$

In order to conduct tests of significance, we must ensure that $AR_{i,t}$ has identical standard deviation. Assuming that the forecast values are normally distributed, we can scale $AR_{i,t}$ by the standard deviation of the prediction, S_{it}, given by Eq. 5.4 In particular,

$$S_{it} = \left\{\sigma_i^2 + (1/ED) + \left[\left(R_{mt} - \overline{R}_m\right)^2 / \Sigma\left(R_{mi} - \overline{R}_m\right)^2\right]\right\}^{1/2} \tag{5.6}$$

where σ^2_i is the variance of the error term of Eq. 5.3, ED is the number of days to estimate the market model for firm I, R_{mt} is the return of the market portfolio, and R_m is the mean market return in the estimation period.

Hence the standardized abnormal return, $SAR_{i,t}$, is equal to $AR_{i,t}/S_{it}$. $SAR_{i,t}$ is distributed normally with a mean of zero and a standard deviation equal to one. The mean standardized abnormal return for time t, SAR_t, is the sum of the $SAR_{i,t}$ divided by n. SAR_t is normally distributed with a mean of zero and a standard deviation of the square root of l/n. The cumulative abnormal returns for days 1 to k, CAR*k*, is the sum of mean abnormal returns for $t = 1$ to k. The standardized cumulative mean excess returns for the *k* days after month $t = 0$, *SCARt*, is equal to the sum of SAR_t for days 1 to k. *SCART* is normally distributed with a standard deviation of square root of k/n. Please see Campbell et al. (1996) for a detailed discussion of these tests as well as event studies in general.

In our cross-sectional regressions where we regress the CARs on firm-specific variables, we estimate the model using GMM. However, since we are not interested in conditional estimates, our regressors are the instruments. Therefore our parameters estimates are the same that would be obtained by OLS. However, we use the White (1980) heteroscedastic-consistent estimator. In the standard regression model,

$$Y = X\beta + \varepsilon \tag{5.7}$$

The OLS estimator of $\beta = (X'X)^{-1}(X'y)$ with the covariance matrix

$$\mathrm{VaR}(\beta) = (X'X)^{-1}(X'\Omega X)(X'X)^{-1} \tag{5.8}$$

When the errors are homoscedastic, $\Omega = \sigma^2 I$ and the variance reduces to $\sigma^2(X'X)^{-1}$.

However when Ω is unknown as shown by White (1980), Eq. 5.8 can be used using a consistent estimator of Ω. White showed that can be done using the residual from Eq. 5.7.

So when the heteroscedasticity is of the unknown form,

$$\text{Using } (X'X)^{-1}\left(X'\mathrm{diag}\left(\varepsilon^2\right)X\right)(X'X)^{-1} \tag{5.9}$$

gives us heteroscedastic-consistent White standard errors and White t-statistics.

Appendix 4: Generalized Method of Moments (GMM)

GMM is a generalization of the method of moments developed by Hansen (1982). The moment conditions are derived from the model. Suppose Y_t is a multivariate independently and identically distributed (i.i.d.) random variable. The econometric model specifies the relationship between Y_t and the true parameters of the model (θ_0). To use GMM there must exist a function $f(Y_t, \theta_0)$ so that

$$m(\theta_0) \equiv E[f(Y_t, \theta_0)] = 0. \tag{5.10}$$

In GMM, the theoretical expectations are replaced by sample analogs:

$$g(\theta, Y_t) = 1/T \sum f(Y_t, \theta). \tag{5.11}$$

The law of large numbers ensures that the RHS of above equation is the same as

$$E[f(Y_t, \theta_0)]. \tag{5.12}$$

The sample GMM estimator of the parameters may be written as (see Hansen 1982)

$$\Theta = \arg\min \left[1/T \sum f(Y_t, \theta)\right]' W_T 1/T \sum f(Y_t, \theta)]. \tag{5.13}$$

So essentially GMM finds the values of the parameters so that the sample moment conditions are satisfied as closely as possible. In our case for the regression model,

$$y_t = X_t{}'\beta + \varepsilon_t. \tag{5.14}$$

The moment conditions include

$$E[(y_t - X_t{}'\beta)x_t] = E[\varepsilon_t x_t] = 0 \text{ for all t}. \tag{5.15}$$

So the sample moment condition is

$$1/T \sum (y_t - X_t{}'\beta)x_t$$

and we want to select β so that this is as close to zero as possible. If we select β as $(X'X)^{-1}(X'y)$, which is the OLS estimator, the moment condition is exactly satisfied. Thus, the GMM estimator reduces to the OLS estimator and this is what we estimate. For our case the instruments used are the same as the independent variables. If, however, there are more moment conditions than the parameters, the GMM estimator above weighs them. These are discussed in detail in Greene (2008, Chap. 15).

The GMM estimator has the asymptotic variance

$$\left(X'Z\,(Z'\Omega Z)^{-1}Z'X\right)^{-1} \tag{5.16}$$

In our case Z = X since we use the independent variables as the instruments Z. The White robust covariance matrix may be used for Ω as discussed in Appendix 3 when heteroscedasticity is present. Using this approach, we estimate GMM with White heteroscedasticity consistent t-stats.

References

Bae, S. C., Klein, D. P., & Padmaraj, R. (1994). Event risk bond covenants, agency costs of debt and equity, and stockholder wealth. *Financial Management, 23*, 28–41.

Bae, S. C., Klein, D. P., & Padmaraj, R. (1997). Firm characteristics and the presence of event risk covenants in bond indentures. *Journal of Financial Research, 20*, 373–388.

Barnea, A., Haugen, R. A., & Senbet, L. W. (1980). A rationale for debt maturity structure and call provisions in the agency theoretic framework. *Journal of Finance, 25*, 1223–1234.

Billett, M. T., King, T. D., & Mauer, D. C. (2007). Growth opportunities and the choice of leverage, debt maturity, and covenants. *Journal of Finance, 62*, 697–730.

Bodie, Z., & Taggart, R. A. (1978). Future investment opportunities and the value of the call provision on a bond. *Journal of Finance, 33*, 1187–1200.

Brennan, M. J., & Schwartz, E. S. (1988). The case for convertibles. *Journal of Applied Corporate Finance, 1*, 55–64.

Brick, I. E., & Palmon, O. (1993). The tax advantages of refunding debt by calling, repurchasing, and putting. *Financial Management, 22*, 96–106.

Brick, I. E., & Wallingford, B. A. (1985). The relative tax benefits of alternative call features in corporate debt. *Journal of Financial and Quantitative Analysis, 20*, 95–106.

Campbell, J., Lo, A. W., & MacKinlay, C. (1996). *The econometrics of financial markets*. Princeton: Princeton University Press.

Chatfield, R. E., & Moyer, R. C. (1986). Putting away bond risk: An empirical examination of the value of the put option on bonds. *Financial Management, 15*(2), 26–33.

Cook, D. O., & Easterwood, J. C. (1994). Poison put bonds: An analysis of their economic role. *Journal of Finance, 49*, 1905–1920.

Crabbe, L. (1991). Event risk: An analysis of losses to bondholders and 'super poison put' bond covenants. *Journal of Finance, 46*, 689–706.

Crabbe, L., & Nikoulis, P. (1997). The putable bond market: Structure, historical experience, and strategies. *The Journal of Fixed Income, 7*, 47–60.

Dann, L. Y., & Mikkelson, W. H. (1984). Convertible debt issuance, capital structure change and financing-related information: Some new evidence. *Journal of Financial Economics, 13*, 157–186.

David, A. (2001). Pricing the strategic value of putable securities in liquidity crises. *Journal of Financial Economics, 59*, 63–99.

Eckbo, B. E. (1986). Valuation effects of corporate debt offerings. *Journal of Financial Economics, 5*, 119–152.

Fama, E., Fisher, L., Jensen, M., & Roll, R. (1969). The adjustment of stock prices to new information. *International Economic Review, 10*, 1–21.

Green, R. (1984). Investment incentives, debt, and warrants. *Journal of Financial Economics, 13*, 115–136.

Greene, W. (2008). *Econometric analysis*. Boston: Pearson.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica, 50*, 1029–1054.

Hansen, R. S., & Chaplinsky, S. (1993). Partial anticipation, the flow of information and the economic impact of corporate debt sales. *Review of Financial Studies, 6*, 709–732.

Hansen, R. S., & Crutchley, C. (1990). Corporate earnings and financings: An empirical analysis. *Journal of Business, 63*, 347–371.

Haugen, R. A., & Senbet, L. W. (1978). The insignificance of bankruptcy costs to the theory of optimal capital structure. *Journal of Finance, 33*, 383–393.

Haugen, R. A., & Senbet, L. W. (1981). Resolving the agency problems of external capital through options. *Journal of Finance, 26*, 629–647.

Haugen, R. A., & Senbet, L. W. (1988). Bankruptcy and agency costs: Their significance to the theory of optimal capital structure. *Journal of Financial and Quantitative Analysis, 23*, 27–38.

Jensen, M. C. (1986). Agency costs of free cash flow, corporate finance, and takeovers. *American Economic Review, 76*, 323–329.
Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics, 3*, 305–360.
Jung, K., Kim, Y., & Stulz, R. M. (1996). Timing, investment opportunities, managerial discretion, and the security issue decision. *Journal of Financial Economics, 42*, 159–185.
Kim, E. H. (1978). A mean-variance theory of optimal capital structure and corporate debt capacity. *Journal of Finance, 33*, 45–63.
Lehn, K., & Poulsen, J. (1989). Free cash flow and stockholders gain in going private transactions. *Journal of Finance, 44*, 771–788.
Lewis, C., Rogalski, R., & Seward, J. K. (1998). *Industry conditions, growth opportunities and the market reactions to convertible debt financing decisions*. (Unpublished mimeo, Vanderbilt University).
Mackie-Mason, J. K. (1990). Do taxes affect corporate financing decisions? *Journal of Finance, 45*, 1471–1493.
Meulbroek, L. K., Mitchell, M. L., Mulherin, J. H., Netter, J. M., & Poulsen, A. B. (1990). Shark repellants and managerial myopia: An empirical test. *Journal of Political Economy, 98*, 1108–1117.
Mikkelson, W. H., & Partch, M. M. (1986). Valuation effects of security offerings and the issuance process. *Journal of Financial Economics, 15*, 31–60.
Miller, M. H. (1977). Debt and taxes. *Journal of Finance, 32*, 261–275.
Miller, M. H., & Rock, K. (1985). Dividend policy under asymmetric information. *Journal of Finance, 40*, 1031–1051.
Modigliani, F., & Miller, M. H. (1963). Corporate income taxes and the cost of capital: A correction. *American Economic Review, 63*, 433–443.
Myers, S. (1977). Determinants of corporate borrowing. *Journal of Financial Economics, 5*, 147–175.
Nash, R. C., Netter, J. M., & Poulsen, A. B. (2003). Determinants of contractual relations between shareholders and bondholders: Investment opportunities and restrictive covenants. *Journal of Corporate Finance, 9*, 201–232.
Ravid, S. A., & Sarig, O. H. (1991). Financial signaling by committing to cash outflows. *Journal of Financial and Quantitative Analysis, 26*, 165–180.
Robbins, E. H., & Schatzberg, J. D. (1986). Callable bonds: A risk-reducing signaling mechanism. *Journal of Finance, 41*, 935–950.
Ross, S. (1977). The determination of financial structure: The incentive signaling approach. *Bell Journal of Economics, 8*, 23–40.
Roth, G., & McDonald, C. G. (1999). Shareholder-management conflict and event risk covenants. *Journal of Financial Research, 22*, 207–225.
Scott, J. (1976). A theory of optimal capital structure. *Bell Journal of Economics, 7*, 33–54.
Shyam-Sunder, L. (1991). The stock price effect of risky versus safe debt. *Journal of Financial and Quantitative Analysis, 26*, 549–558.
Spiegel, Y., & Wilkie, S. (1996). Investment in a new technology as a signal of firm value under regulatory opportunism. *Journal of Economics and Management Strategy, 5*, 251–276.
Stein, J. C. (1988). Takeover threats and managerial myopia. *Journal of Political Economy, 96*, 61–80.
Tewari, M., & Ramanlal, P. (2010). Is the put option in U.S. structured bonds good news for both bondholders and stockholders? *International Research Journal of Finance and Economics, 52*, 50–59.
Wahal, S., & McConnell, J. J. (2000). Do institutional investors exacerbate managerial myopia. *Journal of Corporate Finance: Contracting, Governance and Organization, 6*, 307–329.
White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica, 48*, 817–838.

Multi-Risk Premia Model of US Bank Returns: An Integration of CAPM and APT

6

Suresh Srivastava and Ken Hung

Contents

Abstract

Interest rate sensitivity of bank stock returns has been studied using an augmented CAPM, a multiple regression model with market returns and interest rate as independent variables. In this paper, we test an asset-pricing model in which the CAPM is augmented by three orthogonal factors which are proxies for the innovations in inflation, maturity risk, and default risk. The model proposed

S. Srivastava (✉)
University of Alaska Anchorage, Anchorage, AK, USA
e-mail: afscs@uaa.alaska.edu

K. Hung
Division of International Banking & Finance Studies, Texas A&M International University, Laredo, TX, USA
e-mail: ken.hung@tamiu.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_6,

is an integration of CAPM and APT. The results of the two models are compared to shed light on sources of interest rate risk.

Our results using the integrated model indicate the inflation beta to be statistically significant. Hence, innovations in short-term interest rates contain valuable information regarding inflation premium; as a result the interest rate risk is priced with respect to the short-term interest rates. Further, it also indicates that innovations in long-term interest rates contain valuable information regarding maturity premium. Consequently the interest rate risk is priced with respect to the long-term interest rates. Using the traditional augmented CAPM, our investigation of the pricing of the interest rate risk is inconclusive. It shows that interest rate risk was priced from 1979 to 1984 irrespective of the choice of interest rate variable. However, during the periods 1974–1978 and 1985–1990, bank stock returns were sensitive only to the innovations in the long-term interest rates.

Keywords
CAPM • APT • Bank stock return • Interest rate risk • Orthogonal factors • Multiple regression

6.1 Introduction

Interest rate sensitivity of commercial bank stock returns has been the subject of considerable academic research. Stone (1974) proposed a multiple regression model incorporating both the market return and interest rate variables as return-generating independent variables. While some studies have found the interest rate variable to be an important determinant of common stock returns of banks (Fama and Schwert 1977; Lynge and Zumwalt 1980; Christie 1981; Flannery and James 1984; Booth and Officer 1985), others have found the returns to be insensitive (Chance and Lane 1980) or only marginally explained by the interest rate factor (Lloyd and Shick 1977). A review of the early literature can be found in Unal and Kane (1988). Sweeney and Warga (1986) used the APT framework and concluded that the interest rate risk premium exists but varies over time. Flannery et al. (1997) tested a two-factor model for a broad class of security returns and found the effect of interest rate risk on security returns to be rather weak. Bae (1990) examined the interest rate sensitivity of depository and nondepository firms using three different maturity interest rate indices. His results indicate that depository institutions' stocks are sensitive to actual and unexpected interest rate changes, and the sensitivity increases for longer-maturity interest rate variables. Song (1994) examined the two-factor model using time-varying betas. His results show that both market beta and interest rate beta varied over the period 1977–1987. Yourougou (1990) found the interest rate risk to be high during a period of great interest rate volatility (post-October 1979) but low during a period of stable interest rates (pre-October 1979). Choi et al. (1992) tested a three-factor model of bank stock returns using market, interest, and exchange rate variables. Their findings about interest rate risk are consistent with the observations of Yourougou (1990).

The issue of interest rate sensitivity remains empirically unresolved. Most of the studies use a variety of short-term and long-term bond returns as the interest rate factor without providing any rationale for their use. The choice of bond market index seems to affect the pricing of the interest rate risk. Yet, there is no consensus on the choice of the interest rate factor that should be used in testing the two-factor model. In this paper, we provide a plausible explanation of why pricing of interest rate risk differs with the choice of interest rate variable. We also suggest a hybrid return-generating model for bank stock returns in which the CAPM is augmented by three APT-type factors to account for unexpected changes in the inflation premium, the maturity-risk premium, and the default-risk premium. The use of three additional factors provides a better understanding of the interest rate sensitivity and offers a plausible explanation for the time-varying interest rate risk observed by other investigators. Our empirical investigation covers three distinction economic and bank regulatory environments: 1974–1978, a period of increasing but only moderately volatile interest rates in a highly regulated banking environment; (2) 1979–1984, a period characterized by high level of interest rates with high volatility, in which there was gradual deregulation of the banking industry; and (3) 1985–1990, a low interest rate and low-volatility period during which many regulatory changes were made in response to enormous bank loan losses and bankruptcies. The results of the multifactor asset-pricing model are compared with those from the two-factor model in order to explain the time-varying interest rate risk.

The rest of this paper is divided into five sections. In Sect. 6.2, we describe the two-factor model of the bank stock return and the pricing of the interest rate risk. The multi-risk premia model and the specification of the factors are discussed in Sect. 6.3. The data for this analysis is described in Sect. 6.4. Section 6.5 presents empirical results, and Sect. 6.6 concludes the paper.

6.2 Multiple Regression Model of Bank Return

Stone (1974) proposed the following two-variable bank stock return-generating model:

$$R_{jt} = \alpha_j + \beta_{1j} R_{mt} + \beta_{2j} R_{It} + \varepsilon_{jt} \tag{6.1}$$

where R_{jt} is the bank common stock return, R_{mt} is the market return, and R_{It} is the innovation in the interest rate variable. Coefficients α_j and β_{1j} are analogous to the alpha and beta coefficients of the market model, and β_{2j} represents interest rate risk. Since then, numerous researchers have studied the pricing of interest rate risk with varying results. While Stone (1974) and others did not place an a priori restriction on the sign of β_{2j}, the nominal contracting hypothesis implies that it should be positive. This is because the maturity of bank assets is typically longer than that of liabilities.[1] Support for this hypothesis was found by Flannery and James (1984) but not by French et al. (1983).

An important issue in the empirical investigation of the two-factor model is the specification of an appropriate interest rate factor. Theoretical consideration of factor analysis requires that two factors, R_{mt} and R_{It}, be orthogonal whereby choice of the second factor (R_{It}) would not influence the first factor loading (β_{1j}). The resolution of this constraint requires a robust technique for determining the unexpected changes in the interest rate that is uncorrelated with the market return. There are three approaches to specify the interest rate factor (Appendix 2). In the first approach, the expected change in the interest rate is estimated using the high correlation between the observed interest rate and the market rate. The residual – difference between observed and estimated rates – is used as the interest rate factor. The second approach is to identify and estimate a univariate ARMA model for the interest rate variable and use the residuals from the ARMA model as the second factor. In the third approach, the interest rate variable (R_{It}) and the market return (R_{mt}) are treated as the components of a bivariate vector, which is modeled as a vector ARMA process. The estimated model provides the unanticipated change in interest rate variable to be used as the second factor in the augmented CAPM, Eq. 6.1. Srivastava et al. (1999) discuss the alternate ways of specifying the innovations in the interest rate variable and its influence on the pricing of the interest rate risk. In this paper, the error term from the regression of interest rates on market returns is used as the orthogonal interest rate factor in Eq. 6.1.

6.2.1 Pricing of Interest Rate Risk

In addition to changes in the level of expected or unexpected inflation, changes in other economic conditions produce effects on interest rate risk. For example, according to the intertemporal model of the capital market (Merton 1973; Cox et al. 1985), a change in interest rates alters the future investment opportunity set; as a result, investors require additional compensation for bearing the risk of such changes. Similarly, changes in the investor's degree of risk aversion, default risk, or maturity risk of bank financial assets cause additional shifts in the future investment opportunities for the bank stockholders. The specific choice of the bond market index for the two-variable model determines what unexpected change is captured by the coefficient $\beta_{2j.}$

The nominal return on a debt security, R, is expressed as

$$R = R_{rf} + MRP + DRP + LP \tag{6.2}$$

where R_{rf} is the real risk-free rate plus an inflation premium, DRP is the default-risk premium, MRP is the maturity-risk premium, and LP is the liquidity-risk premium.

[1]The sign of β_{2j} is negative when changes in bond yields and not the bond market return are used as the interest rate variable (see Sweeney and Warga 1986).

A change in nominal return consists of changes in the risk-free, liquidity-risk, default-risk, and maturity-risk rate. A change in the short-term risk-free rate can be attributed to changes in real rate or short-term inflation. A sizeable change in the real rate takes place over a longer time horizon. Therefore, it should not significantly impact monthly stock returns and cause the interest rate risk.[2] However, Fama and Gibbons (1982) suggested that changes in real rate may cause changes in the magnitude of interest risk from period to period. Fama and Schwert (1977) argued that, in equilibrium, the risk premium in the short-term interest rate is compensation to the investor for changes in the level of *expected* inflation. However, French et al. (1983) pointed out that the interest rate risk is due to the *unexpected* changes in inflation. More specifically, the interest rate risk can be viewed as the compensation for expected or unexpected changes in the level of short-term inflation. Maturity risk is associated with the changes in the slope of the yield curve that caused a number of economic factors, such as supply and demand of long-term credit, long-term inflation, and risk aversion. Estimates of β_{2j} using short-term T-bill returns as the second factor account for the changes in short-term inflation, whereas estimates of β_{2j} using long-term T-note returns account for the unexpected changes in long-term inflation as well as maturity risk. The inflation expectation horizon approximates that of the maturity of the debt security. Estimating coefficient β_{2j} using BAA bond returns as the interest rate factor explains the fluctuations in bank stock return due to unexpected changes in long-term inflation, maturity-risk premium, default-risk premium, or a combination of risk premia.[3]

As the risk premia are additive, in Eq. 6.2, the magnitude β_{2j} depends on the choice of the interest rate index. A priori expectations about the relative size of β_{2j} are:

$$\begin{aligned} \beta 2j\{\text{6-month T-bill}\} &> \beta 2j\{\text{3-month T-bill}\} \\ \beta_{2j}\{\text{7-year T-note}\} &> \beta_{2j}\{\text{6-month T-bill}\} \\ \beta_{2j}\{\text{BAA-rated bond}\} &> \beta_{2j}\{\text{7-year T-note}\} \end{aligned} \tag{6.3}$$

where the debt security within the brackets identifies the choice of the bond market index. A number of researchers have indicated that bank stock returns are sensitive to unexpected changes in long-term but not short-term interest rates. This observation is consistent with the expectations expressed in Eq. 6.3. However, we are unable to isolate and identify which component of the interest rate risk is being

[2]Federal Reserve's intervention changes the short-term interest rates. These rate changes take place after considerable deliberation and are often anticipated by financial markets. Hence, it neither generates interest rate innovations nor produces interest rate risk.

[3]Liquidity premium is ignored in our discussion and subsequent model construction because our sample does not include financially distressed banks, so there are insufficient variations in the liquidity premium.

priced when a long-term bond return is used. It is feasible to observe a significant change in stock returns due to unexpected changes in default-risk premium (or maturity-risk premium), whether or not nominal contracting hypothesis is valid. The two-factor model can ascertain the pricing of interest rate risk without identifying the source. Commercial banks have a variety of nominal assets and liabilities with different sensitivities to unexpected changes in short-term inflation, maturity risk, and default risk. In the next section, we propose an asset-pricing model in which the interest rate factor of the two-factor model is replaced by three factors, each of which represents a different risk component of the bond return.[4]

6.3 Multi-Risk Premia Asset-Pricing Model

We propose a hybrid asset-pricing model to investigate the interest rate sensitivity of bank stock returns.[5] The traditional CAPM is augmented by three additional APT-type factors to account for unexpected changes in the inflation premium, the maturity-risk premium, and the default-risk rating. Hence, the proposed return-generating model is written as

$$\begin{aligned} R_{jt} = \alpha_j + \beta_{1j}R_{mt} + \beta_{2Pj}\Delta P_t + \beta_{2Mj}\Delta MR_t \\ + \beta_{2Dj}\Delta DR_t + \varepsilon_{jt} \end{aligned} \tag{6.4}$$

where ΔP_t, ΔMR_t, and ΔDR_t are the proxies for the innovations in inflation, maturity risk, and default risk, respectively (specification of these principal component factors consistent with APT is discussed in the Appendix). Coefficients β_1, β_{2P}, β_{2M}, and β_{2D} are the measures of market risk, inflation risk, maturity risk, and default risk, respectively. The expected return is given by CAPM:

$$E\left(R_j\right) = \alpha_j + \beta_{1j}E(Rm) \tag{6.5}$$

However, systematic risk is determined by the integrated model:

$$\begin{aligned} \text{Total systematic risk} = \text{market risk} + \text{inflation risk} + \text{maturity risk} \\ + \text{default risk} \end{aligned} \tag{6.6}$$

[4]An appropriate interest rate variable that should be used to examine the pricing of the interest rate risk in the two-variable framework is not easily available. However, one could construct an index composed of long-term US bonds and corporate bonds with duration equal to the net duration of bank assets and appropriate default risk. This interest rate variable will identify the true pricing of the interest rate risk.

[5]This model's conceptual framework is provided by Chen et al. (1986). However, their factors are not orthogonal.

The three APT-type factors have managerial implications regarding bank's asset and liability management. The first factor, ΔP_t, has implications pertaining to the management of the treasury securities portfolio held by the bank. The value of this portfolio is sensitive to the changes in the short-term interest rate. In addition, these changes impact the short-maturity funding gap. The null hypothesis $\beta_{2P} = 0$ implies that management has correctly anticipated future changes in short-term inflation and has taken steps to correctly hedge the pricing risk through the use of derivative contracts and minimize the short-maturity funding gap.

The second factor, ΔMR_t, has implications regarding the bank's long-term assets and liabilities. The market value of a bank's net worth is very sensitive to changes in the slope of the term structure. The null hypothesis $\beta_{2M} = 0$ conjectures that management has correctly anticipated future changes in the slope of the term structure and has immunized the institution's net worth by a sensible allocation assets in the loan portfolio, sale of loans to the secondary markets, securitization of loans, and the use of derivative contracts. The third factor, ΔDR_t, relates to the management of loan losses and overall default-risk rating of bank assets. The null hypothesis $\beta_{2D} = 0$ infers that management has correctly anticipated future loan losses due to changes in exogenous conditions, and subsequent loan losses, however large, will not adversely affect the stock returns.

6.4 Data Description

Month-end yield for 3-month T-bill, 6-month T-bill, 1-year T-note, 7-year T-note, and BAA corporate bonds, for the period January 1974 to December 1990, were obtained, and monthly returns were calculated.[6] Return on the CRSP equally weighted index of NYSE stocks was used as the market return. Month-end closing prices and dividends for a sample of 88 banks were obtained from Compustat's Price, Dividend, and Earning (PDE) data tape, and monthly returns were calculated. Availability of continuous data was the sole sample selection criteria. This selection criterion does introduce a survivorship bias. However, it was correctly pointed out by Elyasiani and Iqbal (1998) that the magnitude of this bias could be small and not affect the pricing of interest risk. Equally weighted portfolios of bank returns were calculated for this study. The total observation period is divided into three contrasting economic and bank regulatory periods: (1) an increasing but moderately volatile interest rate period from January 1974 to December 1978 in a highly regulated environment; (2) a high interest rate and high-volatility period from January 1979 to December 1984, during which there was gradual deregulation of the industry; and (3) a low interest rate and low-volatility period from January 1985 to December 1990, during which many regulatory changes were made in response to banks loan loss problems. The descriptive statistics of the sample are

[6]The return on a bond index is calculated from the yield series as $R_{It} = -(Y_{It} - Y_{I,t-1})/Y_{It}$ where Y_{It} is the bond index yield at time t.

Table 6.1 Summary statistics of portfolio returns and interest rate yields

	1974–1978	1979–1984	1985–1990
Bank portfolio return	9.66%[a]	22.28%	6.48%
	(5.95%)[b]	(4.93%)	(6.17%)
Portfolio beta[c]	0.7091	0.7259	1.0102
CRSP market return	22.05%	22.05%	8.58%
	(7.63%)	(5.39%)	(5.36%)
3-month T-bill yield	6.21%	10.71%	6.92%
	(1.27%)	(2.41%)	(1.00%)
6-month T-bill yield	6.48%	10.80%	7.02%
	(1.27%)	(2.21%)	(0.95%)
1-year T-note yield	7.05%	11.70%	7.62%
	(1.25%)	(2.27%)	(1.00%)
7-year T-note yield	7.23%	11.91%	8.67%
	(0.54%)	(1.79%)	(1.09%)
BAA bond yield	9.66%	14.04%	10.84%
	(0.68%)	(1.97%)	(0.98%)
Monthly inflation[d]	0.78%	0.46%	0.25%
	(0.69%)	(0.50%)	(0.23%)
Maturity-risk premium[e]	0.68%	0.20%	1.05%
	(0.88%)	(1.32%)	(0.72)
Default-risk premium[f]	1.94%	2.13%	2.17%
	(0.77%)	(0.76%)	(0.46%)

[a]Average returns and yields are reported on annualized basis
[b]Standard deviation is in parenthesis
[c]Estimated using single-index market model
[d]Measured by the relative change in consumer price index
[e]Yield differential between 7-year T-note and 1-year T-note
[f]Yield differential between BAA-rated corporate bond and 7-year T-note

summarized in Table 6.1. The average monthly inflation, as measured by the relative change in the consumer price index, was highest during 1974–1978 and lowest during 1985–1990. The average default-risk premium was of the same order of magnitude for all the three periods. The average maturity-risk premium was high during 1985–1990 indicating a steeper yield curve. For the period 1979–1984, the average maturity-risk premium was low, but its standard deviation was high. This indicates a relatively less steep but volatile yield curve.

The bank portfolio's average return (9.66 %), for the period 1974–1978, was much smaller than the average market return (22.05 %). For the period 1979–1984, both returns increased dramatically; the portfolio's average return (22.28 %) was about the same as the average market return (22.05 %). For the period 1985–1990, the portfolio and markets average return both dropped dramatically to 6.48 % and 8.58 %, respectively. The estimated portfolio beta increased from about 0.7 in the two earlier periods to about 1.0 in the latest period.

6.5 Empirical Results

6.5.1 Two-Variable Regression Model and Pricing of Interest Rate Risk

The estimated coefficients of the two-variable regression model are presented in Table 6.2.[7] For the period 1974–1978, a period characterized by low and stable interest rates, the interest rate beta was statistically insignificant with all the interest rate factors except the 7-year T-note returns. For the period 1979–1984,

Table 6.2 Two-variable model of the bank stock returns: $R_{jt} = \alpha_j + \beta_{1j}R_{mt} + \beta_{2j}R_{It} + \varepsilon_{jt}$

		Interest rate variable[a]				
Period	Estimates	3-month T-bill	6-month T-bill	1-year T-note	7-year T-note	BAA Bond
1974–1978	Constant, α	−0.0046	−0.0045	−0.0043	−0.0033	−0.0042
		(−0.18)[b]	(−0.17)	(−0.16)	(−0.14)	(−0.16)
	Market beta, β_1	0.6921	0.6886	0.6897	0.6845	0.6987
		(15.05)	(14.67)	(15.17)	(16.47)	(15.76)
	Interest rate beta, β_2	0.0566	0.0608	0.0700	0.2893	0.2248
		(1.01)	(1.07)	(1.24)	(2.77)	(0.93)
	R^2	0.828	0.828	0.830	0.845	0.828
	F-statistic	137.3	137.7	138.8	156.3	136.9
1979–1984	Constant, α	0.0052	0.0052	0.0052	0.0052	0.0052
		(1.49)	(1.53)	(1.57)	(1.70)	(1.62)
	Market beta, β_1	0.7261	0.7258	0.7262	0.7262	0.7264
		(11.73)	(12.01)	(12.37)	(13.33)	(12.74)
	Interest rate beta, β_2	0.1113	0.1425	0.1714	0.3582	0.6083
		(3.47)	(3.99)	(4.41)	(5.97)	(5.15)
	R^2	0.684	0.699	0.716	0.755	0.732
	F-statistic	74.8	80.1	87.2	106.6	94.4
1985–1990	Constant, α	−0.0018	−0.0018	−0.0018	−0.0018	−0.0018
		(−0.51)	(−0.51)	(−0.52)	(−0.53)	(−0.52)
	Market beta, β_1	1.010	1.010	1.010	1.010	1.010
		(15.31)	(15.30)	(15.45)	(15.70)	(15.57)
	Interest rate beta, β_2	0.0508	0.0374	0.1036	0.1752	0.2741
		(0.51)	(0.40)	(1.22)	(1.95)	(1.62)
	R^2	0.773	0.772	0.777	0.794	0.780
	F-statistic	117.3	117.1	120.0	125.1	122.5

[a]The error term from the regression of interest rate on market return is used as the appropriate orthogonal interest rate variable
[b]t-statistics are in the parenthesis

[7]Test of two-variable model for the period 1991–2007 indicated that interest rate risk is not priced. Tables can be provided to interested readers.

the interest rate beta was statistically significant with all the interest factors, and its magnitude varied substantially. This period was characterized by a relatively flat but volatile yield curve. For the period 1985–1990, the interest rate beta was statistically significant only with the 7-year T-note returns as the interest rate factor. In general, the estimated value of β_{2j} for the periods 1974–1978 and 1985–1990 is smaller and less statistically significant than the value for the period 1979–1984.

The magnitude of interest rate beta, whether significant or not, varies with the choice of the interest rate variable. The size of β_{2j} increases with the choice of securities in the following order: 3-month T-bill, 6-month T-bill, 1-year T-note, 7-year T-note, and BAA-rated bond (except β_{2j} estimated using BAA bond return for the period 1974–1978 and the 6-month T-bill return in 1985–1990). This observation validates inequality (6.3) in all the periods and suggests the expanding nature of the investment opportunity set with increased horizon (Merton 1973). As stated earlier, the difference between β_{2j} estimated using 7-year T-note returns and that using 3-month T-bill returns measures the effect on the bank stock returns due to the unexpected changes in the maturity-risk premium. Further, difference between β_{2j} estimated using BAA bond returns and that using 7-year T-note returns measures the effect on the bank stock returns due to the unexpected changes in the default-risk premium.

The fact that bank stock returns are more sensitive to the long-term interest rates than to short-term interest rates is consistent with our expectation about the size of β_2 expressed in inequalities (6.3). Similar results were reported by other researchers (such as Unal and Kane (1988) and Chen et al. (1986)). A shift of focus from short-term to long-term inflation expectation could explain this result. An alternative explanation is that bank balance sheet returns are better approximated by long-term than by short-term bond returns. To the extent that balance maturity mismatches occur, they should be related to long-term bond returns. The reason is simply that long-term bond returns include the present value of more future period returns than do short-term bond returns. That is, long-term bond returns include price changes not included in short-term bond returns. The price changes in the long-term bond returns represent price changes in long-term bank contracts. The most representative term equals the term of assets or liabilities, whichever is longer. If the maturity mismatch is a net asset (liability) position, then the long-term bond maturity would reflect the asset (liability) maturity. The estimate of a large, positive coefficient (β_{2j}) for 7-year T-notes and BAA bonds implies that banks mismatch in favor of assets with relatively long maturities.

The plausible causes for the change in the interest risk from period to period are (1) changes in real returns, as reported by Fama and Gibbons (1982), (2) unexpected changes in short-term and long-term inflation expectation, (3) shift of focus from short-term to long-term inflation expectation, (4) unexpected changes in risk aversion, (5) unexpected changes in the default risk of bank's nominal contracts, and (6) structural instability of the systematic interest rate equation used to extract interest rate innovations. The estimated coefficients of multi-risk premia model will shed some light on this issue.

6.5.2 Multi-risk Premia Model

The estimates of the first systematic interest rate Eq. 6.11, which specifies innovations in the short-term default-free bond returns, are reported in Table 6.3. The ordinary least square (OLS) estimates were rejected because of the presence of serial autocorrelation as indicated by the Durbin-Watson statistic (see Woolridge 2009). A Durbin-Watson statistic equal to 2 indicates the absence of any serial autocorrelation. The generalized least square (GLS) estimates seemed more appropriate because the Durbin-Watson statistic was approximately equal to 2 and the value of R^2 was higher. Results reported in Table 6.3 exhibit a significant correlation between market return and short-term interest rate variable in the period 1974–1978, but not in the periods 1979–1984 and 1985–1990. However, a low value of R^2 indicates that the relationship expressed in Eq. 6.4 is not as robust as one would have preferred. This approach of extracting changes in short-term inflation was used by Fama and Schwert (1977) and French et al. (1983). Their estimation period overlapped our estimation period 1974–1978 but not the later ones. In Table 6.4 we present the results of the second systematic interest rate Eqs. 6.12a and 6.12b. The estimated coefficient, θ_1, determined the average slope of the yield curve during the estimation period. The OLS estimates were rejected because of the presence of serial correlation indicated by the Durbin-Watson statistics. The term structure coefficient θ_1 was significant in all the periods for *Specifications a* and *b*. However, coefficients θ_2 of *Specification b* was

Table 6.3 Estimating a proxy for ex-post unexpected inflation: $R_{ST,t} = \delta_0 + \delta_1 R_{mt} + \varepsilon_t$

	Estimated coefficients		
	1974–1978	1979–1984	1985–1990
Ordinary least square			
Constant, δ_0	−0.0068	−0.0070	0.0037
	(−0.87)[a]	(−0.45)	(0.89)
Market linkage coefficient, δ_1[b]	0.2987	0.5791	−0.0891
	(2.97)	(2.19)	(−1.13)
R^2	0.132	0.076	0.018
Durbin-Watson statistic	1.90	1.19	1.12
Generalized least square			
Constant, δ_0	−0.0075	−0.0013	0.0044
	(−0.82)	(−0.08)	(0.83)
Market linkage coefficient, δ_1[b]	0.3173	0.3136	−0.0806
	(3.16)	(1.44)	(−1.15)
R^2	0.145	0.306	0.242
Durbin-Watson statistic	1.99	1.97	1.92

Ordinary Least Square estimates were rejected because of the presence of serial correlation indicated by Durbin-Watson statistic

$R_{ST,t}$ and R_{mt} are 3-month T-bill and CRSP equally-weighted market returns

[a]t-statistics are in the parenthesis

[b]Coefficient δ_1 measures the stock–bond market linkage

Table 6.4 Estimating alternate proxies for maturity-risk premia: Specification a : $R_{LT,t} = \theta_0 + \theta_1 R_{ST,t} + \varepsilon_t$, Specification b : $R_{LT,t} = \theta_0 + \theta_1 R_{ST,t} + \theta_2 R_{mt} + \varepsilon_t$

	1974–1978	1979–1984	1985–1990
Specification a: generalized least square			
Constant, θ_0	−0.0028	−0.0053	0.0017
	(−1.02)[a]	(−2.11)	(0.66)
Term structure coefficient, θ_1[b]	0.3988	0.5201	0.8183
	(10.30)	(16.29)	(14.96)
R^2	0.675	0.846	0.820
Durbin-Watson statistic	1.98	1.83	1.84
Specification b: generalized least square			
Constant, θ_0	−0.0024	−0.0064	0.0011
	(−0.84)	(−2.47)	(0.48)
Term structure coefficient, θ_1[b]	0.4070	0.5122	0.8105
	(10.02)	(15.78)	(15.34)
Market linkage coefficient, θ_2[c]	−0.0216	0.0576	0.0944
	(−0.68)	(1.07)	(2.56)
R^2	0.677	0.849	0.835
Durbin-Watson statistic	1.98	1.82	1.80

Ordinary least square estimates were rejected because of the presence of serial correlation indicated by Durbin-Watson statistic
$R_{LT,t}$ and $R_{ST,t}$ are 7-year T-note and 1-year T-note returns
[a]t-statistics are in the parenthesis
[b]Coefficient θ_1 measures the average slope of the yield curve
[c]Coefficient θ_2 accounts for the stock–bond market linkage

significant only during 1985–1990. The errors from Specification a for periods 1974–1978 and 1979–1984 and from *Specification b* for the period 1985–1990 were used to specify the second risk factor ΔMR_t. Estimates of the third systematic interest rate Eqs. 6.13a and 6.13b are presented in Table 6.5. As before, the GLS estimates were deemed appropriate. The errors from *Specification a* for the periods 1974–1978 and 1979–1984 and from *Specification b* for the period 1985–1990 were used to specify the unexpected change in default risk, ΔDR_t.

Results of the multi-risk premia model are presented in Table 6.6. The inflation beta is statistically significant in the periods 1979–1984 and 1985–1990 but not in the period 1974–1978. It was shown in Table 6.3 that quasi-differenced short-term interest rates were correlated with the market return for the period 1974–1978 but not for the periods 1979–1984 and 1985–1990. Hence, one could argue that when short-term interest rates are correlated with the market return (i.e., 1974–1978), the error term from Eq. 6.8 contains no systematic information. This results in the inflation beta being insignificant and the interest rate risk not priced with respect to the short-term rates within the context of the two-factor model. A corollary is that, when short-term interest rates are uncorrelated with the market return (i.e., 1979–1984 and 1985–1990), the error term from Eq. 6.11 contains valuable information leading to the inflation beta being significant. The maturity beta was found

Table 6.5 Estimating alternate proxies for the default-risk premia: Specification a : $R_{BAA,t} = \varphi_0 + \varphi_1 R_{LT,t} + \varepsilon_t$, Specification b : $R_{BAA,t} = \varphi_0 + \varphi_1 R_{LT,t} + \varphi_2 R_{mt} + \varepsilon_t$

	1974–1978	1979–1984	1985–1990
Specification a: generalized least square			
Constant, φ_0	−0.0021	−0.0039	0.0010
	(−0.60)[a]	(−1.12)	(0.66)
Default-risk coefficient, φ_1[b]	0.0842	0.4129	0.4734
	(2.27)	(12.27)	(13.47)
R^2	0.553	0.800	0.772
Durbin-Watson statistic	1.88	1.97	1.99
Specification a: generalized least square			
Constant, φ_0	−0.0019	−0.0051	0.0006
	(−0.54)	(−1.41)	(0.48)
Term structure coefficient, φ_1[b]	0.0861	0.3976	0.4637
	(2.30)	(11.35)	(13.83)
Market linkage coefficient, φ_2[c]	−0.0066	0.0491	0.0577
	(−0.49)	(1.49)	(2.39)
R^2	0.555	0.806	0.785
Durbin-Watson statistic	1.87	1.97	2.00

Ordinary least square estimates were rejected because of the presence of serial correlation indicated by Durbin-Watson statistic

$R_{BAA,t}$ and $R_{LT,t}$ are BAA corporate bond and 7-year T-note returns

[a]t-statistics are in the parenthesis

[b]Coefficient φ_1 measures the average yield differential between $R_{BAA,t}$ and $R_{LT,t}$

[c]Coefficient φ_2 accounts for the stock–bond market linkage

to be statistically significant in the periods 1974–1978 and 1979–1984 but not for the period 1985–1990. Results in Table 6.4 (Specification b) showed that long-term interest rates were correlated with the market return for the period 1985–1990 but not for the periods 1974–1978 and 1979–1984. Hence, we posit that when long-term interest rates are correlated with the market return (i.e., 1985–1990), the error term from Eq. 6.12b contains no systematic information. This results in the maturity beta being insignificant. A corollary is that, when long-term interest rates are uncorrelated with the market return (i.e., 1974–1978 and 1979–1984), the error term from Eq. 6.12b contains valuable information producing a significant maturity beta and the interest rate risk is priced with respect to the long-term rates within the context of the two-factor model. The default beta was found to be statistically significant in the period 1985–1990 but not for the periods 1974–1978 and 1979–1984. The economic factors that lead to significant correlation between market returns and long-term interest rates (Eq. 6.12b) or between market returns and BAA-rated bond returns (Eq. 6.13b) caused the interest rate risk to be priced with respect to the long-term rates within the context of the two-factor model (1985–1990). Since the correlation between market return and interest rate changes over time, the interest rate risk also changes over time.

Table 6.6 Multi-risk premia model of the bank stock returns: $R_{jt} = \alpha_{0j} + \beta_{1j}R_{mt} + \beta_{2Qj}\Delta P_t + \beta_{2Mj}\Delta MR_t + \beta_{2Dj}\Delta DR_t + \varepsilon_{jt}$

	Estimated coefficients[a]		
	1974–1978	1979–1984	1985–1990
Constant, α_0	−0.0056	0.0041	−0.0047
	(−1.84)[b]	(1.04)	(−1.31)
Market beta, β_1	0.7397	0.6833	0.7185
	(16.94)	(10.55)	(17.55)
Inflation beta, β_{2P}	0.0698	0.1292	−0.2157
	(1.27)	(3.45)	(−2.29)
Maturity beta, β_{2M}	0.5350	0.4881	−0.2476
	(2.87)	(2.93)	(−1.38)
Default beta, β_{2D}	−0.1973	0.1153	−0.5696
	(−0.56)	(0.43)	(−1.94)
R^2	0.849	0.754	0.851
Durbin-Watson statistic	2.00	2.04	1.98

Ordinary least square estimates were rejected because of the presence of serial correlation indicated by Durbin-Watson statistic

Portfolio beta estimated using single-index model are 0.7091, 0.7259, and 1.0102 for the periods 1974–1978, 1979–1984, and 1985–1990, respectively

[a]Generalized least square estimates

[b]t-statistics are in the parenthesis

Negative inflation and maturity betas for the period 1985–1990 need some explanation because it is contrary to a priori expectation. For the period 1985–1990, the bank portfolio and market return both dropped dramatically to 6.48 % and 8.58 %, respectively. However, the estimated portfolio beta increased from about 0.7 in the two earlier periods to about 1.0 in this period (beta estimated independently by the single-index model). Consequently, the market factor alone will overestimate the bank portfolio's expected return. The negative values of inflation beta and maturity beta (though insignificant) correct the overestimation. Economic factors and regulatory changes that fuelled M&A activities during this period must have been such that they increased the portfolio beta without increasing the *ex-post* portfolio return. Some of these unidentifiable factors are negatively correlated with the interest rates. One of the shortcomings of the factor analytic approach is that factors are at times unidentifiable. In spite of difficulties in explaining some of the results for the period 1985–1990, the multi-risk premia model does provide greater insight into the pricing of interest rate risk.

6.6 Conclusions

In this paper, we examine the interest rate sensitivity of commercial bank returns covering three distinct economic and regulatory environments. First, we investigate the pricing of the interest rate risk within the framework of the two-factor model.

Our results indicate that interest rate risk was priced during 1979–1984 irrespective of the choice of interest rate variable. However, during the periods 1974–1978 and 1985–1990, bank stock returns were sensitive only to the unexpected changes in the long-term interest rates. Next, we tested to an asset-pricing model in which the traditional CAPM is augmented by three additional factors to account for unexpected changes in the inflation, the maturity premium, and default premium. Our results show that the inflation beta was significant for the periods 1979–1984 and 1985–1990, but not for the period 1974–1978; the maturity beta was significant for the periods 1974–1978 and 1979–1984 but not for the period 1985–1990; and the default beta was significant for the period 1985–1990 but not for the periods 1974–1978 and 1978–1984.

We can infer that when short-term interest rates are correlated with the market return, the innovations in short-term interest rate are indeed white noise. However, innovations in short-term interest rates contain valuable information when short-term interest rates are uncorrelated with the market return. This will lead to a significant inflation beta and the interest rate risk will be priced with respect to the short-term rates within the context of the two-factor model. We can also infer that when long-term interest rates are correlated with the market return, the innovations in long-term interest rate are indeed white noise. However, innovations in long-term interest rates contain valuable information when long-term interest rates are uncorrelated with the market return. This results in a significant maturity beta and priced interest rate risk with respect to the long-term rates within the context of the two-variable model.

Appendix 1: An Integration of CAPM and APT

The traditional CAPM is augmented by three additional APT-type principal component factors (Johnson and Wichern 2007) to account for unexpected changes in the inflation premium, the maturity-risk premium, and the default-risk rating. Hence, the proposed return-generating model is written as

$$R_{jt} = \alpha_j + \beta_{1j}R_{mt} + \beta_{2Pj}\Delta P_t + \beta_{2Mj}\Delta MR_t + \beta_{2Dj}\Delta DR_t + \varepsilon_{jt} \tag{6.7}$$

where ΔP_t, ΔMR_t, and ΔDR_t are the proxies for the innovations in inflation, maturity risk, and default risk, respectively. Coefficients β_1, β_{2P}, β_{2M}, and β_{2D} are the measures of systematic market risk, inflation risk, maturity risk, and default risk, respectively, and are consistent with APT.

Principal Component Factor Specification

An important issue in the empirical investigation of this model is the specification of an appropriate inflation and maturity-risk and default-risk factors. Being innovations in economic variables, these factors cannot be predicted using past information. Hence, they must meet the following conditions:

$$\begin{aligned} E(\Delta P_t|t-1) &= 0 \\ E(\Delta MR_t|t-1) &= 0 \\ E(\Delta DR_t|t-1) &= 0 \end{aligned} \tag{6.8}$$

where $E(.|t-1)$ is the expectation based on available information at $t-1$. Theoretical consideration dictates that the choice of a factor (say ΔMR_t) should not influence any other factor loadings (say $\exists_{1j}$). Hence, there should not be any contemporaneous cross-correlations between the factors. Hence,

$$\begin{aligned} E(\Delta P_t.\Delta MR_t) &= 0 \\ E(\Delta P_t.\Delta DR_t| &= 0 \\ E(\Delta R_t.\Delta DR_t) &= 0 \end{aligned} \tag{6.9}$$

In addition, there are no common market shocks that may influence any of the risk factors. So the following conditions

$$\begin{aligned} E(R_{mt}.\Delta P_t) &= 0 \\ E(R_{mt}.\Delta MR_t) &= 0 \\ E(R_{mt}.\Delta DR_t) &= 0 \end{aligned} \tag{6.10}$$

must be satisfied. We use a generated regressor approach to construct orthogonal inflation-risk, maturity-risk, and default-risk factors.

Economic factors that lead to changes in the equity market return also induce term structure movements. This leads to a significant correlation between the stock market index and the bond market index. Hence, we use the stock market return to forecast systematic changes in the short-term interest rates. The innovations in the short-term default-free bond return specify the unexpected changes in inflation. Hence, the first systematic interest rate equation is written as

$$R_{ST,t} = \delta_0 + \delta_1 R_{mt} + \varepsilon_{It} \tag{6.11}$$

where R_{mt} is the return on the stock market index and $R_{ST,t}$ is the return on the short-term default-free bond return. The error term in Eq. 6.11, ε_{It}, specifies the unexpected change in inflation and is the generated regressor which serves as a proxy for factor ΔP_t. Fama and Schwert (1977) and French et al. (1983) employ similar approaches to specify a proxy for changes in inflation.

The yield differential between short-term and long-term default-free bond represents the slope of the yield curve. The relationship used to construct the maturity-risk factor is given by the second systematic interest rate equation:

$$R_{LT,t} = \theta_0 + \theta_1 R_{ST,t} + \varepsilon_{Mt} \tag{6.12a}$$

where θ_1 measures the average slope of the yield curve. Alternately, we will also test the following systematic interest rate equation:

$$R_{LT,t} = \theta_0 + \theta_1 R_{ST,t} + \theta_2 R_{mt} + \varepsilon_{Mt} \tag{6.12b}$$

Inclusion of the independent variable, R_{mt}, will control for the stock–bond market linkage. The estimated coefficient θ_1 measures the slope of the yield curve, and it determines the maturity-risk premium on long-term assets and liabilities. The error term in Eqs. 6.12a or 6.12b, ε_{Mt}, specifies the unexpected change in maturity risk and will be used as the generated regressor which serves as a proxy for factor ΔMR_t.

The portfolio of commercial banks used for this study consisted of money center and large national and regional banks. Most of these banks were well capitalized and had acquired a balanced portfolio of assets. The average default-risk rating of the portfolio of banks should be close to the default-risk rating of a high-grade corporate bond. Hence, the yield differential between BAA corporate bond and long-term treasury security is used to construct the default-risk factor. Our third systematic interest rate equation is written as

$$R_{BAA,t} = \varphi_0 + \varphi_1 R_{LT,t} + \varepsilon_{Dt} \tag{6.13a}$$

Alternately, we will also test the following systematic interest rate equation:

$$R_{BAA,t} = \varphi_0 + \varphi_1 R_{LT,t} + \varphi_2 R_{mt} + \varepsilon_{Dt} \tag{6.13b}$$

where φ_1 measures the average default risk on BAA corporate bond. Inclusion of the independent variable, R_{mt}, will control for the stock–bond market linkage. The error term in Eqs. 6.13a or 6.13b, ε_{Dt}, specifies the unexpected change in default risk and serves as the generated regressor which will be used as the proxy for factor ΔDR_t.

Appendix 2: Interest Rate Innovations

An important issue in the empirical investigation of the two-factor model is the specification of an appropriate interest rate factor. Theoretical consideration of factor analysis requires that two factors, R_{mt} and R_{It}, be orthogonal whereby choice of the second factor (R_{It}) would not influence the first factor loading (β_{Ij}). The resolution of this constraint requires a robust technique for determining the unexpected changes in the interest rate that is uncorrelated with the market return. The three approaches to specify the interest rate factor are presented here.

Orthogonalization Procedure

Economic factors producing changes in the market return also induced term structure movements. This leads to a high correlation between the market factor and the interest rate factor. Hence, market return can be used as an instrument variable to forecast the expected changes in the interest rates. To find the unexpected component of interest rates, the expected interest rate is purged by regressing R_{It} on R_{mt} and using the residuals. The systematic interest rate risk equation is

$$R_{It} = \delta_0 + \delta_1 R_{mt} + \varepsilon it \tag{6.14}$$

The residuals, εit, are the unsystematic interest rates and are used to replace R_{It} in Eq. 6.14. The validity of this approach has been questioned on methodological grounds. It is pointed out that this orthogonalization procedure produces biased estimates of coefficients (intercept and β_{1j} in Eq. 6.14) and that the deficiency of β_{2j} is not improved. On the other hand, use of an unorthogonal interest rate factor leads to the errors-in-variable problem, i.e., the estimated coefficient β_{2j} also captures some of the effects responsible for changing the market factor, and hence, it is not a true measure of the interest rate risk. Another problem using an unorthogonal interest rate factor stems from the fact that interest rate (R_{It}) is usually autoregressive. Therefore, residuals, εit, from Eq. 6.14 are autocorrelated unless GLS parameter estimation procedure is employed. To use the GLS procedure, a variance-covariance matrix has to be specified, which is not an easy task.

Univariate ARMA Model

The second approach is to identify and estimate an ARMA model for the interest rate variable, R_{It} (Flannery and James (1984)). The unanticipated change in interest rate from the estimated model (i.e., residuals) is used to replace R_{It} in Eq. 6.14. In general, the ARMA model of order (p, q) for the univariate time series, R_{It}, is written as

$$R_{It} = \phi_1 R_{It-1} + \phi_2 R_{It-2} + \ldots + \phi_p R_{It-p} + \mu - \theta_1 e_{I,t-1} - \ldots - \theta_q \varepsilon_{1,t-q} + \varepsilon_{It} \tag{6.15}$$

where ε_{It}, $\varepsilon_{I,t-1}$, ... are identically and independently distributed random errors with mean zero. The ARMA procedure for the modeling of time series data is outlined in Box and Jenkins (1976). The modeling is usually done in three steps. First, a tentative parsimonious model is identified. Second, the parameters are estimated, and diagnostic (Box-Pierce Q) statistics and residual auto correlation plots are examined. The model is acceptable if the time series of residuals is white noise and the Box-Pierce Q statistics are significant.

This approach leads to unbiased estimates of all the coefficients in Eq. 6.14. The shortcoming of the univariate ARMA approach is that valuable information contained in the stock and bond market linkage is ignored. Consequently, coefficient β_{2j} captures some of the effects of economic factors producing stock market changes.

Vector ARMA Model

In recent years vector autoregressive moving average (VARMA) models have proved to be useful tools to describe the dynamic relationship between economic variables. A vector autoregressive moving average model of order (p, q) for k-dimensional time series, $Rt = (R_{1t}, R_2t, \ldots R_{kt})^T$, is generated by the following equation:

$$\Phi(\mathbf{B})R_t = \Theta(\mathbf{B})\varepsilon_t \tag{6.16}$$

$$\Phi(\mathbf{B}) = \left(\mathbf{I} - \phi_1\mathbf{B} - \phi_2\mathbf{B}^2 - \ldots - \phi_p\mathbf{B}^p\right) \tag{6.17}$$

$$\Theta(\mathbf{B}) = \left(\mathbf{I} - \theta_1\mathbf{B} - \theta_2\mathbf{B}^2 - \ldots - \theta_q\mathbf{B}^q\right) \tag{6.18}$$

where **B** is the back shift operator, **I** is kxk unit matrix, and ε_t is a sequence of an independent k-dimensional vector with zero mean and positive definite covariance matrix. The interest rate variable R_{It} and the market return R_{mt} are treated as the components of a bivariate vector. Then vector $(R_{It}, R_{mt})^T$ is modeled as a vector AR process. The estimated model provides the unanticipated change in interest rate variable to be used as the second factor in the augmented CAPM model.

References

Bae, S. C. (1990). Interest rate changes and common stock returns of financial institutions: Revised. *Journal of Financial Research, 13*, 71–79.

Booth, J. R., & Officer, D. T. (1985). Expectations, interest rates, and commercial bank stocks. *Journal of Financial Research, 8*, 51–58.

Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis, forecasting and control*. San Fran: Holden-Day.

Chance, D. M., & Lane, W. R. (1980). A Re-examination of interest rate sensitivity in the common stock of financial institutions. *Journal of Financial Research, 3*, 49–55.

Chen, N., Roll, R., & Ross, S. (1986). Economic forces and the stock market. *Journal of Business, 59*, 383–404.

Choi, J. J., Elyasiani, E., & Kopecky, K. (1992). The sensitivity of bank stock returns to market, interest, and exchange rate risks. *Journal of Banking and Finance, 16*, 983–1004.

Christie, A. A. (1981). The stochastic behavior of common stocks variances: Value, leverage and interest rate effects. *Journal of Financial Economics, 10*, 407–432.

Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985). An intertemporal general equilibrium model of asset prices. *Econometrica, 53*, 363–84.

Elyasiani, E., & Iqbal, M. (1998). Sensitivity of the bank stock returns distribution to changes in the level and volatility of interest rate: A GARCH-M model. *Journal of Banking & Finance, 22*, 535–563.

Fama, E. F., & Gibbons, M. R. (1982). Inflation, real returns and capital investment. *Journal of Monetary Economics, 9*, 297–323.

Fama, E. F., & Schwert, W. G. (1977). Asset returns and inflation. *Journal of Financial Economics, 4*, 115–146.

Flannery, M. J., & James, C. M. (1984). The effect of interest rate changes on the common stock returns of financial institutions. *Journal of Finance, 39*, 1141–1153.

Flannery, M. J., Hameed, A. S., & Harjes, R. H. (1997). Asset pricing time-varying risk premia and interest rate risk. *Journal of Banking and Finance, 21*, 315–335.

French, K. R., Ruback, R. C., & Schwert, G. W. (1983). Effects of nominal contracting hypothesis on stock returns. *Journal of Political Economy, 91*, 70–96.

Johnson, R., & Wichern, D. (2007). *Applied multivariate statistical analysis* (6th ed.). Upper Saddle River: Pearson Education.

Llyod, W. P., & Shick, R. (1977). A test of Stone's two index models of returns. *Journal of Financial and Quantitative Analysis, 12*, 363–376.

Lynge, M., & Zumwalt, J. (1980). An empirical study of interest rate sensitivity of commercial bank returns: A multi index approach. *Journal of Financial and Quantitative Analysis, 15*, 731–742.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica, 41*, 867–887.

Song, F. (1994). A two factor ARCH model for deposit-institution stock returns. *Journal of Money, Credit, and Banking, 26*, 323–340.

Srivastava, S. C., Hamid, S., & Choudhury, A. H. (1999). Stock and bond market linkage in the empirical study of interest rate sensitivity of bank returns. *Journal of Applied Business Research, 15*, 47–58.

Stone, B. (1974). Systematic interest rate risk in a two – index model of returns. *Journal of Financial and Quantitative Analysis, 9*, 709–721.

Sweeney, R. J., & Warga, A. D. (1986). The pricing of interest-rate risk: Evidence from the stock market. *Journal of Finance, 41*, 393–410.

Unal, H., & Kane, E. J. (1988). Two approaches to assessing the interest-rate sensitivity of deposit-institutions' equity returns. In C. Andrew (Ed.), *Research in finance* (Vol. 7, pp. 113–132). Greenwich Conn.: Jai Press.

Woolridge, J. (2009). *Introductory econometrics: A modern approach* (4th ed.). Mason: Cengage Learning.

Yourougou, P. (1990). Interest rate and the pricing of depository financial intermediary common stock: Empirical evidence. *Journal of Banking and Finance, 14*, 803–320.

Nonparametric Bounds for European Option Prices

7

Hsuan-Chu Lin, Ren-Raw Chen, and Oded Palmon

Contents

Abstract

There is much research whose efforts have been devoted to discovering the distributional defects in the Black-Scholes model, which are known to cause severe biases. However, with a free specification for the distribution, one can only find upper and lower bounds for option prices. In this paper, we derive a new nonparametric lower bound and provide an alternative interpretation of Ritchken's (1985) upper bound to the price of the European option. In a series of numerical examples, our new lower bound is substantially tighter than previous lower bounds.

The financial support of National Science Council, Taiwan, Republic of China (NSC 96-2416-H-006-039-), is gratefully acknowledged.

H.-C. Lin (✉)
Graduate Institute of Finance and Banking, National Cheng-Kung University, Tainan, Taiwan
e-mail: hsuanchu@mail.ncku.edu.tw

R.-R. Chen
Graduate School of Business Administration, Fordham University, New York, NY, USA
e-mail: rchen@fordham.edu

O. Palmon
Department of Finance and Economics, Rutgers Business School – Newark and New Brunswick, Piscataway, NJ, USA
e-mail: palmon@business.rutgers.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_7,

This is prevalent especially for out-of-the-money (OTM) options where the previous lower bounds perform badly. Moreover, we present that our bounds can be derived from histograms which are completely nonparametric in an empirical study. We first construct histograms from realizations of S&P 500 index returns following Chen, Lin, and Palmon (2006); calculate the dollar beta of the option and expected payoffs of the index and the option; and eventually obtain our bounds. We discover violations in our lower bound and show that those violations present arbitrage profits. In particular, our empirical results show that out-of-the-money calls are substantially overpriced (violate the lower bound).

Keywords

Option bounds • Nonparametric • Black-Scholes model • European option • S&P 500 index • Arbitrage • Distribution of underlying asset • Lower bound • Out-of-the-money • Kernel pricing

7.1 Introduction

In a seminal paper, Merton (1973) presents for the first time the no-arbitrage bounds of European call and put options. These bounds are nonparametric and do not rely on any assumption.[1] Exact pricing formulas such as the Black and Scholes (1973) model and its variants, on the other hand, rely on strong assumptions on the asset price process and continuous trading. Due to the discreteness of actual trading opportunities, Perrakis and Ryan (1984) point out that option analyses in continuous time limit the accuracy and applicability of the Black-Scholes and related formulas. Relying on Rubinstein's (1976) approach, the single-price law, and arbitrage arguments, they derive upper and lower bounds for option prices with both a general price distribution and discrete trading opportunities. Their lower bound is tighter than that of Merton.

Levy (1985) applies stochastic dominance rules with borrowing and lending at the risk-free interest rate to derive upper and lower option bounds for all unconstrained utility functions and alternatively for concave utility functions. The derivation of these bounds can be applied to any kinds of stock price distribution as long as the stock is "nonnegative beta," which is identical to the assumption of Perrakis and Ryan (1984). Moreover, Levy claims that Perrakis and Ryan's bounds can be obtained by applying the second-degree stochastic dominance rule. However, Perrakis and Ryan do not cover all possible combinations of the risky asset with the riskless asset, and their bounds are therefore wider than those of Levy. Levy also applies the first-degree stochastic dominance rule (FSDR) with riskless assets to prove that Merton's bounds are in fact FSDR bounds and applies the second-degree stochastic dominance rule to strengthen Merton's bounds on the option value. At the same time, Ritchken (1985) uses a linear programming methodology to derive option bounds based on primitive prices in incomplete markets and claims that his bounds are tighter than those of Perrakis and Ryan (1984).

[1]The only assumption is that both option and its underlying stock are traded securities.

With an additional restriction that the range of the distribution of the one-period returns per dollar invested in the optioned stock is finite and has a strictly positive lower limit, Perrakis (1986) extends Perrakis and Ryan (1984) to provide bounds for American options. Instead of assuming that no opportunities exist to revise positions prior to expiration in Levy (1985) and Ritchken (Ritchken 1985), Ritchken and Kuo (1988) obtain tighter bounds on option prices under an incomplete market by allowing for a finite number of opportunities to revise position before expiration and making more restrictive assumptions on probabilities and preferences. The single-period linear programming option model is extended to handle multiple periods, and the stock price is assumed to follow a multiplicative multinomial process. Their results show that the upper bounds are identical to those of Perrakis, while the lower bounds are tighter. Later, Ritchken and Kuo (1989) also add suitable constraints to a linear programming problem to derive option bounds under higher orders of stochastic dominance preferences. Their results show that while the upper bounds remain unchanged beyond the second-degree stochastic dominance, the lower bounds become sharper as the order of stochastic dominance increases.[2]

Claiming that Perrakis and Ryan (1984), Levy (1985), Ritchken (1985), and Perrakis (1986) are all parametric models, Lo (1987) derives semi-parametric upper bounds for the expected payoff of call and put options. These upper bounds are semi-parametric because they depend on the mean and variance of the stock price at maturity but not on its entire distribution. In addition, the derivation of corresponding semi-parametric upper bounds for option prices is shown by adopting the risk-neutral pricing approach of Cox and Ross (1976).[3] To continue the work of Lo (1987), Zhang (1994) and De La Pena et al. (2004), both of which assume that the underlying asset price must be continuously distributed, sharpen the upper option bounds of Lo (1987). Boyle and Lin (1997) extend the results of Lo (1987) to contingent claims on multiple underlying assets. Under an intertemporal setting, Constantinides and Zariphopoulou (2001) derive bounds for derivative prices with proportional transaction costs and multiple securities. Frey and Sin (1999) examine the sufficient conditions of Merton's bounds on European option prices under random volatility. More recently, Gotoh and Konno (2002) use the semi-definite programming and a cutting plane algorithm to study upper and lower bounds of European call option prices. Rodriguez (2003) uses a nonparametric method to derive lower and upper

[2]To further explore the research work of Ritchken and Kuo (1989) under the decreasing absolute risk aversion dominance rule, Basso and Pianca (1997) obtain efficient lower and upper option pricing bounds by solving nonlinear optimization problem. Unfortunately, neither model provides enough information of their numerical examples for us to compare our model with. The Ritchken-Kuo model provides no Black-Scholes comparison, and the Basso-Pianca model provides only some partial information on the Black-Scholes model (we find the Black-Scholes model under 0.2 volatility to be 13.2670 and under the 0.4 volatility to be 20.3185, which are different from what are reported in their paper (12.993 and 20.098, respectively)) which is insufficient for us to provide any comparison.

[3]Inspired by Lo (1987), Grundy (1991) derives semi-parametric upper bounds on the moments of the true, other than risk-neutral, distribution of underlying assets and obtains lower bounds by using observed option prices.

bounds and contributes a new tighter lower bound than previous work. Huang (Huang 2004) puts restrictions on the representative investor's relative risk aversion and produces a tighter call option bound than that of Perrakis and Ryan (1984). Hobson et al. (2005) derive arbitrage-free upper bounds for the prices of basket options. Peña et al. (2010) conduct static-arbitrage lower bounds on the prices of basket options via linear programming. Broadie and Cao (2008) introduce new and improved methods based on simulation to obtain tighter lower and upper bounds for pricing American options. Lately, Chung et al. (2010) also use an exponential function to approximate the early exercise boundary to obtain tighter bounds on American option prices. Chuang et al. (2011) provide a more complete review and comparison of theoretical and empirical development on option bounds.[4]

In this paper we derive a new and tighter lower bound for European option prices under a nonparametric framework. We show that Ritchken's (1985) upper bound is consistent with our nonparametric framework. Both bounds are nonparametric because the price distribution of underlying asset is totally flexible, can be arbitrarily chosen, and is consistent with any utility preference.[5] We compare our lower bound with those in previous studies and show that ours dominate those models by a wide margin. We also present the lower bound result on the model with random volatility and random interest rates (Bakshi et al. 1997; Scott 1997) to demonstrate how easily our model can be made consistent with any parametric structure.[6] Finally, we present how our bounds can be derived from histograms which are nonparametric in an empirical study. We discover violations of our lower bound and show that those violations present arbitrage profits.[7] In particular, our empirical results show that out-of-the-money calls are substantially overpriced (violate the lower bound).

7.2 The Bounds

A generic and classical asset pricing model with a stochastic kernel is

$$S_t = E_t\left[M_{t,T} S_T\right], \tag{7.1}$$

where $M_{t,T}$ is the marginal rate of substitution, also known as the pricing kernel that discounts the future cash flow at time T; $E_t[.]$ is the conditional expectation

[4]Since our paper only provides a nonparametric method on examining European option bounds, our literature review is much limited. For a more complete review and comparison on prior studies of option bounds, please see Chuang et al. (2011).

[5]Christoffersen et al. (2010) provide results for the valuation of European-style contingent claims for a large class of specifications of the underlying asset returns.

[6]Given that our upper bound turns out to be identical to Ritchken's (1985), we do not compare with those upper bound models that dominate Ritchken (e.g., Huang (2004), Zhang (1994) and De La Pena et al. (2004)). Also, we do not compare our model with those models that require further assumptions to carry out exact results (e.g., Huang (2004) and Frey and Sin (1999)), since it is technically difficult to do.

[7]For the related empirical studies of S&P 500 index options, see Constantinides et al. (2009, 2011).

under the physical measure $\mathbb{P}$ taken at time t; and S_t is the value of an arbitrary asset at time t. The standard kernel pricing theory (e.g., Ingersoll (1989)) demonstrates that

$$D_{t,T} = E_t\left[M_{t,T}\right], \tag{7.2}$$

where $D_{t,T}$ is the risk-free discount factor that gives the present value of $1 over the period (t,T). The usual separation theorem gives rise to the well-known risk-neutral pricing result:

$$\begin{aligned} S_t &= E_t\left[M_{t,T}S_t\right] \\ &= E_t\left[M_{t,T}\right]\hat{E}_t^{(T)}[S_T], \\ &= D_{t,T}\hat{E}_t^{(T)}[S_T] \end{aligned} \tag{7.3}$$

If the risk-free interest rate is stochastic, then $\hat{E}_t^{(T)}[.]$ is the conditional expectation under the T-forward measure $\hat{\mathbb{P}}^{(T)}$. When the risk-free rate is non-stochastic, then the forward measure reduces to the risk-neutral measure $\hat{\mathbb{P}}$ and will not depend upon maturity time, i.e., $E_t^{(T)}[\cdot] \rightarrow \hat{E}_t[\cdot]$.[8]

Note that Eqs. 7.1 and 7.3 can be applied to both the stock and the option prices. This leads to the following theorem which is the main result of this paper.

Theorem 7.1 The following formula provides a lower bound for the European call option $\underline{C}_t$:

$$\underline{C}_t = D_{t,T}E_t[C_T] + \beta_C\left\{S_t - D_{t,T}E_t[S_T]\right\} \tag{7.4}$$

where $\beta_C = \frac{\text{cov}[C_T, S_T]}{\text{var}[S_T]}$.

Proof By Eq. 7.1, the option price must follow $C_t = E_t[M_{t,T}\, C_T]$, and hence

$$\begin{aligned} C_t &= E_t\left[M_{t,T}C_T\right] \\ &= E_t\left[M_{t,T}\right]E_t[C_T] + \text{cov}\left[M_{t,T}, C_T\right] \\ &= D_{t,T}E_t[C_T] + \text{cov}\left[M_{t,T}, C_T\right] \end{aligned} \tag{7.5}$$

or

$$C_t - D_{t,T}E_t[C_T] = \text{cov}\left[M_{t,T}, C_T\right]. \tag{7.6}$$

Similarly,

$$S_t - D_{t,T}E_t[S_T] = \text{cov}\left[M_{t,T}, S_T\right]. \tag{7.7}$$

[8] Without loss of generality and for the ease of exposition, we take non-stochastic interest rates and proceed with the risk-neutral measure $\hat{\mathbb{P}}$ for the rest of the paper.

Hence, to prove Eq. 7.4, we only need to prove

$$\operatorname{cov}\left[M_{t,T}, C_T\right] \geq \beta_C \operatorname{cov}\left[M_{t,T}, S_T\right]. \tag{7.8}$$

Note that cov$[M_{t,T}, C_T]$ = cov$[M_{t,T}, \max\{S_T - K, 0\}]$ is monotonic in strike price K and has a minimum value when $K = 0$ in which case cov$[M_{t,T}, C_T]$ = cov$[M_{t,T}, S_T]$ and a maximum value when $K \to \infty$ in which case cov$[M_{t,T}, C_T] = 0$. Hence, cov$[M_{t,T}, S_T] \leq$ cov$[M_{t,T}, C_T]$ which is less than 0. Note that $0 < \beta_C < 1$ (see the proof in the following Corollary). In Appendix 1, it is proved that cov$[M_{t,T}, C_T] \geq \beta_C$ cov$[M_{t,T}, S_T]$.

The put lower bound takes the same form and is provided in the following corollary:

Corollary 7.1 The lower bound of the European put option $\underline{P}_t$ can be obtained by the put-call parity and satisfies the same functional form in Theorem 7.1:

$$\underline{P}_t = D_{t,T} E_t[P_T] + \beta_P \{S_t - D_{t,T} E_t[S_T]\} \tag{7.9}$$

where $\beta_P = \dfrac{\operatorname{cov}[P_T, S_T]}{\operatorname{var}[S_T]}$.

[Proof] By the Put-Call Parity:

$$P_t = C_t + D_{t,T} K - S_t \geq \underline{C}_t + D_{t,T} K - S_t = \underline{P}_t. \tag{7.10}$$

We then substitute in the result of Theorem 7.1 to get

$$\begin{aligned}
\underline{P}_t &= D_{t,T} E_t[C_T] + \beta_C \{S_t - D_{t,T} E_t[S_T]\} + D_{t,T} K - S_t \\
&= D_{t,T} E_t[S_T + P_T - K] + \beta_C \{S_t - D_{t,T} E_t[S_T]\} + D_{t,T} K - S_t \\
&= D_{t,T} E_t[P_T] + \beta_P \{S_t - D_{t,T} E_t[S_T]\}
\end{aligned} \tag{7.11}$$

where $\beta_P = \beta_C - 1$. Note that $\beta_P < 0 < \beta_C$. This also implies that $\beta_c < 1$. Finally, it is straightforward to show that $\beta_P = \frac{\operatorname{cov}[P_T, S_T]}{\operatorname{var}[S_T]}$ using the put-call parity.

Therefore, for both call and put options, since the relationship between the pricing kernel M and stock price S is convex, the theorem provides a lower bound.[9] Merton (1973) shows that the stock price will be the upper bound for the call option and the strike price should be the upper bound for the put option; otherwise, arbitrage should occur. Ritchken (1985) provides a tighter upper bound[10] than that of Merton, which is stated in the following theorem, although we provide an alternative proof to Ritchken that is consistent with our derivation of the lower bound.

[9] In the Appendix, $\varepsilon > 0$.

[10] Perrakis and Ryan (1984) and Ritchken (1985) obtain the identical upper bound.

Theorem 7.2 The following formulas provide upper bounds for the European call and put options (Ritchken (1985))[11]:

$$\begin{aligned}\overline{C}_t &= \frac{S_t}{E_t[S_T]}E_t[C_T]\\ \overline{P}_t &= \frac{S_t}{E_t[S_T]}E_t[P_T] + \left(D_{t,T} - \frac{S_t}{E_t[S_T]}\right)K.\end{aligned} \tag{7.12}$$

Proof Similar to the proof of the lower bound, the upper bound of the call option is provided as follows:

$$\begin{aligned}\frac{S_t}{E_t[S_T]}E_t[C_T] &= \frac{E_t\left[M_{t,T}S_T\right]}{E_t[S_T]}E_t[C_T]\\ &= \frac{E_t\left[M_{t,T}S_TC_T\right] - \mathrm{cov}\left[M_{t,T}S_T, C_T\right]}{E_t[S_T]}\\ &> \frac{E_t\left[M_{t,T}S_TC_T\right]}{E_t[S_T]}\\ &= C_t\frac{E_t^{(C)}[S_T]}{E_t[S_T]}\\ &> C_t.\end{aligned} \tag{7.13}$$

The third line of the above equation is a result from the fact that $\mathrm{cov}[M_{t,T}S_T, C_T] < 0$. The fourth line of the above equation is a change of measure with the call option being the numeraire. The last line of the above equation is a result based upon $E_t^{(C)}[S_T]/E_t[S_T] > 1$.[12]

By the put-call parity, we can show that the upper bound of the put option requires an additional term:

$$\begin{aligned}\overline{P}_t &= \overline{C}_t + D_{t,T}K - S_t = \frac{S_t}{E_t[S_T]}E_t[P_T + S_T - K] + D_{t,T}K - S_t\\ &= \frac{S_t}{E_t[S_T]}E_t[P_T] + \left(D_{t,T} - \frac{S_t}{E_t[S_T]}\right)K.\end{aligned} \tag{7.14}$$

The lower and upper bounds we show in this paper have two important advantages over the existing bounds. The bounds will converge to the true value of the option if:

- The expected stock return, $\frac{E_t[S_T]}{S_t}$, approaches the risk-free rate.
- The correlation between the stock and the call or put option (ρ_{SC} or ρ_{SP}) approaches 1 or -1.

[11]This is same as Proposition 3-i (Eq. 7.26) in Ritchken (1985).

[12]By the definition of measure change, we have $E_t[C_TS_T] = E_t[C_T]E_t^{(C)}[S_T]$ which implies $E_t^{(C)}[S_T]/E[S_T] = E_t[C_TS_T]/\{E_t[C_T]E_t[S_T]\} > 1$.

These advantages help us identify when the bounds are tight and when they are not. The first advantage indicates that the bounds are tight for low-risk stocks and not tight for high-risk stocks. The second advantage indicates that the bounds are tighter for in-the-money options than out-of-the-money options.

7.3 Comparisons

The main purpose of this section is to compare our lower bound to several lower bounds in previous studies, namely, Merton (1973), Perrakis and Ryan (1984), Ritchken (1985), Ritchken and Kuo (1988), Gotoh and Konno (2002), and Rodriguez (2003) using the Black-Scholes model as the benchmark for its true option value. We also compare Ritchken's upper bound (which is also our upper bound) with more recent works by Gotoh and Konno (2002) and Rodriguez (2003).

The Black-Scholes model has five variables: stock price, strike price, volatility (standard deviation), risk-free rate (constant), and time to maturity. In addition to the five variables, the lower bound models need the physical expected stock return. The following is the base case for the comparison:

Current stock S_0	50
Strike K	50
Volatility σ	0.2
Risk-free rate r	0.1
Time to maturity T	1
Stock expected return μ	0.2

In the Black-Scholes model, stock price (S) evolution follows a log normal process:

$$dS = \mu S dt + \sigma S dW \tag{7.15}$$

where instantaneous expected rate of stock return μ and volatility of stock price σ are assumed to be constants and where dW is a wiener process. The call option price is computed as

$$C = S_0 N(d_+) - e^{-rT} K N(d_-) \tag{7.16}$$

where

$$d_{\pm} = \frac{\ln S_0 - \ln K + (r \pm 0.5\sigma^2)T}{\sigma\sqrt{T}}.$$

The lower bound is computed by simulating the stock price using Eq. 7.15 via a binomial distribution approximation:

$$S_j = S_0 u^j d^{n-j}. \tag{7.17}$$

As n approaches infinity, S_j approaches the log normal distribution, and the binomial model converges to the Black-Scholes model. Under the risk-neutral measure, the probability associated with the jth state is set as

$$\widehat{\Pr}[j] = \binom{n}{j} \hat{p}^j (1 - \hat{p})^{n-j} \tag{7.18}$$

where

$$\hat{p} = \frac{e^{r\Delta t} - d}{u - d},$$

$$u = e^{\sigma\sqrt{\Delta t}},$$

$$d = e^{-\sigma\sqrt{\Delta t}},$$

and $\Delta t = T/n$ represents the length of the partition. Under the actual measure, the formula will change to

$$\Pr[j] = \binom{n}{j} p^j (1 - p)^{n-j} \tag{7.19}$$

where

$$p = \frac{e^{\mu\Delta t} - d}{u - d}.$$

Finally, the pricing kernel in our model is set as

$$M_{0T}[j] = \frac{\widehat{\Pr}[j]}{\Pr[j]} e^{-rT}. \tag{7.20}$$

In our results, we let n be great enough so that the binomial model price is 4-digit accurate to the Black-Scholes model. We hence set n to be 1,000. The results are reported in Table 7.1. The first panel presents the results for various moneyness levels, the second panel presents the results for various volatility levels, the third panel presents the results for various interest rates, the fourth panel presents the results for various maturity times, and the last presents the results for various stock expected returns. In general, the lower bounds are tighter (for all models) when the moneyness is high (in-the-money), volatility is high, risk-free rate is high, time to maturity is short, and the expected return of stock is low.

This table presents comparisons between our lower bound and existing lower bounds by Merton (1973), Perrakis and Ryan (1984), and Ritchken (1985). The base case parameter values are:

- Stock price = 50
- Strike price = 50
- Volatility = 0.2
- Risk-free rate = 10 %
- Time to maturity = 1 year
- Stock expected return (μ) = 20 %

The highlighted rows represent the base case.

As we can easily see, universally our model for the lower bound is tighter than any of the comparative models. One result particularly worth mentioning is that our lower bound performs better than the other lower bound models in out-of-the-money options. For example, our lower bound is much better than Ritchken's (1985) lower bound when the option is 20 % out-of-the-money and continues to show value when Ritchken's lower bound returns to 0 (see the first panel of Table 7.1).

While Ritchken and Kuo (1988) claim to obtain tighter lower bounds than Perrakis (1986) and Perrakis and Ryan (1984), they do not show direct comparisons in their paper. Rather, they present through a convergence plot (Figure 3 on page 308 in Ritchken and Kuo (1988)) of a Black-Scholes example with the true value being \$5.4532 and the lower bound approaching roughly \$5.2. The same parameter values with our lower bound show a lower bound of \$5.4427, which demonstrates a substantial improvement over the Ritchken and Kuo model.

The comparisons with more recent studies of Gotoh and Konno (2002) and Rodriguez (2003) are given in Tables 7.2 and 7.3.[13] Gotoh and Konno use semi-definite programming and a cutting plane algorithm to study upper and lower bounds of European call option prices. Rodriguez uses a nonparametric method to derive lower and upper bounds. As we can see in Tables 7.2 and 7.3, except for very few upper bound cases, none of the bounds under the Gotoh and Konno's model and Rodriguez's model are very tight, compared to our model. Furthermore, note that our model requires no moments of the underlying distribution.[14]

The base case parameter values are:

- Stock price = 40
- Risk-free rate = 6 %
- Stock expected return (μ) = 20 %

[13] We also compare with the upper bound by Zhang (1994), which is an improved upper bound by Lo (1987), and show overwhelming dominance of our upper bound. The results (comparison to Tables 7.1, 7.2, and 7.3 in Zhang) are available upon request.

[14] The upper bounds by the Gotoh and Konno model perform well in only in-the-money, short maturity, and low volatility scenarios, and these scenarios are where the option prices are close to their intrinsic values, and hence the percentage errors are small.

Table 7.1 Lower bound comparison

S	Blk-Sch	Merton	PR	Ritch.	Our	$error	%error
20	0.0000	0	0	0	0	0.0000	
25	0.0028	0	0	0	0	0.0028	
30	0.0533	0	0	0	0	0.0533	
35	0.3725	0	0	0	0.0008	0.3717	99.78
40	1.3932	0	0	0.272	0.8224	0.5708	40.97
45	3.4746	0	1.441	2.278	2.8957	0.5789	16.66
50	6.6322	4.758	5.449	5.791	6.1924	0.4398	6.63
55	10.6248	9.758	10.017	10.143	10.3535	0.2714	2.55
60	15.1288	14.758	14.849	14.888	14.9851	0.1437	0.95
65	19.9075	19.758	19.788	19.797	19.8395	0.0680	0.34
70	24.8156	24.758	24.768	24.767	24.7860	0.0296	0.12
75	29.7794	29.758	29.761	29.76	29.7673	0.0121	0.04
80	34.7658	34.758	34.759	34.754	34.7611	0.0047	0.01

Vol	Blk-Sch	Merton	PR	Ritch.	Our	$error	%error
0.1	5.1526	4.7580	4.7950	4.8630	4.8930	0.2596	5.04
0.15	5.8325	4.7580	5.0290	5.2400	5.4367	0.3958	6.79
0.2	6.6322	4.7580	5.4490	5.7890	6.1924	0.4398	6.63
0.25	7.4847	4.7580	5.9700	6.4360	7.0247	0.4600	6.15
0.3	8.3633	4.7580	6.5370	7.1010	7.8864	0.4769	5.70
0.35	9.2555	4.7580	7.1180	7.7760	8.7601	0.4954	5.35
0.4	10.1544	4.7580	7.6990	8.4570	9.6382	0.5162	5.08
0.45	11.0559	4.7580	8.2680	9.1250	10.5170	0.5389	4.87
0.5	11.9574	4.7580	8.8220	9.7780	11.3945	0.5629	4.71
0.55	12.8569	4.7580	9.3570	10.4240	12.2692	0.5877	4.57

Rate	Blk-Sch	Merton	PR	Ritch.	Our	$error	%error
0.02	4.4555	0.9901	1.7380	2.7680	3.0243	1.4313	32.12
0.04	4.9600	1.9605	2.6940	3.5010	3.8402	1.1198	22.58
0.06	5.4923	2.9118	3.6310	4.2490	4.6400	0.8523	15.52
0.08	6.0504	3.8442	4.5490	5.0200	5.4240	0.6264	10.35
0.1	6.6322	4.7581	5.4490	5.7970	6.1924	0.4398	6.63
0.12	7.2355	5.6540	6.3310	6.5810	6.9456	0.2900	4.01
0.14	7.8578	6.5321	7.1960	7.3670	7.6839	0.1739	2.21
0.16	8.4965	7.3928	8.0430	8.1470	8.4076	0.0889	1.05
0.18	9.1488	8.2365	8.8740	8.9350	9.1169	0.0319	0.35
0.2	9.8122	9.0635	9.6880	9.7170	9.8122	0.0000	0.00
0.22	10.4841	9.8741	10.4870	10.4830	10.4841	0.0000	0.00

T	Blk-Sch	Merton	PR	Ritch.	Our	$error	%error
0.1	1.5187	0.4975	1.2850	1.3600	1.4976	0.0211	1.39
0.2	2.3037	0.9901	1.8870	2.0240	2.2469	0.0568	2.47
0.3	2.9693	1.4777	2.4000	2.5940	2.8696	0.0997	3.36
0.4	3.5731	1.9605	2.8730	3.0990	3.4266	0.1465	4.10
0.5	4.1371	2.4385	3.3250	3.5830	3.9416	0.1955	4.72
0.6	4.6726	2.9118	3.7640	4.0510	4.4272	0.2454	5.25
0.7	5.1862	3.3803	4.1940	4.5020	4.8909	0.2953	5.69

(*continued*)

Table 7.1 (continued)

0.8	5.6822	3.8442	4.6170	4.9460	5.3375	0.3447	6.07
0.9	6.1635	4.3034	5.0350	5.3710	5.7705	0.3930	6.38
1	6.6322	4.7581	5.4490	5.7890	6.1924	0.4398	6.63

Mu (μ)	Blk-Sch	Merton	PR	Ritch.	Our	$error	%error
0.1	6.6322	4.7580	6.3740	6.4310	6.6322	0.0000	0.00
0.15	6.6322	4.7580	5.8380	6.0610	6.4703	0.1620	2.44
0.2	6.6322	4.7580	5.4490	5.7890	6.1924	0.4398	6.63
0.25	6.6322	4.7580	5.1800	5.5820	5.8689	0.7633	11.51
0.3	6.6322	4.7580	5.0040	5.4360	5.5572	1.0750	16.21
0.35	6.6322	4.7580	4.8940	5.3090	5.2936	1.3386	20.18
0.4	6.6322	4.7580	4.8300	5.2130	5.0929	1.5393	23.21
0.45	6.6322	4.7580	4.7940	5.1330	4.9536	1.6786	25.31

Note: *S* is the stock price; *vol* is the volatility; *rate* is the risk-free rate; *Mu* (μ) is the expected rate of return of stock; *Blk-Sch* is the Black-Scholes (1973) solution; *Merton* is the Merton (1973) model; *PR* is Perrakis and Ryan (1984) model; *Ritch.* is the Ritchken (1985) model; *Our* is our model; *$error* is error in dollar; *%error* is error in percentage

The base case parameter values are:

- Strike price = 50
- Volatility = 0.2
- Risk-free rate = 10 %
- Time to maturity = 1 year
- Stock expected return (μ) = 20 %

7.4 Extensions

In addition to a tight lower bound, another major contribution of our model is that it makes no assumption on the distribution of the underlying stock (unlike Lo (1984) and Gotoh and Konno (2002) who require moments of the underlying distribution) or any assumption on interest rates and volatility (unlike Rodriguez (2003) who requires constant interest rates). As a result, our lower bound can be used with models that assume random volatility and random interest rates or any arbitrary specification of the underlying stock. Note that our model needs only the dollar beta of the option and expected payoffs of the stock and the option.[15] In this section, we extend our numerical experiment to a model with random volatility and random interest rates.

Option models with random volatility and random interest rates can be derived with closed form solutions under the Scott (1997) and Bakshi et al. (1997)

[15]The term "dollar beta" is originally from Page 173 of Black (1976). Here we mean β_c and β_p.

Table 7.2 Comparison of upper and lower bounds with the Gotoh and Konno (2002) model

	Lower bound			Upper bound	
Stk	Our	GK	Blk-Sch	Our	GK
$S = 40$; rate $= 6$ %; vol $= 0.2$; $t = 1$ week					
30	10.0346	10.0346	10.0346	10.1152	10.0349
35	5.0404	5.0404	5.0404	5.1344	5.0428
40	0.4628	0.3425	0.4658	0.5225	0.5771
45	0.0000	0.0000	0.0000	0.0000	0.0027
50	0.0000	0.0000	0.0000	0.0000	0.0003
$S = 40$; rate $= 6$ %; vol $= 0.8$; $t = 1$ week					
30	10.0400	10.0346	10.0401	10.1202	10.1028
35	5.2644	5.0404	5.2663	5.3483	5.4127
40	1.7876	1.2810	1.7916	1.8428	2.2268
45	0.3533	0.0015	0.3548	0.3717	0.5566
50	0.0412	0.0000	0.0419	0.0444	0.1021
$S = 40$; rate $= 6$ %; vol $= 0.8$; $t = 12$ week					
30	11.9661	10.4125	12.0278	12.7229	12.8578
35	8.7345	6.2980	8.8246	9.4774	9.7658
40	6.2141	3.8290	6.3321	6.8984	7.5165
45	4.3432	2.5271	4.4689	4.9421	6.8726
50	2.9948	1.5722	3.1168	3.4990	4.5786

Note: S is the stock price; *Stk* is the strike price; *vol* is the volatility; *rate* is the risk-free rate; *Blk-Sch* is the Black-Scholes (1973) solution; *GK* is the Gotoh and Konno (2002) model; *Our* is our model; *\$error* is error in dollar; *%error* is error in percentage

specifications. However, here, given that there is no closed form solution to the covariance our model requires, we shall use Monte Carlo to simulate the lower bound. In order to be consistent with the lower bound, we must use the same Monte Carlo paths for the valuation of the option. For the ease of exposition and simplicity, we assume the following joint stochastic processes of stock price S, interest rate r, and volatility V under the actual measure, respectively:

$$\begin{aligned} dS &= \mu S dt + \sqrt{V} S d\hat{W}_2 \\ dr &= \alpha(\theta - r)dt + v d\hat{W}_2 \\ dV &= \eta V d\hat{W}_2 \end{aligned} \tag{7.21}$$

where dW is a wiener process, $dW_i dW_j = 0$, μ, α, θ, υ, and η are constants. The processes under the actual measure are used for simulating the lower and upper bounds. The Monte Carlo simulations are performed under the risk-neutral measure in order to compute the option price:

$$\begin{aligned} dS &= rS dt + \sqrt{V} S d\hat{W}_1 \\ dr &= \alpha(\theta - r)dt + v d\hat{W}_2 \\ dV &= \eta V d\hat{W}_3. \end{aligned} \tag{7.22}$$

Table 7.3 Comparison of upper and lower bounds with the Rodriguez (2003) model

	Lower bound			Upper bound	
S	Our	Rodriguez	Blk-Sch	Our	Rodriguez
30	0.0221	0	0.0538	0.1001	0.1806
32	0.0725	0.0000	0.1284	0.2244	0.3793
34	0.1828	0.0171	0.2692	0.4451	0.7090
36	0.3878	0.1158	0.5072	0.7973	1.2044
38	0.7224	0.3598	0.8735	1.3100	1.8900
40	1.2177	0.7965	1.3950	2.0044	2.7767
42	1.8982	1.4521	2.0902	2.8927	3.8619
44	2.7711	2.3329	2.9676	3.9697	5.1315
46	3.8319	3.4286	4.0255	5.2211	6.5640
48	5.0709	4.7177	5.2535	6.6302	8.1337
50	6.4703	6.1724	6.6348	8.1753	9.8149
52	8.0072	7.7635	8.1494	9.8327	11.5835
54	9.6574	9.4631	9.7758	11.5794	13.4187
56	11.3974	11.2462	11.4933	13.3950	15.3032
58	13.2067	13.0922	13.2832	15.2621	17.2235
60	15.0683	14.9845	15.1292	17.1674	19.1693
62	16.9716	16.9101	17.0179	19.1024	21.1328
64	18.9035	18.8593	18.9384	21.0573	23.1086
66	20.8559	20.8250	20.8822	23.0262	25.0926
68	22.8239	22.8020	22.8429	25.0056	27.0822
70	24.8018	24.7867	24.8157	26.9915	29.0754

Note: *S* is the stock price; *Blk-Sch* is the Black-Scholes (1973) solution; *Rodriguez* is the Rodriguez (2003) model; *Our* is our model; *\$error* is error in dollar; *%error* is error in percentage

To simplify the problem without loss of generosity, we assume that investors charge no risk premiums on interest rate risk and volatility risk, i.e., $dW_2 = d\hat{W}_2$ and $dW_3 = d\hat{W}_3$.

The simulations are done by the following integrals under the actual (top equation) and risk-neutral (bottom equation) measures:

$$
\begin{aligned}
S_t &= S_0 \exp\left(\int_0^t \mu_u du - \int_0^t ½ V_u du + \int_0^t \sqrt{V_u} dW_u\right) \\
\hat{S}_t &= S_0 \exp\left(\int_0^t r_u du - \int_0^t ½ V_u du + \int_0^t \sqrt{V_u} d\hat{W}_u\right).
\end{aligned} \tag{7.23}
$$

The no-arbitrage price of the call option is computed under the risk-neutral measure as

$$
C = E\left[\exp\left(-\int_0^t r_u du\right) \max\{\hat{S}_t - K, 0\}\right]. \tag{7.24}
$$

Table 7.4 Lower bound under the random volatility and random interest rate model

S	BCC/Scott	Our	\$error	%error
25	1.9073	1.1360	0.7713	40.44
30	3.2384	2.4131	0.8253	25.49
35	5.2058	4.1891	1.0167	19.53
40	7.5902	6.5864	1.0039	13.23
45	10.2962	9.5332	0.7630	7.41
50	13.6180	12.8670	0.7510	5.52
55	17.1579	16.5908	0.5671	3.30
60	21.1082	20.5539	0.5543	2.63
65	25.0292	24.7251	0.3040	1.21
70	29.3226	29.0742	0.2484	0.85

Note: *S* is the stock price; *BCC/Scott* are Bakshi et al. (1997) and Scott (1997) models; *Our* is our model; *\$error* is error in dollar; *%error* is error in percentage

The bounds are computed under the actual measure. For example,

$$E[C_t] = E[\max\{S_t - K, 0\}]. \tag{7.25}$$

Given that $dW_i dW_j = 0$, we can simulate the interest rates and volatility separately and then simulate the stock price. That is, conditional on known interest rates and volatility, under independence, the stock price is log normally distributed.

We perform our simulations using 10,000 paths over 52 weekly periods. The parameters are given as follows:

Strike K	50
Time to maturity T	1
Stock expected return μ	0.2
Reverting speed α	0.5
Reverting level θ	0.1
Interest rate volatility υ	0.03
Initial interest rate r_0	0.1
Initial variance V_0	0.04[16]
Volatility on variance η	0.2

Note that implicitly we assume the price of risk for both interest rate process and volatility process to be 0, for simplicity and without loss of generality. The results are shown in Table 7.4. Compared to the model of the Black-Scholes (i.e., the first panel of Table 7.4), the lower bound performs similarly in the random volatility and random interest rate model. Take the base case as an example where the Black-Scholes price is 6.6322, the Bakshi-Cao-Chen/Scott price is 13.6180 as a result of extra uncertainty in the stock price due to random volatility and interest

[16]This is so because the initial volatility is 0.2.

rates. The error of the lower bound of our model is 0.7510 in the Bakshi-Cao-Chen/Scott case as opposed to 0.4398 in the Black-Scholes case.[17] The percentage error is 5.52 % in the Bakshi-Cao-Chen/Scott case versus 6.63 % in the Black-Scholes case. The in-the-money options have larger percentage errors than those of the out-of-the-money options.

The parameters are given as follows:

Strike price	50
Time to maturity	1
Stock expected return	0.2
Reverting speed	0.5
Reverting level	0.1
Interest rate volatility	0.03
Price of risk	0
Initial interest rate	0.1
Initial volatility	0.2
Volatility on volatility	0.2

The Monte Carlo paths are 10,000. The stock price, volatility, and interest rate processes are assumed to be independent

7.5 Empirical Study

In this section, we test the lower and upper bounds against data. Charles Cao has generously provided us with the approximated prices of S&P 500 index call option contracts, matched levels of S&P 500 index, and approximated risk-free 90-day T-Bill rates for the period of June 2, 1988 through December 31, 1991.[18] For each day, the approximated option prices are calculated as the average of the last bid and ask quotes. Index returns are computed using daily closing levels for the S&P 500 index that are collected and confirmed using data obtained from Standard and Poor's, CBOE, Yahoo, and Bloomberg.[19]

The dataset contains 46,540 observations over 901 days (from June 2, 1988, to December 31, 1991). Hence, on average there are over 50 options for various maturities and strikes. The shortest maturity of the dataset is 7 days and the longest is 367 days. 15 % of the data are less than 30 days to maturity, 32 % are between

[17]This Black-Scholes case is from the highlighted row in the first panel of Table 7.1.

[18]The data are used in Bakshi et al. (1997).

[19]The (ex-dividend) S&P 500 index we use is the index that serves as an underlying asset for the option. For option evaluation, realized returns of this index need not be adjusted for dividends unless the timing of the evaluated option contract is correlated with lumpy dividends. Because we use monthly observations, we think that such correlation is not a problem. Furthermore, in any case, this should not affect the comparison of the volatility smile between our model and the Black-Scholes model.

30 and 60 days, 30 % are between 60 and 180 days, and 24 % are more than 180 days to maturity. Hence, these data do not have maturity bias.

The deepest out-of-the-money option is -18.33 %, and the deepest in-the-money option is 47.30 %. Roughly half of the data are at-the-money options (46 % of the data are within 5 % in-the-money and out-of-the-money). 10 % are deep-in-the-money (more than 15 % in-the-money), but less than 1 % of the data are deep-out-of-the-money (more than 15 % out-of-the-money). Hence, the data have disproportional fraction of in-the-money options. This is clearly a reflection of the bull market in the sample period.

The best way to test the lower and upper bounds derived in this paper is to use a nonparametric, distribution-free model. Note that the lower bound in Theorem 7.1 requires only the expected return of the underlying stock and the covariance between the stock and the option. There is no further requirement for the lower and the upper bounds. In other words, our lower bound model can permit any arbitrary distribution of the underlying stock and any parametric specification of the underlying stock such as random volatility, random interest rates, and jumps. Hence, to best test the bounds with a parsimonious empirical design, we adopt the histogram method introduced by Chen et al. (2006) where the underlying asset is modelled by past realizations, i.e., histogram.

We construct histograms from realizations of S&P 500 index (SPX) returns. We calculate the price on day t of an option that settles on day T using a histogram of S&P 500 index returns for a holding period of T–t, taken from a 5-year window immediately preceding time t.[20] For example, an x-calendar-day option price on any date is evaluated using a histogram of round$\left[\frac{252}{365}x\right]$ -trading-day holding period returns where round [.] is rounding the nearest integer.[21] The index levels used to calculate these returns are taken from a window that starts on the 1260th ($\approx 5 \times 252$) trading day before the option trading date and ends 1 day before the trading date. Thus, this histogram contains $1,260 - \text{round}\left[\frac{252}{365}x\right]$ -trading-day return realizations. Formally, we compute histogram of the (unannualized) returns by the following equation:

$$R_{t,\,t+x,\,i} = \ln S_{t-i-x} - \ln S_{t-i}, \tag{7.26}$$

where each i is an observation in time and t is the last i. For example, if t is 1988/06/02 and x is 15 calendar days (or ten business days). We further choose our histogram horizon to be 5 years or 1,260 business days. Fifteen business days after 1988/06/02 is 1988/06/17. To estimate a distribution of the stock return for 1988/06/17, we look back a series of 10-business-day returns. Since we choose a 5-year historical window, or 1,260-business-day window, the histogram will contain

[20]We use three alternative time windows, 2-year, 10-year, and 30-year, to check the robustness of our procedure and results.

[21]The conversion is needed because we use trading-day intervals to identify the appropriate return histograms and calendar-day intervals to calculate the appropriate discount factor.

1,260 observations. The first return in the histogram, $R_{1933/06/02,1933/06/07,1}$, is the difference between the log of the stock price on 1988/06/01, ln S_{t-1}, and the log of the stock price 15 calendar days (ten business days) earlier on 1988/05/17, ln $S_{t-1-\alpha}$. The second observation in the histogram, $R_{1933/06/02/1933/06/07,2}$ is computed as ln S_{t-2}–ln $S_{t-2-\alpha}$.

After, we complete the histogram of returns, we then convert it to the histogram of prices by multiplying every observation in the histogram by the current stock price:

$$S_{t+x,\,i} = S_t R_{t,\,t+x,\,i}. \tag{7.27}$$

The expected option payoff is calculated as the average payoff where all the realizations in the histogram are given equal weights. Thus, $E_t[C_{T,T,K}]$ and $E_t[S_T]$ are calculated as

$$\begin{cases} E_t\left[C_{T,T,K}\right] = \dfrac{1}{N}\sum_{i=1}^{N} \max\left\{S_{T,\,i} - K, 0\right\} \\ E_t[S_T] = \dfrac{1}{N}\sum_{i=1}^{N} S_{T,\,i} \end{cases}, \tag{7.28}$$

where N is the total number of realized returns and $C_{t,T,K}$ is the price observed at time t, of an option that expires at time T with strike price K. Substituting the results in Eq. 7.28 in the approximation pricing formula of Eq. 7.4, we obtain our empirical model:

$$\begin{aligned} \underline{C}_t &= P_{t,T}E_t\left[C_{T,T,K}\right] + \beta_C\left\{S_t - P_{t,T}E_t[S_T]\right\} \\ &= P_{t,T}\frac{1}{N}\sum_{i=1}^{N} \max\left\{S_{T,\,i} - K, 0\right\} + \beta_C\left\{S_t - P_{t,T}\frac{1}{N}\sum_{i=1}^{N} S_{T,\,i}\right\} \end{aligned} \tag{7.29}$$

where the dollar beta is defined as $\beta_C = \frac{\mathrm{cov}[C_T, S_T]}{\mathrm{var}[S_T]}$ as defined in Eq. 7.4.

Note that option prices should be based upon projected future volatility levels rather than historical estimates. We assume that investors believe that the distribution of index returns over the time to maturity follows the histogram of a particular horizon with a projected volatility. In practice, traders obtain this projected volatility by calibrating the model to the market price. We incorporate the projected volatility, $v^*_{\mathrm{t,T,K}}$, into the histogram by adjusting its returns:

$$R^*_{t,T,K,\,i} = \frac{v^*_{t,T,K}}{v_{t,T}}\left(R_{t,T,\,i} - \overline{R}_{t,T}\right) + \overline{R}_{t,T}.., \quad i = 1, \cdots, N, \tag{7.30}$$

where the historical volatility $v_{t,T}$ is calculated as the standard deviation of the historical returns as follows:

$$v^2_{t,T} = \frac{1}{N-1}\sum_{i=1}^{N}\left(R_{t,T,\,i} - \overline{R}_{t,T}\right)^2 \tag{7.31}$$

where $R_{t,T,i} = S_{T,i}/S_t$ and $\overline{R}_{t,T} = \frac{1}{N}\Sigma_{i=1}^{N} R_{t,T,\,i}$ is the mean return.

Note that the transformation from R to R^* changes the standard deviation from $v_{t,T}$ to $\nu^*_{t,T,K}$, but does not change the mean, skewness, or kurtosis. In our empirical study, we approximate the true volatility by the Black-Scholes implied volatility. For the upper bound calculations, we also need an expected mean return of the stock. In our empirical study, we simply use the histogram mean for it.

The selection of the time horizon is somewhat arbitrary. Empiricists know well that too long horizons reduce the impact of recent events and yet too short understate the impact of long time effects. Given that there is no consensus on a most proper horizon, we perform our test over a variety of choices, namely, 5-year, 10-year, and 30-year, and the results are similar. To conserve space, we provide the results on the 10-year and leave the others available on request.

The results are shown in Table 7.5. Columns (1) and (2) define the maturity and moneyness buckets. Short maturity is less than 30 days to maturity, medium is between 31 and 90 days, long is between 91 and 180 days, and real long is over 180 days to maturity. At-the-money is between 5 % in-the-money and 5 % out-of-the-money (or −5 %), near-in-the-money/near-out-of-the-money is between 5 % and 15 %, and deep-in-the-money/deep-out-of-the-money is over 15 %. Moneyness is defined as $S/K-1$. Column (3) is a frequency count of the number of observations in each bucket. Column (4) represents the average value of the ratios of the lower bound over the market price of the option. Column (5) shows the number of violations when the lower bound is higher than the market price.

Out of the entire sample (46,540 observations), on average, the lower bound is 9.57 % below the market value and the upper bound is 9.28 % above the market value. When we look into subsamples, the performances vary. In general, the lower bound performs better in-the-money than out-of-the-money and medium maturity than other maturities. The best lower bound performance is when the option is near-in-the-money and short-term maturity (2.83 % below market value).

To visualize the upper and lower bounds, we plot selected contracts in Fig. 7.1. In Fig. 7.1, we plot at-the-money (ATM) options from four maturities, 1 month (Fig. 7.1a), 2 months (Fig. 7.1b), 3 months (Fig. 7.1c), and 6 months (Fig. 7.1d). As we move from short maturity to long maturity, the number of observations drops (51, 38, 21, 15 observations, respectively). The bounds are wider as we move from short maturity to long maturity, consistent with the analysis in the previous sections.

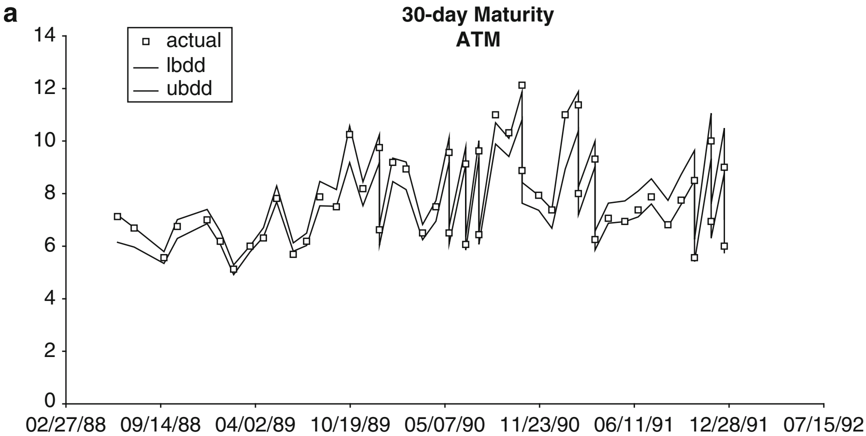
As we can see, in general, the lower bound using histograms is best for in-the-money and short-dated options and worst for out-of-the-money options. On average the lower bound is 9.57 % below the market value (the ratio is 90.43 %). However, there are violations. For example, medium-term, near-out-of-the-money options have five violations of the lower bound, and there are a total of 4,233 violations. Theoretically, there should be arbitrage opportunities when the bounds are violated. To test the lower bound does imply arbitrage opportunities, we

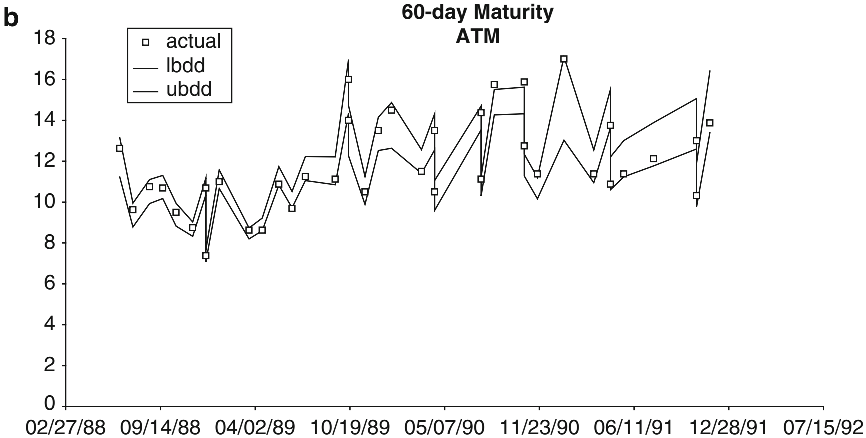
Table 7.5 Empirical results on the lower bound

(1)	(2)	(3)	(4)	(5)
Maturity	Money	Total	% lbdd	Violation
Short	Deep out	0		0
Medium	Deep out	0		0
Long	Deep out	84	57.26	0
Real long	Deep out	295	61.18	0
Short	Near out	145	65.33	0
Medium	Near out	1,925	69.27	5
Long	Near out	3,554	82.29	30
Real long	Near out	3,317	86.88	271
Short	At	4,106	89.73	492
Medium	At	7,732	92.45	962
Long	At	5,650	92.82	380
Real long	At	3,857	92.61	584
Short	Near in	2,220	97.17	218
Medium	Near in	3,924	97.08	785
Long	Near in	2,951	94.50	234
Real long	Near in	2,018	90.31	130
Short	Deep in	660	92.48	0
Medium	Deep in	1,296	94.52	29
Long	Deep in	1,352	93.14	46
Real long	Deep in	1,454	89.42	67
Total		46,540	90.43	4,233

Note:
1. This is based upon 2,520 business-day (10 years) horizon. Results of other horizons are similar and are available on request
2. Columns (1) and (2) define the maturity and moneyness buckets. Short maturity is less than 30 days to maturity, medium is between 31 and 90 days, long is between 91 and 180 days, and real long is over 180 days to maturity. At-the-money is between 5 % in-the-money and 5 % out-of-the-money (or −5 %), near-in-the-money/near-out-of-the-money is between 5 % and 15 %, and deep-in-the-money/deep-out-of-the-money is over 15 %. Moneyness is defined as $S/K - 1$. Column (3) is a frequency count of the number of observations in each bucket. Column (4) represents the average value of the ratios of the lower bound over the market price of the option. Column (5) shows the number of violations when the lower bound is higher than the market price
3. The best lower bound performance is when the option is near-in-the-money and short-term maturity (2.83 % below market value)
The underlying stock return distribution is a 10-year historical return histogram with the volatility replaced by the Black-Scholes implied volatility

perform a simple buy and hold trading strategy. If the lower bound is violated, we will buy the option and hold it till maturity. For the 4,233 (out of 46,540) violations, the buy and hold strategy generated \$22,809 or an average of \$5.39 per contract. Given that the buy and hold strategy can be profitable simply due to the bull market, we compute those that had no violation of the lower bound. For the 42,307 cases that violated no lower bound, the average profit is \$1.83. Hence, the options that violated the lower bound imply a trading profit 200 % above the average.





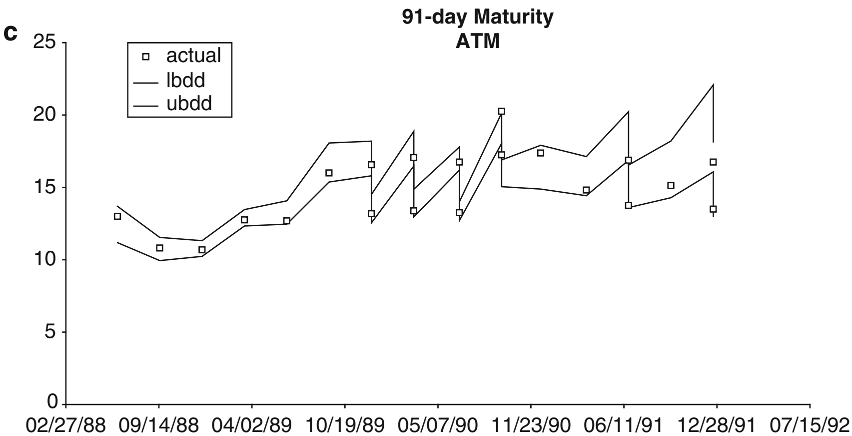


Fig. 7.1 (continued)

d

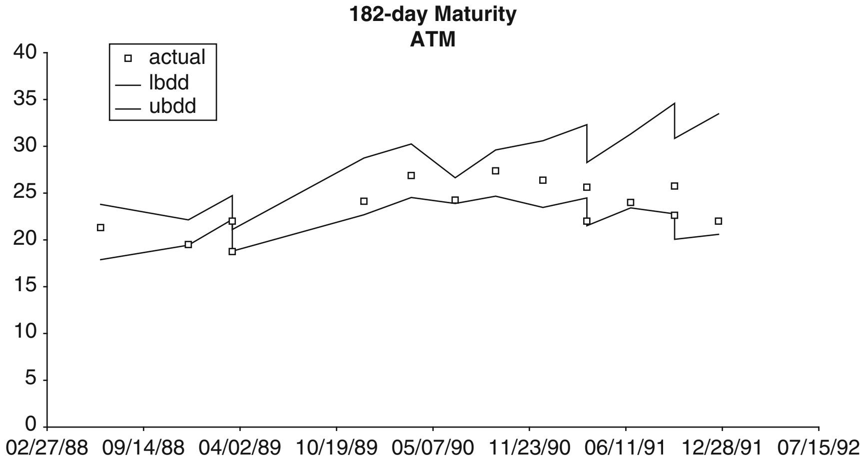


Fig. 7.1 (**a**) Plot of 30-day maturity actual ATM option prices and its upper and lower bound values. (**b**) Plot of 60-day maturity actual ATM option prices and its upper and lower bound values. (**c**) Plot of 91-day maturity actual ATM option prices and its upper and lower bound values. (**d**) Plot of 182-day maturity actual ATM option prices and its upper and lower bound values

7.6 Conclusion

In this paper, we derive a new and tighter lower bound for European option prices comparing with those of previous studies. We further reinterpret Ritchken's (1985) upper bound under a nonparametric framework. Our model contributes to the literature in two different ways. First, our bounds require no parametric assumption of the underlying stock or the moments of the distribution. Furthermore, our bounds require no assumptions on interest rates or volatility. The only requirements of our model are the dollar beta of the option and expected payoffs of the stock and the option. Hence, our bounds can be applied to any model such as the random volatility and random interest rate model by Bakshi et al. (1997) and Scott (1997). Second, despite of much looser and flexible assumptions, our bounds are significantly tighter than the existing upper and lower bound models. Most importantly, our bounds are tighter for the out-of-the-money options that cannot be bounded efficiently by previous models. Finally, we apply our model to real data using histograms of the realized stock returns. The results show that nearly 10 % of the observations violate the lower bound. These violations are shown to generate significant arbitrage profits, after correction of the bull market in the sample period.

Appendix 1

In this appendix, we prove Theorem 7.1. Without loss of generality, we prove Theorem 7.1 by a three-point convex function. The extension of the proof to multiple points is straightforward but tedious. Let the distribution be trinomial and the relationship between the pricing kernel M and the stock price S be convex. In the following table, $x \leq 0$; $y > \varepsilon \geq 0$.

Probability	M	S	$C = \max\{S - K, 0\}$
p^2	$M + x$	$S + y > K$	$S + y - K$
$2p(1 - p)$	M	$S - \varepsilon < K$	0
$(1 - p)^2$	$M - x$	$S - y < K$	0

When $\varepsilon = 0$, the relationship between the pricing kernel M and stock price S is linear, and we obtain equality. When $\varepsilon > 0$, the relationship is convex. We first calculate the mean values:

$$\begin{aligned} &E[M] = M - x + 2px \\ &E[S] = S - y + 2p(y - \varepsilon) + 2p^2\varepsilon \\ &E[C] = p^2(S + y - K). \end{aligned}$$

The three covariances are computed as follows:

$$\begin{aligned} &\mathrm{cov}[M, S] = 2p(1 - p)x(y + \varepsilon(2p - 1)) \\ &\mathrm{cov}[M, C] = 2p^2(1 - p)x(S + y - K) \\ &\mathrm{cov}[S, C] = 2p^2(1 - p)(S + y - K)(y + p\varepsilon). \end{aligned}$$

The variance of the stock price is more complex:

$$\mathrm{var}[S] = 2p(1 - p)z$$

where

$$z = \left(y^2 + \varepsilon^2\left(1 - 2p + 2p^2\right) + 2\varepsilon y(2p - 1)\right) > 0.$$

As a result, it is straightforward to show that

$$\begin{aligned} \frac{\mathrm{cov}[S, C]}{\mathrm{var}[S]}\mathrm{cov}[M, S] &= \frac{2p^2(1 - p)(S + y - K)(y + p\varepsilon)}{2p(1 - p)z} 2p(1 - p)x(y + \varepsilon(2p - 1)) \\ &= 2p^2(1 - p)x(S + y - K)\frac{(y + p\varepsilon)(y + \varepsilon(2p - 1))}{z} \\ &= 2p^2(1 - p)x(S + y - K)\left[1 + \frac{\varepsilon(1 - p)(y - \varepsilon)}{z}\right] \\ &\leq 2p^2(1 - p)x(S + y - K) =: \mathrm{cov}[M, C]. \end{aligned}$$

The fourth line is obtained because cov$[M,C] < 0$ and $1 + \frac{\varepsilon(1-p)(y-\varepsilon)}{z} > 1$. Note that the result is independent of p since all it needs is $0 < p < 1$ for $1 + \frac{\varepsilon(1-p)(y-\varepsilon)}{z}$ to be greater than 1. Also note that when $\varepsilon = 0$ the equality holds.

References

Bakshi, G., Cao, C., & Chen, Z. (1997). Empirical performance of alternative option pricing models. *Journal of Finance, 52*, 2003–2049.

Basso, A., & Pianca, P. (1997). Decreasing absolute risk aversion and option pricing bounds. *Management Science, 43*, 206–216.

Black, F. (1976). The pricing of commodity contacts. *Journal of Financial Economics, 3*, 167–179.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy, 81*, 637–654.

Boyle, P., & Lin, X. (1997). Bounds on contingent claims based on several assets. *Journal of Financial Economics, 46*, 383–400.

Broadie, M., & Cao, M. (2008). Improved lower and upper bound algorithms for pricing American options by simulation. *Quantitative Finance, 8*, 845–861.

Chen, R. R., Lin, H. C., & Palmon, O. (2006). Explaining the volatility smile: Reduced form versus structural option models. *2006 FMA Annual Meeting*. Salt Lake City, UT.

Christoffersen, P., Elkamhi, R., Feunou, B., & Jacobs, K. (2010). Option valuation with conditional heteroskedasticity and nonnormality. *Review of Financial Studies, 23*, 2139–2189.

Chuang, H., Lee, C. F. & Zhong, Z. K. (2011). *Option bounds: A review and comparison* (Working paper).

Chung, S. L., Hung, M. W., & Wang, J. Y. (2010). Tight bounds on American option prices. *Journal of Banking and Finance, 34*, 77–89.

Constantinides, G., & Zariphopoulou, T. (2001). Bounds on derivative prices in an intertemporal setting with proportional transaction costs and multiple securities. *Mathematical Finance, 11*, 331–346.

Constantinides, G. M., Jackwerth, J. C., & Perrakis, S. (2009). Mispricing of S&P 500 index options. *Review of Financial Studies, 22*, 1247–1277.

Constantinides, G. M., Czerwonko, M., Jackwerth, J. C., & Perrakis, S. (2011). Are options on index futures profitable for risk-averse investors? Empirical evidence. *Journal of Finance, 66*, 1407–1437.

Cox, J., & Ross, S. (1976). The valuation of options for alternative stochastic process. *Journal of Finance Economics, 3*, 145–166.

De La Pena, V., Ibragimov, R., & Jordan, S. (2004). Option bounds. *Journal of Applied Probability, 41*, 145–156.

Frey, R., & Sin, C. (1999). Bounds on European option prices under stochastic volatility. *Mathematical Finance, 9*, 97–116.

Gotoh, J. Y., & Konno, H. (2002). Bounding option prices by semidefinite programming: A cutting plane algorithm. *Management Science, 48*, 665–678.

Grundy, B. (1991). Option prices and the underlying assets return distribution. *Journal of Finance, 46*, 1045–1069.

Hobson, D., Laurence, P., & Wang, T. H. (2005). Static-arbitrage upper bounds for the prices of basket options. *Quantitative Finance, 5*, 329–342.

Huang, J. (2004). Option pricing bounds and the elasticity of the pricing kernel. *Review of Derivatives Research, 7*, 25–51.

Ingersoll, J. (1989). *Theory of financial decision making*. Totowa, New Jersey: Rowman & Littlefield.

Levy, H. (1985). Upper and lower bounds of put and call option value: Stochastic dominance approach. *Journal of Finance, 40*, 1197–1217.

Lo, A. (1987). Semi-parametric upper bounds for option prices and expected payoff. *Journal of Financial Economics, 19*, 373–388.

Merton, R. (1973). The theory of rational option pricing. *The Bell Journal of Economics and Management Science, 4*, 141–183.

Peña, J., Vera, J. C., & Zuluaga, L. F. (2010). Static-arbitrage lower bounds on the prices of basket options via linear programming. *Quantitative Finance, 10*, 819–827.

Perrakis, S. (1986). Option pricing bounds in discrete time: Extensions and the pricing of the American put. *Journal of Business, 59*, 119–141.

Perrakis, S., & Ryan, J. (1984). Option pricing in discrete time. *Journal of Finance, 39*, 519–525.

Ritchken, P. (1985). On option pricing bounds. *Journal of Finance, 40*, 1219–1233.

Ritchken, P., & Kuo, S. (1988). Option bounds with finite revision opportunities. *Journal of Finance, 43*, 301–308.

Ritchken, P., & Kuo, S. (1989). On stochastic dominance and decreasing absolute risk averse option pricing bounds. *Management Science, 35*, 51–59.

Rodriguez, R. (2003). Option pricing bounds: Synthesis and extension. *Journal of Financial Research, 26*, 149–164.

Rubinstein, M. (1976). The valuation of uncertain income stream and the pricing of options. *The Bell Journal of Economics, 7*, 407–425.

Scott, L. (1997). Pricing stock options in a jump diffusion model with stochastic volatility and interest rates: Applications of Fourier inversion methods. *Mathematical Finance, 7*, 413–424.

Zhang, P. (1994). Bounds for option prices and expected payoffs. *Review of Quantitative Finance and Accounting, 4*, 179–197.

Can Time-Varying Copulas Improve the Mean-Variance Portfolio?

8

Chin-Wen Huang, Chun-Pin Hsu, and Wan-Jiun Paul Chiou

Contents

C.-W. Huang
Department of Finance, Western Connecticut State University, Danbury, CT, USA
e-mail: huangc@wcsu.edu

C.-P. Hsu (✉)
Department of Accounting and Finance, York College, The City University of New York, Jamaica, NY, USA
e-mail: chsu@york.cuny.edu

W.-J.P. Chiou
Department of Finance and Law College of Business Administration, Central Michigan University, Mount Pleasant, MI, USA
e-mail: Chiou1P@cmich.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_8,

Abstract

Research in structuring asset return dependence has become an indispensable element of wealth management, particularly after the experience of the recent financial crises. In this paper, we evaluate whether constructing a portfolio using time-varying copulas yields superior returns under various weight updating strategies. Specifically, minimum-risk portfolios are constructed based on various copulas and the Pearson correlation, and a 250-day rolling window technique is adopted to derive a sequence of time-varied dependencies for each dependence model. Using daily data of the G7 countries, our empirical findings suggest that portfolios using time-varying copulas, particularly the Clayton dependence, outperform those constructed using Pearson correlations. The above results still hold under different weight updating strategies and portfolio rebalancing frequencies.

Keywords

Copulas • Time-varying dependence • Portfolio optimization • Bootstrap • Out-of-sample return • Performance evaluation • GARCH • Gaussian copula • Student's *t*-copula • Gumbel copula • Clayton copula

8.1 Introduction

The return of a portfolio depends heavily on its asset dependence structure. Over the past decade, copula modeling has become a popular alternative to Pearson correlation modeling when describing data with an asymptotic dependence structure and a non-normal distribution.[1] However, several critical issues attached to the applications of copulas emerge: Do portfolios using time-varying copulas outperform those constructed with Pearson correlations? How does the risk return of copula-based portfolios change over the business cycle? The estimation of parameters has become particularly critical for finance academics and professionals on the heels of the recent financial crises. In this chapter, we model the time-varying dependence of an international equity portfolio using several copula functions and the Pearson correlation. We investigate whether a portfolio constructed with copula dependence yields superior returns as compared to a portfolio constructed using a Pearson correlation under various weight updating strategies.

This paper extends the existing literature in two ways. First, we estimate our time-varying copulas using a rolling window of the latest 250 trading days. It is well accepted that the dependencies between asset returns are time varying (Kroner and Ng 1998; Ang and Bekaert 2002). Differing from the regime-switching type used in Rodriguez (2007) and Okimoto (2008) or the time-evolving type GARCH model used in Patton (2006a), we estimate time-varying copulas via a rolling window

[1]See Chan et al. (1999), Dowd (2005), Patton (2006a), Engle and Sheppardy (2008), Chollete et al. (2009), Bauer and Vorkink (2011).

based on daily data from the previous year. The rolling window method has several benefits. First, it is a method frequently adopted by practitioners. Second, the rolling window method considers only the past year's information when forming dependencies, thus avoiding disturbances that may have existed in the distant past. Several studies have applied this technique, such as Aussenegg and Cech (2011). However, Aussenegg and Cech (2011) considered only the daily Gaussian and Student's *t*-copulas in constructing models, and it is reasonable to consider monthly and quarterly frequencies, given that portfolio managers do not adjust their portfolios on a daily basis. Our research also extends Aussenegg and Cech's (2011) study by including the Archimedean copulas to govern the strength of dependence.

Second, our study investigates how the choice of copula functions affects portfolio performance during periods of economic expansion and recession. The expansion and recession periods we define are based on announcements from the National Bureau of Economic Research (NBER). While the use of copula functions in financial studies has grown enormously, little work has been done in comparing copula dependencies under different economic states.

Using daily US dollar-denominated Morgan Stanley Capital International (MSCI) indices of the G7 countries, our empirical results suggest that the copula-dependence portfolios outperform the Pearson-correlation portfolios. The Clayton-dependence portfolios, for most scenarios studied, deliver the highest portfolio returns, indicating the importance of lower-tail dependence in building an international equity portfolio. Moreover, the choice of weight updating frequency matters. As we increase the weight updating frequency from quarterly to monthly, the portfolio returns for the full sample and recession periods also increase, regardless of the choice of dependence measure. Our finding supports the value of active portfolio reconstruction during recession periods.

This paper is organized as follows. Section 8.2 reviews the literature on copula applications in portfolio modeling. Section 8.3 describes the empirical models. Section 8.4 presents the data used. The main empirical results are reported in Sect. 8.5. Section 8.6 concludes.

8.2 Literature Review

Copulas, implemented in either static or time-varying fashion, are frequently seen in options pricing, risk management, and portfolio selection. In this section, we review some copula applications in portfolio selection. Patton (2006a) pioneered time-varying copulas by modifying the copula functional form to allow its parameters to vary. Patton (2006a) used conditional copulas to examine asymmetric dependence in daily Deutsche mark (DM)/US dollar (USD) and Japanese yen (Yen)/US dollar (USD) exchange rates. His empirical results suggest that the correlation between DM/USD and Yen/USD exchange rates is stronger when the DM and Yen are depreciating against the dollar. Hu (2006) adopted a mixture of a Gaussian copula, a Gumbel copula, and a Gumbel survival copula to examine the various dependence structures of four stock indices. His results demonstrate the underestimation problem due to multivariate normality correlations and the importance of incorporating both the structure and the degree of dependence into the portfolio evaluation. Kole et al. (2007) compared the Gaussian, Student's *t*, and Gumbel copulas to illustrate the importance of selecting an appropriate copula to manage the risk of a portfolio composed of stocks, bonds, and real estate. Kole et al. (2007) empirically demonstrated that the Student's *t*-copula, which considers the dependence both in the center and the tail of the distribution, provides the best fit for the extreme negative returns of the empirical probabilities under consideration.

Rodriguez (2007) studied financial contagions in emerging markets using switching Frank, Gumbel, Clayton, and Student's *t*-copulas. Rodriguez (2007) found evidence that the dependence structures between assets change during a financial crisis and that a good asset allocation strategy should allow dependence to vary with time. Chollete et al. (2009) modeled asymmetric dependence in international equity portfolios using a regime-switching, canonical vine copula approach, which is a branch of the copula family first described by Aas et al. (2007). Chollete et al. (2009) documented that the canonical vine copula provides better portfolio returns and that the choice of copula dependencies affects the VaR of the portfolio return. Chollete et al. (2011) investigated international diversification benefits using the Pearson correlation and six copula functions. Their results show that dependence increases over time and that the intensity of the asymmetric dependence varies in different regions of the world.

While some existing studies have applied copulas to the optimization of portfolio selection, most have tended to focus on portfolio risks, i.e., value at risk, rather than portfolio returns. Empirically, however, investors pay at least equal attention to portfolio returns. Our study is among the few that have focused on equity portfolio returns using time-varying copulas.

8.3 Empirical Methods

8.3.1 Copulas

A copula C is a function that links univariate distribution functions into a multivariate distribution function. Let F be an n-dimensional joint distribution function and let $U = (u_1, u_2, \ldots, u_n)^{\mathrm{T}}$ be a vector of n random variables with marginal distributions $F_1, F_2, \ldots, F_n$. According to Sklar's (1959) theorem, if the marginal distributions $F_1, F_2, \ldots, F_n$ are continuous, then a copula C exists, where C is a multivariate distribution function with all uniform (0,1) marginal distributions.[2] That is,

$$F(u_1, u_2, \ldots, u_n) = C(F_1(u_1), F_2(u_2), \ldots, F_n(u_n)), \text{ for all } u_1, u_2, \ldots, u_n \in \mathbb{R}^n. \quad (8.1)$$

[2]For detailed derivations, please refer to Cherubini et al. (2004), Demarta and McNeil (2005), Embrechts et al. (2003), Embrechts et al. (2005), Franke et al. (2008), Nelson (2006), and Patton (2009).

Table 8.1 The characteristics of different copulas

Dependence model	Tail dependence	Parameter range
Pearson correlation	No	$\rho \in (-1, 1)$
Gaussian copula	No	$\rho \in (-1, 1)$
Student's *t*-copula	Yes (symmetry)	$\rho \in (-1, 1)$, $v > 2$
Gumbel copula	Yes (upper tail)	$\delta \in (0, 1)$
Clayton copula	Yes (lower tail)	$\alpha \in [-1, \infty)\backslash\{0\}$

For a bivariate case, the model can be defined as

$$F(x, y) = C(F_X(\mathrm{x}), F_Y(\mathrm{y})) \tag{8.2}$$

8.3.2 Copula Specifications

In this paper, we consider four copula functions: the Gaussian, Student's *t*, Gumbel, and Clayton. The Gaussian copula focuses on the center of the distribution and assumes no tail dependence. The Student's *t*-copula stresses both the center of the distribution and symmetric tail behaviors. The Clayton copula emphasizes lower-tail dependence, while the Gumbel copula focuses on upper-tail dependence. Table 8.1 summarizes the characteristics of each copula in detail.

8.3.3 Gaussian Copula

The Gaussian copula is frequently seen in the finance literature due to its close relationship to the Pearson correlation. It represents the dependence structure of two normal marginal distributions. According to Nelson (2006), the bivariate Gaussian copula can be expressed as

$$\begin{aligned} C(x, y) &= \int_{-\infty}^{\Phi^{-1}(x)} ds \int_{-\infty}^{\Phi^{-1}(y)} dt \frac{1}{2\pi\sqrt{1-\rho^2}} \exp\left\{-\frac{s^2 - 2\rho st + t^2}{2(1-\rho^2)}\right\} \\ &= \Phi_\rho\left(\Phi^{-1}(x), \Phi^{-1}(y)\right), \end{aligned} \tag{8.3}$$

where Φ denotes the univariate standard normal distribution function and Φ_ρ is the joint distribution function of the bivariate standard normal distribution with correlation coefficient $-1 \le \rho \le 1$. The Gaussian copula has no tail dependence unless $\rho = 1$.

8.3.4 Student's *t*-Copula

Unlike the Gaussian copula, which fails to capture tail behaviors, the Student's *t*-copula depicts the dependence in the center of the distribution as well as in the tails. The Student's *t*-copula is defined using the multivariate t distribution and can be written as

$$C^t_{v,\rho}(x,y) = \int_{-\infty}^{t_v^{-1}(x)}\int_{-\infty}^{t_v^{-1}(y)} \frac{1}{2\pi\sqrt{1-\rho^2}}\left\{1+\frac{s^2-2\rho st+t^2}{v(1-\rho^2)}\right\}^{-\frac{v+2}{2}} dsdt$$
$$= t^2_{v,r}\left(t_v^{-1}(x), t_v^{-1}(y)\right), \tag{8.4}$$

where $t^2_{v,r}$ indicates the bivariate joint t distribution, $t_v{}^{-1}$ is the inverse of the distribution of a univariate t distribution, v is the degrees of freedom, and ρ is the correlation coefficient of the bivariate t distribution when $v > 2$.

8.3.5 Archimedean Copulas

According to Embrechts et al. (2005), the coefficient of upper-tail dependence (λ_u) of two series, X and Y, can be defined as

$$\lambda_u(x,y) = \lim_{q-1^-} P\left[Y > F_y^{\leftarrow}(q) \middle| X > F_x^{\leftarrow}(q)\right]. \tag{8.5}$$

The upper-tail dependence presents the probability that Y exceeds its qth quantile given that X exceeds its qth quantile, considering the limit as q goes to its infinity. If the limit $\lambda_u \in [0,1]$ exists, then X and Y are said to show upper-tail dependence. In the same manner, the coefficient of lower-tail dependence (λ_1) of X and Y is described as

$$\lambda_l(x,y) = \lim_{q-0^+} P\left[Y \le F_y^{\leftarrow}(q) \middle| X \le F_x^{\leftarrow}(q)\right]. \tag{8.6}$$

Since both F_{U_1} and F_{U_2} are continuous density functions, the upper-tail dependence can be presented as

$$\lambda_u = \lim_{q-0^-} \frac{P\left[Y > F_y^{\leftarrow}(q) \middle| X > F_x^{\leftarrow}(q)\right]}{P\left[X > F_x^{\leftarrow}(q)\right]}. \tag{8.7}$$

For lower-tail dependence, it can be described as

$$\lambda_l = \lim_{q-0^+} \frac{P\left[Y \le F_y^{\leftarrow}(q) \middle| X \le F_x^{\leftarrow}(q)\right]}{P\left[X \le F_x^{\leftarrow}(q)\right]}. \tag{8.8}$$

8.3.5.1 Gumbel Copula

The Gumbel copula is a popular upper-tail dependence measure as suggested in Embrechts et al. (2005). The Gumbel copula can be written as

$$\mathrm{c}(x,y) = \exp\left[-\left\{\left(-\ln(\mathrm{x})^{\frac{1}{\delta}} + (-\ln(\mathrm{y})^{\frac{1}{\delta}}\right\}^{\delta}\right], \tag{8.9}$$

where $0 < \delta \leq 1$ measures the degree of dependence between X and Y. When $\delta = 1$, X and Y do not have upper-tail dependence (i.e., X and Y are independent at upper tails), and when $\delta \to 0$, X and Y have perfect dependence.

8.3.5.2 Clayton Copula

The Clayton copula is used to measure lower-tail dependence. The Clayton copula is defined as

$$\mathrm{c}(x, y) = \max\left[(\mathrm{x}^{-\alpha} + \mathrm{y}^{-\alpha} - 1)^{-\frac{1}{\alpha}}, 0\right], \tag{8.10}$$

where α describes the strength of dependence. If $\alpha \to 0$, X and Y do not have lower-tail dependence. If $\alpha \to \infty$, X and Y have perfect dependence.

8.3.6 Portfolio Construction

The selection of optimal portfolios draws on the seminal work of Markowitz (1952). Specifically, we adopt the variance minimization strategy with no short selling and transaction cost assumptions.[3] An optimal portfolio allocation can be found by solving the following optimization problem:

$$\mathrm{Min}_{\{w\}} w' V w$$

$$\text{Subject to} \sum \mathrm{w_i} = 1, \mathrm{w}_i \geq 0, \tag{8.11}$$

where w_i is the weight of asset i, and V is the covariance matrix of asset returns.

Because dependence is a time-varying parameter, the data from a subset of 250 trading days prior to the given sample date t is used to derive the dependence for date t. With 1,780 daily data points in our sample, we calculate a total of 1,531 dependencies for each copula method and the Pearson correlation. With these dependencies, optimal portfolio weightings can be obtained by solving a quadratic function subject to specified constraints. The optimal weightings for time t are used to calculate the realized portfolio returns for $(t + 1)$.[4]

In practice, portfolio managers periodically reexamine and update the optimal weights of their portfolios. If the asset allocation of an existing portfolio has deviated from the target allocation to a certain degree and if the benefit of updating exceeds its costs, a portfolio reconstruction action is executed. In this paper, we construct a comprehensive study of portfolio returns by varying the state of

[3]Short selling usually involves other service fees, which vary depending on the creditability of the investors. Because the focus of this study is on the effect of the dependence structure on portfolio performance, we assume that short selling is not allowed to simplify the comparison.

[4]For example, we use return data from t_1 to t_{250} to calculate the optimal portfolio weights with dependencies estimated from the copulas and the Pearson correlation. The optimal portfolio weights are applied to the return data at t_{251} to calculate the realized portfolio returns.

Table 8.2 The summary statistics of the G7 indices

	Canada	France	Germany	Italy	Japan	UK	USA
Mean (%)	0.0301	0.0064	0.0082	0.0003	−0.0017	−0.0039	−0.0082
Std. dev.	0.0164	0.0171	0.0178	0.0161	0.0155	0.0159	0.0144
Skewness	−0.8781	0.0740	0.0666	0.0477	−0.1475	−0.0535	−0.1365
Kurtosis	14.1774	10.7576	8.6920	12.9310	7.4592	12.9143	12.1182
Jarque-Bera	9494	4465	2404	7315	1481	7290	6171
JB P-value	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
Observations	1780	1780	1780	1780	1780	1780	1780

The results indicate that the daily returns of the G7 indices are not normally distributed

the economy (i.e., expansion or recession), the dependence structure, and the frequency of weight updating, i.e., quarterly, monthly, and daily. Quarterly updating allows investors to update the optimal weights on the first trading days of March, June, September, and December; monthly updating allows investors to update the optimal weights on the first trading days of each month. Under daily updating, investors update the optimal weights every trading day.

8.4 Data

The data are the US dollar-denominated daily returns of the Morgan Stanley Capital International (MSCI) indices for the G7 countries, including Canada, France, Germany, Italy, Japan, the UK, and the USA. The sample period covers from the first business day in June 2002 to the last business day in June 2009, for a total of 1,780 daily observations. Based on the definitions provided by the National Bureau of Economic Research, we separate the data into an expansion period from June 2002 to November 2007 and a recession period from December 2007 to June 2009.

Table 8.2 presents the descriptive statistics. Among the G7 countries, Canada had the highest daily returns, while the USA had the lowest. Germany, however, experienced the most volatile returns. All return series exhibit high kurtosis, suggesting fat tails on return distributions. The results of the Jarque-Bera test reject the assumption that the G7 indices have normal distributions.

8.5 Empirical Results

8.5.1 Dependence

Using 1,780 daily data points from the G7 countries, for each dependence model, we estimate 21 dependence pairs, each containing a sequence of 1,531 dependencies. The parameters for the Gaussian, Student's *t*, Gumbel, and Clayton copula functions are estimated using the two-stage *inference for the margins* (IFM) method proposed by Joe and Xu (1996) and Joe (1997).

The dependencies from the Pearson correlation are calculated using the standard method. Appendix 1 shows the maximum and the minimum of the 21 dependence pairs of each dependence model.

The graphs in Fig. 8.1 show the dependencies between the USA and other countries as estimated via the various copulas and the Pearson correlation. In general, the Gaussian copula estimation is similar to that of the corresponding Pearson correlation, but the Student's *t*-copulas show significant jumps over time. For our sample period, Japan shows a low dependence with the US market as compared to other economies.

8.5.2 Average Portfolio Returns

Table 8.3 presents the average portfolio returns for the full sample period, the expansion period, and the recession period for the quarterly, monthly, and daily weight updating strategies. Under the quarterly weight updating, the Clayton-dependence portfolios have the highest average returns at 6.07 % in the expansions and −12.52 % in the recessions, and the Pearson-correlation portfolios have the lowest average returns, 5.48 % in the expansions and −14.25 % in the recessions. The order of portfolio performance, in the form of its dependence model regardless of the state of economy, is as follows: the Clayton copula, the Gumbel copula, the Student's *t*-copula, the Gaussian copula, and the Pearson correlation. Because both the Clayton and Gumbel copulas highlight the tail dependence between assets, the empirical evidence suggests that with a quarterly weighting strategy, tail dependence, particularly lower-tail dependence, is important for obtaining superior average portfolio returns across different states of the economy.

As we increase the updating frequency from quarterly to monthly, similar empirical results are observed. That is, the Clayton-copula portfolios yield the highest average returns, while the Pearson-correlation portfolios provide the lowest average returns. In the expansion periods, the order of portfolio performance, in the form of its dependence model, is as follows: the Clayton copula, the Student's *t*-copula, the Gaussian copula, the Gumbel copula, and the Pearson correlation. In recession periods, the order of portfolio performance, in the form of its dependence model, is as follows: the Clayton copula, the Gumbel copula, the Student's *t*- copula, the Gaussian copula, and the Pearson correlation. According to Kole et al. (2007), the Gaussian copula, which does not consider lower-tail dependence, tends to be overly optimistic on the portfolio's diversification benefits, and the Gumbel copula, which focuses on the upper tail and pays no attention to the center of the distribution, tends to be overly pessimistic on the portfolio's diversification benefits. We verify this argument by observing that the Gumbel-copula portfolio outperforms only the Pearson-correlation portfolio in the expansion periods, while the Gaussian-copula-dependence portfolio outperforms the Pearson-correlation portfolio only in recession periods. Interestingly, as we increase the weight updating frequency from quarterly to monthly, the average portfolio returns for the full sample and recession periods also

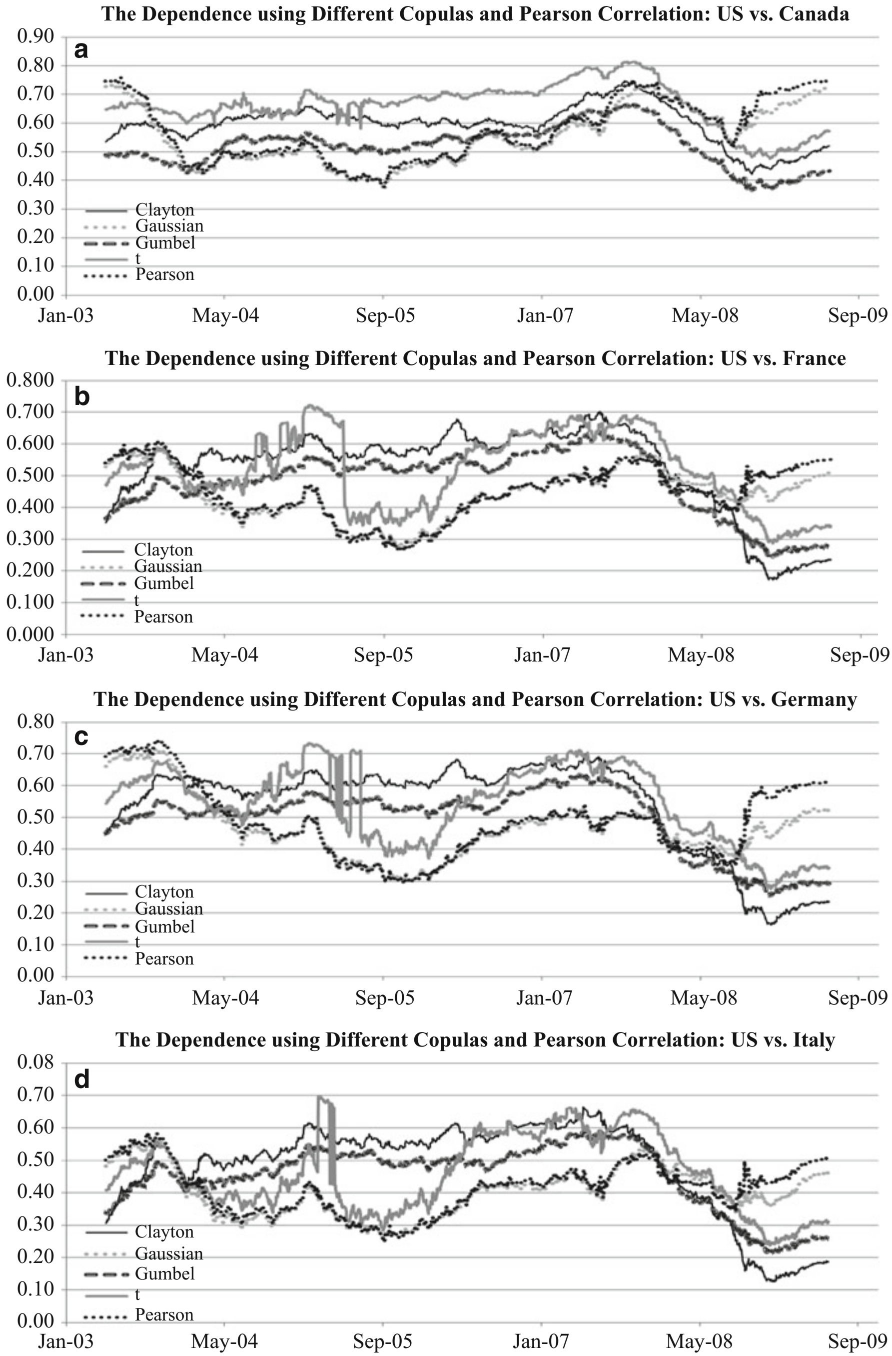


Fig. 8.1 (continued)

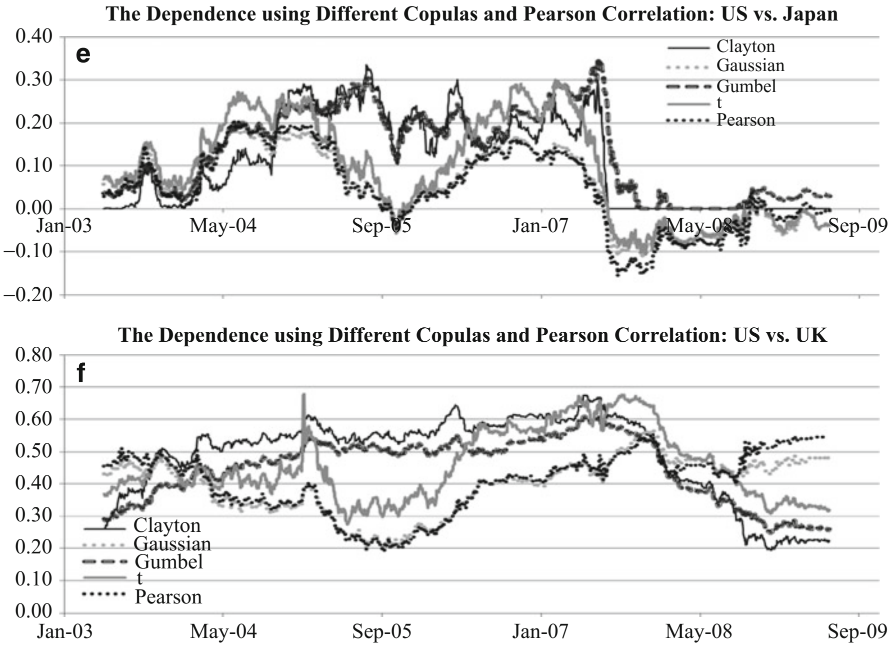


Fig. 8.1 The dependence using different copulas and Pearson correlation. Panel (**a**) USA versus Canada. Panel (**b**) USA versus France. Panel (**c**) USA versus Germany. Panel (**d**) USA versus Italy. Panel (**e**) USA versus Japan. Panel (**f**) USA versus UK

increase, regardless of the choice of dependence measures. Thus, the empirical results seem to support the need for active portfolio reconstruction during recessions.

As the weight updating frequency increases to daily, the Clayton copula delivers only the highest average portfolio returns during the expansion period. The Student's *t*-copula, by contrast, generates the highest portfolio average returns for the full sample and recession periods. The influence of the lower-tail dependence seems to diminish under daily weight reconstruction. The Gaussian-copula portfolio delivers the worst portfolio performance in both expansion and recession periods.

8.5.3 Testing Significance of Return Difference

The results reported in the previous section show the average portfolio returns for different dependencies and weight updating frequencies. One issue with average returns is that if extreme values exist over the examined period, the empirical results may be biased and relatively high-standard deviations will be reported. Previous methods of examining the robustness of portfolio performance usually build on the data normality assumption (Jobson and Korkie 1981; Memmel 2003), which goes against the empirical facts.

Table 8.3 Average portfolio returns

	Clayton	Gaussian	Gumbel	Student's t	Pearson
Panel A: quarterly adjustments					
Full sample returns	1.44 %	0.91 %	1.16 %	1.11 %	0.57 %
	(1.1903)	(1.1749)	(1.1922)	(1.1763)	(1.0644)
Expansion returns	6.07 %	5.63 %	5.70 %	5.66 %	5.48 %
	(0.6946)	(0.0798)	(0.6962)	(0.7014)	(0.6613)
Recession returns	−12.52 %	−13.35 %	−12.56 %	−12.64 %	−14.25 %
	(2.0549)	(2.0025)	(2.0578)	(1.7922)	(2.0153)
Panel B: monthly adjustments					
Full sample returns	1.63 %	1.08 %	1.22 %	1.22 %	0.75 %
	(1.1764)	(1.1812)	(1.1835)	(1.1736)	(1.0199)
Expansion returns	6.15 %	5.67 %	5.66 %	5.69 %	5.23 %
	(0.6800)	(0.6928)	(0.6870)	(0.6864)	(0.6401)
Recession returns	−12.01 %	−12.78 %	−12.19 %	−12.30 %	−12.80 %
	(2.0375)	(2.0355)	(2.0474)	(2.0247)	(1.1708)
Panel C: daily adjustments					
Full sample returns	1.34 %	0.85 %	1.16 %	1.47 %	1.03 %
	(1.1737)	(1.1651)	(1.1749)	(1.1733)	(1.0253)
Expansion returns	5.84 %	5.35 %	5.59 %	5.70 %	5.39 %
	(0.6733)	(0.6859)	(0.6839)	(0.6766)	(0.6367)
Recession returns	−12.27 %	−12.74 %	−12.22 %	−11.34 %	−12.14 %
	(2.0382)	(2.0053)	(2.0277)	(2.0346)	(1.7280)

The average portfolio returns are presented in an annualized, percentage format. Three weight updating frequencies are considered: quarterly, monthly, and daily. Within each frequency, we report the returns for the full sample period, the expansion period, and the recession period. The numbers in the parentheses are standard errors

To cope with this problem, Ledoit and Wolf (2008) proposed an alternative testing method using the inferential studentized time-series bootstrap. Ledoit and Wolf's (2008) method is as follows.[5] Let a and b be two investment strategies, and let r_{at} and r_{bt} be the portfolio returns for strategies a and b, respectively, at time t, where t ranges from 1 to i. The mean vector μ and the covariance matrix Σ for the return pairs $(r_{a1},r_{b1})',\ldots,(r_{at},r_{bt})'$ are denoted by

$$\mu = \begin{pmatrix} \mu_a \\ \mu_b \end{pmatrix} \text{ and } \sum = \begin{pmatrix} \sigma_a^2 & \sigma_{ab} \\ \sigma_{ab} & \sigma_b^2 \end{pmatrix}. \tag{8.12}$$

The performance of strategies a and b can be examined by checking whether the difference between the Sharpe ratios for strategies a and b are statistically different from 0. That is,

$$\Delta = S_a - S_b = \frac{\mu_a}{\sigma_a} - \frac{\mu_b}{\sigma_b} \tag{8.13}$$

[5]For detailed derivations and computer codes, please refer to Ledoit and Wolf (2008).

and

$$\hat{\Delta} = \widehat{S_a} - \widehat{S_b} = \frac{\widehat{\mu_a}}{\widehat{\sigma_a}} - \frac{\widehat{\mu_b}}{\widehat{\sigma_b}}, \tag{8.14}$$

where Δ is the difference between the two Sharpe ratios, and S_a and S_b are the Sharpe ratios for strategies a and b, respectively.

Let the second moments of the returns from strategies a and b be denoted by γ_a and γ_b. Then $\gamma_a = E({\gamma_{at}}^2)$ and $\gamma_b = E({\gamma_{bt}}^2)$. Let υ and $\hat{\upsilon}$ be $(\mu_a, \mu_b, \gamma_a, \gamma_b)$' and $(\hat{\mu_a}, \hat{\mu_b}, \hat{\gamma_a}, \hat{\gamma_a})'$, respectively. Then Δ and $\hat{\Delta}$ can be expressed as

$$\Delta = f(\upsilon) \text{ and } \hat{\Delta} = f(\hat{\upsilon}) \tag{8.15}$$

where $f(\upsilon) = \frac{\mu_a}{\sqrt{\gamma_a - \mu_a^2}} - \frac{\mu_b}{\sqrt{\gamma_b - \mu_b^2}}$ and $\sqrt{i}(\hat{\upsilon} - \upsilon) \xrightarrow{d} N(0; \Psi)$.

For time-series data, Ledoit and Wolf (2008) have argued that Ψ can be evaluated by the studentized bootstrap as $\widehat{\Psi}^* = \frac{1}{\vartheta}\sum_{j=1}^{\vartheta} \xi_j \xi'_j$, where $\xi_j = \frac{1}{\sqrt{b}}\sum_{t=1}^{b} y^*_{(j-1)b+t}\, t = 1, \ldots, \vartheta. \vartheta$ is the integer part of the fraction of the total observations divided by the blocks b. Also,

$$y_t^* = \left(r_{ta}^* - \hat{\mu}_a^*, r_{tb}^* - \hat{\mu}_b^*, r_{ta}^{*2} - \hat{\gamma}_a^*, r_{tb}^{*2} - \hat{\gamma}_b^*\right) t = 1, \ldots, i. \tag{8.16}$$

Following Ledoit and Wolf's (2008) method, we examine the significance of 60 pairs of portfolio performance. The size of the bootstrap iteration is 10,000 to ensure a sufficient sample.[6] Table 8.4 presents the results from Ledoit and Wolf's (2008) portfolio performance test.

The results indicate that during the recession periods and using quarterly weight updating, the Pearson correlation underperforms all the copula dependencies at a confidence level of 90 % or greater. During recession periods and adopting monthly weight updating, the superiority of the copula dependencies jumps to a 99 % confidence level. Moreover, during the recession periods and assuming daily updating, the Student's t-copula outperforms the Pearson correlation at the 99 % confidence level.

Overall, Ledoit and Wolf's (2008) empirical tests illustrate the superiority of copulas during recession periods, regardless of the frequency of weight updating. During a bullish market, this advantage seems not as statistically significant as during a bearish market.

[6]Ledoit and Wolf (2008) suggested that 5,000 iterations guarantee a sufficient sample. We adopt a higher standard of 10,000 iterations to strengthen our testing results.

Table 8.4 Ledoit and Wolf portfolio performance test

Panel A: quarterly adjustments					
	CL-GA	CL-GU	CL-PE	CL-t	GA-GU
Expansion	0.821	0.812	0.788	0.677	0.916
Recession	0.060[a]	0.987	0.054[a]	0.839	0.0760[a]
	GA-PE	GA-t	GU-PE	GU-t	PE-t
Expansion	0.892	0.930	0.766	0.912	0.778
Recession	0.0030[c]	0.0727[a]	0.0267[b]	0.943	0.026[b]
Panel B: monthly adjustments					
	CL-GA	CL-GU	CL-PE	CL-t	GA-GU
Expansion	0.193	0.415	0.568	0.744	0.881
Recession	0.249	0.803	0.021[c]	0.295	0.092[a]
	GA-PE	GA-t	GU-PE	GU-t	PE-t
Expansion	0.850	0.892	0.795	0.896	0.809
Recession	0.001[c]	0.318	0.019[c]	0.475	0.008[c]
Panel C: daily adjustments					
	CL-GA	CL-GU	CL-PE	CL-t	GA-GU
Expansion	0.389	0.814	0.732	0.929	0.913
Recession	0.000[c]	0.119	0.173	0.371	0.000[c]
	GA-PE	GA-t	GU-PE	GU-t	PE-t
Expansion	0.928	0.915	0.460	0.831	0.301
Recession	0.718	0.001[c]	0.045[c]	0.778	0.019[c]

The performance tests are conducted using the approach suggested by Ledoit and Wolf (2008). The tests examine whether the returns from two portfolios are significantly different at the 95 % level

CL stands for the Clayton copula, *GA* stands for the Gaussian copula, *GU* stands for the gumbel copula, *PE* stands for Pearson correlation, and *t* stands for the Student's *t*-copula

[a]Represents 90 % statistical significance

[b]Represents 95 % statistical significance

[c]Represents 99 % statistical significance

8.6 Conclusions

In this paper, we study whether adopting time-varying copulas as a measure of dependence of asset returns can improve portfolio performance. This study was motivated by the fact that the traditional Pearson correlation is inadequate in describing most financial returns. Moreover, the robustness of copula functions has not been fully examined under different states of the economy and weight updating scenarios. We evaluate the effectiveness of various copulas in managing portfolios while considering portfolio rebalance frequencies and the business cycle. The significance of return difference is tested using the studentized time-series bootstrap method suggested by Ledoit and Wolf (2008).

The main findings are as follows: first, an international equity portfolio modeled using the Pearson correlations underperforms those modeled using copula-based dependencies, especially during recession periods. Our findings remain robust regardless of the rebalancing frequency. Second, the importance of lower-tail behaviors in portfolio modeling is highlighted by the higher average portfolio returns of the Clayton-dependence portfolios. Third, the choice of weight updating frequency affects portfolio returns. The portfolios using monthly weight updating frequency provide better portfolio returns than those using quarterly and daily weight adjustments.

We add to the current literature by thoroughly evaluating the effectiveness of asymmetric conditional correlations in managing portfolio risk. This paper synthesizes the major concepts and modi operandi of previous research and maximizes the practicality of applying copulas under a variety of scenarios. Future research into copulas can be extended to the contagion of various asset classes and interest rates and evaluations of the impact of certain economic events.

Appendix 1

Appendix 1 illustrates the dependence of the G7 countries from different dependence models. Note that to ease the comparison between dependencies, we transform Gumbel dependencies by $(1 - \delta)$. Therefore, the range for the Clayton and the Gumbel copulas is between 0 and 1, with 0 meaning no dependence and 1 standing for perfect dependence. The range for the Gaussian copula, the Student's t-copula, and the Pearson correlation is -1 to 1, with 0 meaning no dependence and 1 or -1 standing for complete dependence.

	CA	FR	DE	IT	JP	UK	USA
Panel A: Gaussian dependence							
CA							
Max							
Min							
FR							
Max	0.7064						
Min	0.3914						
DE							
Max	0.6487	0.9686					
Min	−0.2016	0.7856					
IT							
Max	0.6320	0.9407	0.9274				
Min	−0.2143	−0.2303	0.7001				
JP							
Max	0.1814	0.2761	0.2478	0.4023			
Min	−0.2115	−0.2238	−0.2229	0.0119			
UK							
Max	0.6393	0.9100	0.8665	0.8575	0.4664		
Min	−0.1945	−0.2384	−0.2008	−0.2050	0.0329		

(*continued*)

	CA	FR	DE	IT	JP	UK	USA
USA							
Max	0.7221	0.5031	0.5231	0.4661	0.5831	0.5671	
Min	−0.1864	−0.2127	−0.2087	−0.2414	−0.2241	0.1997	
Panel B: Student's *t* dependence							
CA							
Max							
Min							
FR							
Max	0.7509						
Min	0.3921						
DE							
Max	0.4578	0.9810					
Min	−0.1070	0.7476					
IT							
Max	0.4367	0.9810	0.9586				
Min	−0.1089	0.7476	0.6937				
JP							
Max	0.1093	0.1679	0.1522	0.6846			
Min	−0.1284	−0.1511	−0.1574	0.0093			
UK							
Max	0.4478	0.8055	0.7380	0.7316	0.7335		
Min	−0.1094	−0.1546	−0.1359	−0.1230	0.0284		
USA							
Max	0.5733	0.8055	0.3457	0.3176	0.4375	0.6614	
Min	−0.1107	−0.1546	−0.1343	−0.1360	−0.1519	0.2687	
Panel C: Gumbel dependence							
CA							
Max							
Min							
FR							
Max	0.5947						
Min	0.3220						
DE							
Max	0.3744	0.9063					
Min	0.0000	0.5928					
IT							
Max	0.3666	0.7286	0.8544				
Min	0.0000	0.0000	0.5516				
JP							
Max	0.0961	0.1384	0.1222	0.5356			
Min	0.0000	0.0000	0.0000	0.0200			

(*continued*)

	CA	FR	DE	IT	JP	UK	USA
UK							
Max	0.3650	0.6946	0.6156	0.6086	0.5855		
Min	0.0000	0.0000	0.0000	0.0000	0.0234		
USA							
Max	0.4340	0.2816	0.2952	0.2649	0.3261	0.5967	
Min	0.0000	0.0000	0.0000	0.0000	0.0000	0.2493	
Panel D: Clayton dependence							
CA							
Max							
Min							
FR							
Max	0.6763						
Min	0.2899						
DE							
Max	0.3585	0.9327					
Min	0.0000	0.6696					
IT							
Max	0.3463	0.7635	0.9004				
Min	0.0000	0.0000	0.6006				
JP							
Max	0.0038	0.0408	0.0287	0.6476			
Min	0.0000	0.0000	0.0000	0.0000			
UK							
Max	0.3657	0.7193	0.6567	0.6506	0.6794		
Min	0.0000	0.0000	0.0000	0.0000	0.0000		
USA							
Max	0.5248	0.2450	0.2476	0.1987	0.3472	0.6622	
Min	0.0000	0.0000	0.0000	0.0000	0.0000	0.1982	
Panel E: Pearson correlation							
CA							
Max							
Min							
FR							
Max	0.7002						
Min	0.3966						
DE							
Max	0.6944	0.9726					
Min	0.3645	0.7890					
IT							
Max	0.6833	0.9596	0.9477				
Min	0.3877	0.8204	0.6965				

(*continued*)

	CA	FR	DE	IT	JP	UK	USA
JP							
Max	0.3549	0.4594	0.4702	0.4073			
Min	−0.0411	0.0251	0.0170	−0.0072			
UK							
Max	0.7080	0.9573	0.9298	0.9181	0.4612		
Min	0.3676	0.7791	0.6572	0.6990	0.0154		
USA							
Max	0.7586	0.6096	0.7443	0.5871	0.2078	0.5480	
Min	0.3764	0.2647	0.2921	0.2481	−0.1562	0.1913	

References

Aas, K., Czado, C., Frigessi, A., & Bakken, H. (2007). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics, 2*, 1–25.

Ang, A., & Bekaert, G. (2002). International asset allocation with regime shifts. *Review of Financial Studies, 15*, 1137–1187.

Ang, A., & Chen, J. (2002). Asymmetric correlations of equity portfolios. *Journal of Financial Economics, 63*, 443–494.

Aussenegg, W., & Cech, C. (2011, forthcoming). Simple time-varying copula estimation. In A. S. Barczak, E. Dziwok (Ed.), *Mathematical econometrical and computational methods in finance and insurance*. Katowice: University of Economics in Katowice.

Bauer, G. H., & Vorkink, K. (2011). Forecasting multivariate realized stock market volatility. *Journal of Econometrics, 160*, 93–101.

Chan, L. K., Karceski, J., & Lakonishok, J. (1999). On portfolio optimization: Forecasting covariances and choosing the risk model. *Review of Financial Studies, 12*, 937–974.

Cherubini, U., Luciano, E., & Vecchiato, W. (2004). *Copula methods in finance*. Hoboken: Wiley.

Chollete, L., Heinen, A., & Valdesogo, A. (2009). Modeling international financial returns with a multivariate regime-switching copula. *Journal of Financial Econometrics, 7*, 437–480.

Chollete, L., Peña, V., & Lu, C. (2011). International diversification: A copula approach. *Journal of Banking and Finance, 35*, 403–417.

Demarta, S., & McNeil, A. (2005). The *t* copula and related copulas. *International Statistical Review, 73*, 111–129.

Dowd, K. (2005). Copulas and coherence. *Journal of Portfolio Management, Fall, 32*(1), 123–127.

Embrechts, P., Frey, R., & McNeil, A. (2005). *Quantitative risk management*. Princeton: Princeton University Press.

Embrechts, P., Lindskog, F., & McNeil, A. (2003). Modelling dependence with copulas and applications to risk management. In S. T. Rachev (Ed.), *Handbook of heavy tailed distributions in finance* (pp. 329–384). Amsterdam: Elsevier/North-Holland.

Engle, R. F., & Sheppardy K. (2008). *Evaluating the specification of covariance models for large portfolios* (Working Paper). New York University.

Franke, J., Hardle, W., & Hafner, C. (2010). *Statistics of financial markets: An introduction*. New York: Springer.

Hu, L. (2006). Dependence patterns across financial markets: A mixed copula approach. *Applied Financial Economics, 16*, 717–729.

Jobson, J., & Korkie, B. (1981). Performance hypothesis testing with the Sharpe and Treynor measures. *Journal of Finance, 36*, 889–908.

Joe, H. (1997). *Multivariate models and dependence concepts*. New York: Chapman & Hall/CRC.

Joe, H., & Xu, J. J. (1996). *The estimation method of inference functions for margins for multivariate models* (Technical Report 166). Department of Statistics, University of British Columbia.

Kole, E., Koedijk, K., & Verbeek, M. (2007). Selecting copulas for risk management. *Journal of Banking and Finance, 31*, 2405–2423.

Kroner, K. E., & Ng, V. (1998). Modeling asymmetric comovements of asset returns. *Review of Financial Studies, 11*, 817–844.

Ledoit, O., & Wolf, M. (2008). Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance, 15*, 850–859.

Longin, F., & Solnik, B. (2001). Extreme correlation of international equity markets. *Journal of Finance, 56*, 649–676.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance, 7*, 77–91.

Memmel, C. (2003). Performance hypothesis testing with the Sharpe ratio. *Finance Research Letters, 1*, 21–23.

Nelson, R. (2006). *An introduction to copulas*. New York: Springer.

Okimoto, T. (2008). New evidence of asymmetric dependence structures in international equity markets. *Journal of Financial and Quantitative Analysis, 43*, 787–816.

Patton, A. (2006a). Modelling asymmetric exchange rate dependence. *International Economic Review, 47*, 527–556.

Patton, A. (2006b). Estimation of multivariate models for time series of possibly different lengths. *Journal of Applied Econometrics, 21*, 147–173.

Patton, A. (2009). Copula-based models for financial time series. In T. Mikosch, J. P. Kreiss, T. G. Andersen, R. A. Davis, J. P. Kreiss, & T. Mikosch (Eds.), *Handbook of financial time series*. New York: Springer.

Rodriguez, J. C. (2007). Measuring financial contagion: A copula approach. *Journal of Empirical Finance, 14*, 401–423.

Sklar, A. (1959) "Fonctions de Répartition à n Dimensions et Leurs Marges," Publications de l'Institut de Statistique de l'Université de Paris 8, 229–231.

Determinations of Corporate Earnings Forecast Accuracy: Taiwan Market Experience

9

Ken Hung and Kuo-Hao Lee

Contents

K. Hung (✉)
Division of International Banking & Finance Studies, Texas A&M International University, Laredo, TX, USA
e-mail: ken.hung@tamiu.edu

K.-H. Lee
Department of Finance, College of Business, Bloomsburg University of Pennsylvania, Bloomsburg, PA, USA
e-mail: klee@bloomu.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_9,

Abstract

Individual investors are actively involved in stock market and are making investment decision based on publicly available and nonproprietary information, such as corporate earnings forecasts from the management and the financial analyst. Also, the management forecast is another important index investors might use.

To examine the accuracy of the earnings forecasts, the following test methodology have been conducted. **Multiple regression models** are used to examine the effect of six factors: firm size, market volatility, trading volume turnover, corporate earnings variances, type of industry, and experience. If the two-sample groups are related, **Wilcoxon two-sample test** will be used to determine the relative earnings forecast accuracy.

The results indicate that firm size has no effect on management forecast, voluntary management forecast, mandatory management forecast, and analysts' forecast. There are some indications that forecasting accuracy is affected by market ups and downs. The results also reveal that relative accuracy of earnings forecasts is not a function of trading volume turnover. However, management's earnings forecast and analysts' forecasts are sensitive to earnings variances.

Readers are well advised and referred to the chapter appendix for methodological issues such as sample selection, variable definition, regression model, and Wilcoxon two-sample test.

Keywords

Multiple regression • Wilcoxon two-sample test • Corporate earnings • Forecast accuracy • Management earnings • Firm size • Corporation regulation • Volatility • Trade turnover • Industry

9.1 Introduction

In recent times, individual investors are actively involved in stock market and are making investment decision based on publicly available and nonproprietary information. Corporate earnings forecasts are an important investment tool for investors. Corporate earnings forecasts come from two sources: the company management and financial analyst. As an insider, the management has the advantage of possessing more information and hence provides a more accurate earnings forecast. However, because of the existing relationship of the company with its key investor group, the management may have a tendency to take an optimistic view and overestimate its future earnings. In contrast, the financial analysts are less informed about the company and often rely on management briefings. They have more experiences in the overall market and economies and are expected to analyze companies objectively. Hence, analysts should provide reliable and more accurate earnings forecast.

Whether investors should rely on the earnings forecast made by the management or by the analyst is a debatable issue. Many researchers have examined the accuracy of such earnings forecasts. Because there are differences in methodologies, sample selections and time horizons, the findings and conclusions from the previous studies are conflicting and inconclusive. This motivated us to do a new analysis by using a different methodology.

According to *Regulations for Publishing Corporate Earnings Forecast* imposed by the Department of Treasury,[1] publicly traded Taiwanese companies have to publish their earnings forecasts under the following situations:

1. To issue new stocks or acquired liabilities;
2. When more than one third of the board has been changed;
3. When one of the situations as listed in section 185 of the corporation regulations happens;
4. Merger and acquisitions;
5. Profit gain/loss up to one third of its annual revenue due to an unexpected event;
6. Revenue loss over 30 % compared to last year;
7. Voluntarily publish its earnings forecast

Since management earnings forecasts are mandatory or voluntary, the focus of this research is to examine the accuracy of management's overall earnings forecast, management voluntary earnings forecast, management mandatory earnings forecast, and financial analyst earnings forecast.

9.2 Literature Review

Jaggi (1980) examined the impact of company size on forecast accuracy using management's earnings forecasts from the Wall Street Journal and analysts' earnings forecasts from Value Line Investment Service from 1971 to 1974. He argued that because a larger company has strong financial and human capital resources, its management's earnings forecast would be more accurate than the analyst's. The sample data were classified into six categories based on the size of the firms' total revenue to examine the factors that attribute to the accuracy of management's earnings forecast with the analyst's. The result of his research did not support his hypothesis that management's forecast is more accurate than the analyst's.

Bhushan (1989) assumed that it is more profitable trading large companies' stocks because large companies have better liquidity than the small ones. Therefore, the availability of information is related to company size. His research results support his hypothesis that the larger the company size, the more information is available to financial analysts and the more accurate their earnings forecasts are. Kross and Schreoder (1990) proposed that brokerage firm's characteristics influence analysts' earnings forecasting accuracy. In their analysis, sample analysts'

[1]Securities Regulation Committee, Department of Treasury, series 00588, volume #6, 1997

earnings forecasts from 1980 to 1981 were obtained from Value Line Investment, and the market value of a firm is used as the size of the firm. The results of this study on analysts' earnings forecasts did not find a positive relation between the company size and the analyst's forecast accuracy. Xu (1990) used the data range from 1986 to 1990 and the logarithm of average total revenue as a proxy of company earnings and examined factors associated with the accuracy of analysts' earnings forecast. The hypothesis that the larger the firm size is, the more accurate the analysts' earnings forecast would be was supported.

Su (1996) focuses on comparison of relative accuracy of management and analysts' earnings forecasts by using cross-sectional design method. Samples selection includes forecast data during the time period from 1991 to 1994. The company's "book value" of its total assets is used as a proxy for the size of the company in the regression analysis. The author believes that analysts are more attracted to larger companies, and there are more incentives for them to follow large companies than small companies in their forecasting. Therefore, the size of a company will affect the relative accuracy of analyst earnings forecast. On the other hand, large companies possess excessive human and financial resources and information which analysts have no access to, to allow managers to anticipate the corporate future earnings with high accuracy. The study results show that analyst and voluntary earnings forecast accuracy for larger companies are higher than forecast accuracy for small companies.

Yie (1996) examines the factors influencing management and financial analyst earnings forecasting accuracy. Data used in this study are the earnings forecasts during the years 1991–1995. She uses the company's total assets and market value of the company's equity as proxies for company size. The finding of this research reveals that the relative earnings forecast accuracy (management, voluntary management, mandatory management, and analyst) is not affected by the size of the company when the company's total assets are used as the proxy of company size. The result also indicates that mandatory management's earnings forecast and analysts' earnings forecasts are influenced by company size if market value of company's equity is used.

Xu (1990) examines the relative accuracy of analysts' earnings forecasts, a hypothesis that market volatility is one of factors that influence the relative accuracy of analyst earnings forecast. In upmarket situation, a vast amount of information regarding corporate earnings and overwhelming trading activities may hinder the analyst from getting realistic and objective information; thus, overoptimistic forecast might be a result. In contrast, when market is experiencing a downturn, individual investors are less speculative and more rational; thus, information about corporate earnings tends to be more accurate. Under these circumstances, the analyst tends to provide earnings forecasts with a higher level of accuracy. The results of this study support the author's hypothesis. Jiang (1993) examines the relative accuracy between management's earnings forecast and analyst earnings forecast. He hypothesizes that analysts' earnings forecast has a higher degree of accuracy in down market compared to upmarket situation. He uses samples forecast data from years 1991 to 1993 in his analysis and finds that the result of this research supports his argument.

Das et al. (1998) used a cross-sectional approach to study the optimistic behavior of financial analysts. Especially, they focused on the predicative accuracy of past information analysts' earnings forecast associated with magnetite of the bias in analysts' earnings forecasts. The sample selection covers the time period from 1989 to 1993 with 274 companies' earnings forecasts information. A regression method was used in this research. The term "optimistic behavior" is referred to as the optimistic earnings forecasts made by financial analysts. The authors hypothesize the following scenario: there is higher demand for nonpublic information for firms whose earnings are more difficult to predict than for firms whose earnings can be accurately forecasted using public information. Their finding supports the hypothesis that analysts will make more optimistic forecasts for low-predictability firms with an assumption that optimistic forecast facilitates access to management's nonpublic information. Clement (1999) studies the relation between the quality and forecast accuracy of analysts' reports. It also identifies systematic and time persistence in analysts' earnings forecast accuracy and examines the factors associated with degree of accuracy. Using the I/B/E/S database, the author has found that earnings forecast accuracy is positively related with analysts' experience (a surrogate for analyst ability and skill) and employer size (a surrogate for resources available) and inversely related with the number of firms and industries followed by the analyst. The sample selection covers the time horizon from 1983 to 1994 with earnings forecasts of 9,500 companies and 7,500 analysts. The author believes that as the analyst's experience increases, his earnings forecast accuracy will increase, which implies that the analyst has a better understanding of the idiosyncrasies of a particular firm's reporting practices or he might establish a better relationship with insiders and therefore gain better access to the managers' private information. An analyst's portfolio complexity is also believed to have association with his earnings forecast accuracy. He hypothesizes that forecast accuracy would decrease with the number of industries/firms followed. The effect of available resources impacts analyst's earnings forecast in such a way that analysts employed by a larger broker firm supply more accurate forecasts than smaller ones. The rationale behind this hypothesis is that the analyst hired by a large brokerage firm has better access to the private information of managers at the companies he follows. Large firms have more advanced networks that allow the firms to better disseminate their analyst's recommendations into the capital markets. The results of this research support the hypothesis made by the author.

Xiu (1992) studies the relative accuracy of management and analysts' earnings forecasts using Taiwan database covering the period 1986–1990. The management and analyst's earnings forecasts used in the study are from Business News, Finance News, Central News Paper, and The United Newspaper. The research methodology is to examine management's earnings forecast accuracy with prior and posterior analyst's earnings forecasts. The result reveals that a management's forecast is superior to prior analyst's forecast, but less accurate than posterior analyst's earnings forecasts.

Jiang (1993) examined the determinants associated with analysts' earnings forecast and management and analyst's earnings accuracy under different assumptions. A sample of Taiwan corporations is collected from the Four Seasons newspaper. Jiang uses cross-sectional regression analysis to investigate the relations between the forecast accuracy and firm's size, rate of earnings deviation, forecasting time horizon, market situation, and rate of annual trading volume turnover. His results show that earnings forecasts provided by analysts are more accurate than management earnings forecasts. Chia (1994) focuses a study on mandatory management earnings forecasts and the rate of trading volume turnover in an unpublished thesis.

9.3 Testable Hypotheses

9.3.1 Firm Size

The size of a firm is believed to have influence on the accuracy of analyst's and management's earnings forecast. Jaggi (1980), Bhushan (1989), and Clement (1999) found that the larger the company is, the more accurate the earnings forecast will be. They believe that holding other factors constant, larger companies have more financial and human resources available that allow the management to draw more precise earnings forecast than smaller companies. Thus, forecasts and recommendations supplied by larger firms are more valuable and accurate than the smaller firms:

H1: The accuracy of management's earnings forecast increases with the size of the firm.

H2: The accuracy of management's voluntary earnings forecast increases with the size of the firm.

H3: The accuracy of management's mandatory earnings forecast increases with the size of the firm.

H4: The accuracy of analysts' earnings forecast increases with the size of the firm.

9.3.2 Volatility of Market

The accuracy of earnings forecast will be affected by market situation. When market is very volatile and unstable, investors who are looking for the opportunities to profit will act more, speculative about what would be the next for the market. In this situation, it is more difficult for analysts to figure out the real useful information for their forecasts; they might have a tendency to overoptimistically forecast the earnings and provide recommendations. When a market is in a relative stable period, investors tend to be rational about the next movement of the market; there is less biased information regarding corporate earnings among the general investors; thus, the information accessed by analysts will allow them to be more objective in the earnings forecast. In contrast, the management has the insights on what is really happening in the aspects of operation, finance, top management changes, and profitability of the business. Even if they are less vulnerable regardless of what the market situation is, voluntary management's earnings forecast might be affected by market volatility to some extent:

H5: Management's earnings forecast will not be affected by volatility of market.

H6: Voluntary management's earnings forecast is a function of the of market volatility.

H7: Mandatory management's earnings forecast is not affected by market volatility.

H8: Accuracy of analysts' earnings forecast is affected by market volatility.

9.3.3 Volume Turnover

The relationship between trading volume turnover and accuracy of earnings forecast can be examined based on the hypothesis that daily stock trading volume represents the public investors' perception about a company. Larger trading volume during a day for a particulate stock reflects a higher degree of divergence on confidence about the company's stock, and vice versa. This public perception on a stock might distract management's and analysts' judgment; they need more time and strive extra efforts in order to prove accurate earnings forecasts:

H9: Trading volume turnover affects the accuracy of management's earnings forecast.

H10: Trading volume turnover affects the accuracy of voluntary management's earnings forecast.

H11: Trading volume turnover affects the accuracy of mandatory management's earnings forecast.

H12: Trading volume turnover affects the accuracy of analysts' earnings forecast.

9.3.4 Corporate Earnings Variance

Corporate earnings surprises are an important aspect of analysts' earnings forecast. The larger the earnings surprise is, the less useful the past information will be in earnings forecasting and the harder it is to make accurate forecasts. Corporate earnings variances represent the earnings surprises a company has in the past; it would affect the accuracy of management and analysts' earnings forecasts:

H13: Corporate earnings variances affect the accuracy of management's earnings forecast.

H14: Corporate earnings variances affect the accuracy of voluntary management's earnings forecast.

H15: Corporate earnings variances affect the accuracy of mandatory management's earnings forecast.
H16: Corporate earnings variances affect the accuracy of analysts' earnings forecast.

9.3.5 Type of Industry

There may exist a relationship between the type of industry and earnings forecast accuracy. They hypothesize that the difference between different industries may result in different levels of accuracy on earnings forecast. Some analysts may not possess adequate knowledge necessary in the forecasting in a particular industry; therefore, their forecast may not be as accurate as management's earnings forecast. Hence, the following hypotheses can be tested:
H17: Type of industries influences the accuracy of management's earnings forecast.
H18: Type of industries affects the accuracy of voluntary management's earnings forecast.
H19: Type of industries affects the accuracy of mandatory management's earnings forecast.
H20: Type of industries affects the accuracy of analysts' earnings forecast.

9.3.6 Forecasting Experience

Analysts' accuracy of earnings forecast will improve as their experience and knowledge about companies increase. They learn from their previous forecasts and make the next forecast more accurate. A similar argument can be made about the management's earnings forecast. Hence, the following hypotheses can be tested:
H21: Forecasting experience influences the accuracy of management's earnings forecast.
H22: Forecasting experience affects the accuracy of voluntary management's earnings forecast.
H23: Forecasting experience affects the accuracy of mandatory management's earnings forecast.
H24: Forecasting experience affects the accuracy of analysts' earnings forecast.

9.4 Empirical Results

9.4.1 Comparison of Management and Analyst's Earnings Forecast

To compare the relative accuracy of management and analysts' earnings forecasts, we focus on four major aspects regarding the relative accuracy of earnings forecasts. First, management versus analysts' earnings forecasts is made to compare the relative accuracy of the forecasts. Secondly, a comparison is made between voluntary management's forecasts and analysts' forecasts. Thirdly, mandatory management's forecasts are compared with analysts' forecasts to determine the relative accuracy of the forecasts. Finally, tests of hypothesis have been made to further prove the relative accuracy of management and analysts' earnings forecasts.

Table 9.1 provides descriptive statistics of management and financial analysts' earnings forecasts based from 1987 to 1999. It can be observed that absolute errors of management earnings forecasts are less than the analyst's in the years 1992, 1993, 1995, 1996, 1997, and 1997. It indicates that management's earnings forecasts are superior to analysts during that time period. But, in other time periods, the absolute errors for management's earnings forecast are higher than analysts', indicating analysts provide higher forecast productions. Overall, it is less obvious to conclude who has higher predicate ability for providing more precise earnings forecast.

The last three rows of Table 9.1 list the mean absolute errors of earnings forecasts by management and analysts during three different time periods. From 1988 to 1992, management's forecast absolute mean error is 1.796, whereas analyst's earnings forecast absolute mean error is 1.503. From 1993 to 1999, management's earnings forecast absolute mean error is 2.031, while analysts' forecast absolute mean error shows higher value of 2.236. If we look at the entire time period from 1987 to 1999, the absolute mean error for management's forecast is less than the absolute mean error for analysts' earnings forecast, which is 1.969 and 2.043 respectively. A conclusion can be drawn from the above results that management's earnings forecasts are more accurate than analysts' forecasts from the early 1990s, but less accurate during the late 1980s.

Table 9.2 shows the results of the Wilcoxon signed-rank test used to test the relative accuracy of managements' forecast and analysts' forecast. Comparing the negative ranks and positive ranks in Table 9.2, management's forecasts are less accurate than analysts' forecasts in the years 1987, 1988, 1989, 1990, and 1999, but more accurate in the years 1995 and 1998. There is no significant difference in the absolute errors between management's forecasts and analysts' forecasts.

If we examine the z-values of the test for the entire time period (1986–1999), the z-value for Wilcoxon signed-rank test is −0.346, which is not significant enough to tell the difference between the two samples. This supports the hypothesis H1 that there are no significant differences between management's forecast accuracy and analysts' forecast accuracy. This also agrees with the findings suggested by Imhoff and Pare (1982) and Baretley and Cameron (1991). They believe the reason for that is due to the similar abilities of forecasters and comparable networks to access company information (public/private) between management and analysts; it is possible that both can provide relative accurate earnings forecasts.

If the entire time period is divided into two subsamples, one is from 1987 to 1992 and the other is from 1993 to 1999, the latter subsample shows a significant level of 0.05 with a z-value of −2.138, which indicates that management's forecasts are less reliable than analysts' forecasts. For the former subsample, it shows no contradiction with the results by Imhoff and Pare (1982) and Baretley and Cameron (1991).

Table 9.1 Descriptive statistics of management and analyst's earnings forecast errors

		Management's forecast error				Analyst's forecast error			
Year	Sample size	Mean	Standard error	Maximum value	Minimum value	Mean	Standard error	Maximum value	Minimum value
1999	402	1.15	3.64	38.55	0.0006	1.07	3.18	35.76	0.0006
1998	360	1.61	5.80	63.48	0.0017	2.10	7.10	75.81	0.0037
1997	317	0.74	2.20	29.87	0.0040	0.78	3.17	53.23	0.0002
1996	267	1.12	5.41	55.45	0.0007	1.33	6.49	71.34	0.0009
1995	226	0.86	2.93	33.27	0.0001	0.87	2.31	28.42	0.0018
1994	178	0.62	2.17	18.8	0.0015	0.62	2.17	23.86	0.0025
1993	157	12.71	149.1	1869.1	0.0011	13.80	164.1	2057.4	0.0003
1992	126	0.65	1.48	11.58	0.0013	0.73	1.48	8.14	0.0005
1991	144	0.84	2.99	32.93	0.0014	0.65	1.10	5.67	0.0079
1990	120	5.98	32.60	334.94	0.0017	4.86	20.92	187.3	0.0075
1989	116	1.11	2.28	18.66	0.0009	0.91	1.59	10.56	0.0056
1988	96	1.25	3.29	19.57	0.0001	1.08	2.74	15.43	0.0006
1987	78	0.66	1.88	15.73	0.0171	0.57	1.78	14.62	0.0025
1993–1999	1,907	2.031	0.983	1869.1	0.0001	2.236	1.083	2057.	0.0002
1987–1993	680	1.796	0.535	334.93	0.0001	1.503	0.34	187.368	0.0005
1987–1999	2,587	1.969	37.57	1869.1	0.0001	2.043	40.88	2057.44	0.0002

Table 9.2 Wilcoxon signed-rank test for earnings forecast accuracy of management and analyst's forecasts errors

Year	Sample size	Negative ranks	Positive ranks	Ties	Z-value	Sig.
1999	402	150	230	22	−3.119	0.002***
1998	360	209	140	11	−5.009	0***
1997	317	147	140	30	−0.5	0.96
1996	267	143	119	5	−1.567	0.17
1995	226	117	89	20	−2.857	0.004***
1994	178	82	90	6	−0.339	0.734
1993	157	74	79	4	−0.502	0.616
1992	126	60	62	4	−1.136	0.256
1991	144	69	69	6	−0.407	0.684
1990	120	43	76	1	−2.941	0.003***
1989	116	46	67	3	−2.7	0.007***
1988	96	28	60	8	−3.503	0***
1987	78	27	51	0	−3.315	0.001***
1993–1999	1,299	626	607	66	−1.592	0.111
1987–1993	1,288	569	665	54	−2.138	0.033**
1987–1999	2,587	1,195	1,272	120	−0.346	0.73

Negative ranks: absolute error of management's earnings forecast < absolute error of analyst earnings forecast
Positive ranks: absolute error of management's earnings forecast > absolute error of analyst earnings forecast
Tie: absolute error of management's earnings forecast = absolute error of analyst earnings forecast
*Significant level = 0.1, ** significant level = 0.05, *** significant level = 0.01

9.4.2 Factors Influencing the Absolute Errors of Earnings Forecast

9.4.2.1 Firm Size

We argue that the management possesses the relative advantage of having private insights that the analyst cannot access, and that a larger company has stronger human and financial resources. Therefore, the management forecasts of corporate earnings are much more precise. On the other hand, a larger company tends to draw attentions and is more likely to attract and be followed by financial analysts; analysts' forecasts can be objective and accurate as well; however, the results from our research do not support this argument. Table 9.4 shows that the p-value of t-parameter for the size of a company (0.478) does not reach a significant level, indicating the size of a company is not associated with the accuracy of management's earnings forecast. This result does not support the hypothesis H1: management's earnings forecast accuracy increases with the size of company. Tables 9.5 and 9.6 show results of the regression analysis for two subsamples representing the time period of 1987–1992 and 1993–1999 to investigate the relationship between company size and accuracy of management's earnings forecast.

Table 9.3 Taiwan stock market volatility from 1987 to 1999

Year	Rm (%)	R_f (%)	$R_m - R_f$ (%)	Market volatility
1999	27	4	23	Upmarket
1998	−13	6	−19	Down market
1997	14	5	9	Upmarket
1996	39	5	34	Upmarket
1995	−24	5	−29	Down market
1994	−13	5	−18	Down market
1993	55	6	49	Upmarket
1992	−28	6	−34	Down market
1991	13	6	7	Upmarket
1990	−59	8	−67	Down market
1989	51	7	44	Upmarket
1988	124	4	120	Upmarket
1987	138	4	134	Upmarket

Market volatility measure, Rm, suggested by Pettengill et al. (1995), is the last month's market return minus the first month's market return divided by the first month's market return in a given year. R_f is the risk-free rate

9.4.2.2 Volatility of Market

In Table 9.4, the p-value of t-parameter in column 4 for the *volatility of market* of a company (0.075) does not reach a significant level, which implicates the accuracy of management's earnings forecast is not positively associated with the volatility of market. This result supports the hypothesis H1: management's earnings forecast accuracy will not change with market volatility. Further examining the two sub-tables of Tables 9.4, 9.5, and 9.6, p-value of parameter for the volatility of market of a company (0.310) indicates that management's earnings forecast accuracy will not change with market volatility during 1987–1999. But the t-parameter for the volatility of market is −2.569, indicating the management can provide accurate forecast during upmarket, but less accurate forecast during down market.

9.4.2.3 Trading Volume Turnover

The p-values of t-parameter in column 4 for the trading volume turnover of a company in all three tables do not reach a significant level. The regression analysis does not support the hypothesis H8 that trading volume turnover will affect management earnings forecast accuracy.

9.4.2.4 Corporate Earnings Variances

In Table 9.4, the p-values of t-parameter (2.74) in column 4 for the rate of earnings divination of a company is 0.01, which shows corporate earnings variances affect management's earnings forecast accuracy. Positive value of t-parameter means management earnings forecast accuracy decreases as corporate earnings variance

Table 9.4 Regression model for the absolute errors of management earnings forecasts dating from 1987 to 1999

Independent variable	Correlation coefficient	t-statistic	p-value of t-statistic
Intercept	−11.59	−0.52	0.604
Size	0.69	0.71	0.478
I_1: Market volatility	2.86	−1.55	0.122
TR: Rate of trading volume turnover	0.09	0.10	0.923
CV: Corporate earnings variances	2.74	3.50	0.00***
E: Forecasting experience	1.00	1.25	0.212
I_2: Cement	0.71	0.11	0.196
I_3: Food	−0.38	−0.07	0.943
I_4: Plastic	−1.13	−0.20	0.844
I_5: Textile	−0.87	−0.19	0.853
I_6: Electrical machinery	0.85	0.14	0.888
I_7: Electronic equipment and cable	−0.78	−0.13	0.895
I_8: Chemical	14.83	2.70	0.007***
I_9: Glass and ceramic	−1.36	−0.17	0.867
I_{10}: Paper manufacturing	−0.84	−0.11	0.912
I_{11}: Steel	−0.79	−0.14	0.891
I_{12}: Rubber	−0.46	−0.68	0.946
I_{13}: Auto	−2.74	−0.28	0.782
I_{14}: Electronics	−0.82	−0.17	0.865
I_{15}: Construction	−1.01	−0.18	0.856
I_{16}: Transportation	−0.53	−0.08	0.933
I_{17}: Travel	−0.76	−0.10	0.924
I_{18}: Insurance	−0.92	−0.14	0.886
I_{19}: Grocery	0.65	0.09	0.925
R-square	0.016		

I_2–I_{19}: dummy variables for industry
*Significant level = 0.10, ** significant level = 0.05, *** significant level = 0.01

increases. This supports the hypothesis H12 that corporate earnings variance has an effect on management earnings forecast accuracy. From the examination for Tables 9.5 (1993–1999) and 9.6 (1987–1992), management earnings forecast accuracy is affected by corporate earnings variances during the recent years (1993–1999).

9.4.2.5 Type of Industry

To determine whether and which industry will influence the forecast accuracy, 18 industries are selected and represented by a dummy variable I_j. From Table 9.4, I_8 is the only industry that has a significant level for the p-values of t-parameter of 14.73. According to our assumption, I_8 represents the chemical industry. Thus, we conclude that management's forecasts are reliable for most of the industries studied in this research, except for the chemical industry.

Table 9.5 Regression model for the absolute errors of management earnings forecasts dating from 1993 to 1999

Independent variable	Correlation coefficient	t-statistic	p-value of t-statistic
Intercept	−18.16	−0.59	0.552
Size	1.00	0.76	0.450
I_1: Market volatility	−2.54	−1.02	0.310
TR: Rate of trading volume turnover	0.04	0.03	0.976
CV: Corporate earnings variances	3.48	3.42	0.001***
E: Forecasting experience	1.26	1.23	0.218
I_2: Cement	1.94	0.21	0.836
I_3: Food	−0.96	−0.15	0.885
I_4: Plastic	−2.08	−0.27	0.789
I_5: Textile	−1.39	−0.23	0.815
I_6: Electrical machinery	1.20	0.14	0.875
I_7: Electronic equipment and cable	−1.42	−0.18	0.859
I_8: Chemical	20.36	2.84	0.005***
I_9: Glass and ceramic	−1.91	−0.17	0.854
I_{10}: Paper manufacturing	−2.22	−0.19	0.851
I_{11}: Steel	−1.13	−0.16	0.875
I_{12}: Rubber	−0.74	−0.08	0.935
I_{13}: Auto	−3.29	−0.26	0.798
I_{14}: Electronics	−3.01	−0.50	0.617
I_{15}: Construction	−1.55	−0.23	0.822
I_{16}: Transportation	−0.37	−0.05	0.964
I_{17}: Travel	−0.36	−0.03	0.974
I_{18}: Insurance	−1.99	−0.23	0.815
I_{19}: Grocery	−0.70	−0.08	0.939
R-square	0.016		

I_2–I_{19}: dummy variables for industry
*Significant level = 0.10, ** significant level = 0.05, *** significant level = 0.01

9.4.2.6 Forecasting Experience

The results of regression analyses for investigating the relationship of forecasting experience and management's earnings forecast accuracy are shown in Tables 9.4, 9.5, and 9.6. All three p-values of t-parameter in column 4 for forecasting experience indicate that management earnings forecast accuracy is not affected by previous forecasting experiences. This conclusion does not support the hypothesis H20 that forecasting experiences affect management's forecast accuracy. Test of other hypotheses indicates similar results.

9.5 Conclusions

The results of our research indicate that company size has no effect on any of the following: management forecast, voluntary management forecast, mandatory

Table 9.6 Regression model for the absolute errors of management earnings forecasts dating from 1978 to 1992

Independent variable	Correlation coefficient	t-statistic	p-value of t-statistic
Intercept	15.55	0.87	0.386
Size	−0.68	−0.87	0.383
I_1: Market volatility	−2.57	−1.78	0.075*
TR: Rate of trading volume turn over)	0.25	0.47	0.636
CV: Corporate earnings variances	−0.39	−0.59	0.555
E: Forecasting experience	0.57	0.76	0.450
I_2: Cement	−0.59	−0.11	0.914
I_3: Food	0.54	0.11	0.913
I_4: Plastic	−0.16	−0.03	0.974
I_5: Textile	1.17	0.26	0.799
I_6: Electrical machinery	−0.68	−0.12	0.905
I_7: Electronic equipment and cable	0.40	0.08	0.936
I_8: Chemical	2.05	0.41	0.680
I_9: Glass and ceramic	−0.40	−0.06	0.955
I_{10}: Paper manufacturing	2.05	0.37	0.709
I_{11}: Steel	0.56	0.10	0.923
I_{12}: Rubber	−0.14	−0.02	0.981
I_{13}: Auto	1.94	0.24	0.814
I_{14}: Electronics	7.98	1.64	0.102
I_{15}: Construction	0.04	0.01	0.994
I_{16}: Transportation	−0.36	−0.06	0.949
I_{17}: Travel	−0.96	−0.16	0.876
I_{18}: Insurance	1.63	0.29	0.770
I_{19}: Grocery	3.78	0.66	0.509
R-square	0.035		

I_2–I_{19}: dummy variables for industry
*Significant level = 0.10, ** significant level = 0.05, *** significant level = 0.01

management forecast, and analysts' forecast. This result agrees with Jaggi (1980) and Kross and Schreoder (1990) that analyst's earnings forecast accuracy is not related with the size of company, but differs from the results suggested by Bhushan (1989), Das et al. (1998), Clement (1999), Xu (1990), and Jiang (1993) that company size does influence the relative precision of management or analysts' earnings forecasts.

It can be seen that the relative accuracy of management's earnings forecast and analyst's earnings forecast is not affected by market situation across the entire range of sampled forecasts. There are some indications that forecasting accuracy is affected by market ups and downs. For instance, the relative accuracy of voluntary management's earnings forecast during the entire time period, accuracy of management's forecast, and analysts' earnings forecasts during the years 1978 through 1992 are more accurate when market is up and are less accurate during the down market. This result agrees with what Su has suggested – earnings forecast accuracy is affected by market volatility, but in different ways. We believe that due to the fact that more individual investors who are most likely to chase the market when it is up saturate the Taiwan stock market. They examine corporate earnings with more caution, so their expectations for companies in general are more realistic and rational. Therefore, overall earnings forecast accuracy is increased, and vice versa.

The results of this study reveal that relative accuracy of all four kinds of earnings forecasts is not the functions of trading volume turnover. This agrees with the results obtained by Chai, but it disagrees with the results of Jiang (1993). Results of regression analysis indicate that management's earnings forecast and analysts' forecasts are sensitive to the corporate earnings variances. This conclusion proves the hypothesis supports H13 through H16 and supports the theories of Kross and Schreoder (1990) and Jiang (1993). We postulate that corporate earnings variance of earnings surprises is an important indicator for a company's profitability and its earnings in the future. The management and analysts use past year's earnings surprises to forecast that future earnings, with an assumption of higher forecast inaccuracy, are a result of a high degree of earnings deviation. Therefore, they will need to exercise their highest ability in making earnings forecast more accurate. But we found that the higher the corporate earnings variance is, the lower the forecast accuracy will be for both the management and for analysts. Corporate earnings variances should not be used as an important indicator as how a company is operating, but it represents a complicated business environment it operates in. The higher the complicity of the business environment is, the less accurate the prediction/forecast will be.

Analyst earnings forecast and management's earnings forecast are biased for the chemical industry over the entire time period of sampled forecast; voluntary management's earnings forecasts for textile, electrical machinery, and paper manufacturing industries are inaccurate during 1993–1999. Mandatory management's earnings forecasts are very inaccurate for the food industry, textile industry, and travel industry in the time period of 1993–1999. This supports Kross and Schreoder (1990) who concluded that analyst's earnings forecast is affected by the type of industry he/she follows.

The results reveal that the relative accuracy of management's earnings forecast and analyst's earnings forecast do not respond to the differences of forecasters' previous experiences. But, the relative accuracy of mandatory management's earnings forecast for forecasters affect the entire time period and the subsampled voluntary management's forecast previous earnings forecast experiences. This conclusion agrees with Clement's (1999) finding.

We rationalize that forecast accuracy is positive related to the forecasting experiences as hypotheses H21 through H24 state. The results from this study indicate otherwise. The forecast accuracy of mandatory and voluntary management's earnings forecast has a negative relationship with previous forecasting experiences. We argue that it is because of (1) mis-quantifying variable as a proxy of forecasting experiences and (2) only using the past mandatory management's earnings forecasts as the base of future focusing without paying attention to ways on how to reduce the forecasting errors in those forecasts. Therefore, the more forecasting experiences the forecaster has, the less accurate the forecast will be.

Methodology Appendix

Sample Selection

This research uses cross-sectional design to examine the relative accuracy of management and analysts' earnings forecast. Due to the disclosure regulation in Taiwan, management's earnings forecast is classified as two categories: mandatory earnings forecast and voluntary earnings forecast[1]. Samples used are management's and analysts' earnings forecasts at all publicly traded companies during the time period of 1987–1999. The forecasts are compared with actual corporate earnings on an annual basis. Voluntary and mandatory management's forecast and analysts' forecast are then used to compare with actual corporate earnings to evaluate the effects of management motivation and behaviors on their earnings forecasts.

Management and analysts' pretax earnings forecast data are collected from "Taiwan Business & Finance News" during the time period of 1987–1999. Actual corporate earnings are collected from the Department of Education's "AREMOS" database. Only those firms were included in the samples whose stocks were traded on the Taiwan Stock Exchange before December 31, 1999. Also, forecasts made after accounting year and before announcement of earnings were excluded from the sample.

Management's earnings forecast and analysts' earnings forecast samples for this research are selected to cover the time period from 1987 to 1999. Available database, over the 13-year period, consisted of 5,594 management's earnings forecasts, in which 2,894 management forecasts are voluntary and 2,700 management' forecasts are mandatory. A total of 17,783 analysts' forecasts are in the database. The selected samples, presented in Table 9.7, consist of 2,941 management earnings forecasts, of which 2,046 are voluntary and 1,679 mandatory forecasts and 3,210 analysts' earnings forecasts. Table 9.7 shows that the average number of analysts' earnings forecasts is more than the number of management's earnings forecasts. A higher frequency of analysts' earnings forecasts is expected as an analyst may cover more than one firm. Most of management's earnings forecasts are made after 1991; it may be attributed by the amendment of "Regulation of Financial Report of Stock Issuer" imposed by the Taiwanese government in 1991. In the new regulation, a new section dealing with earnings forecast was added requiring company's management to disclose its earnings forecasts to the general public. Comparing the number of management voluntary forecasts and mandatory forecasts, the latter is about 1.5 times more than the former except during the years 1991–1993.

Table 9.7 Sample of earnings forecasts by management and analysts from Taiwan database selected for the study

Year	Management voluntary	Management mandatory	Management	Analysts
1999	407	236	319	479
1998	408	238	335	430
1997	384	227	288	376
1996	328	201	223	322
1995	281	193	218	279
1994	219	135	112	247
1993	172	137	84	225
1992	139	78	84	200
1991	156	154	16	175
1990	125	125	NA[a]	154
1989	123	123	NA	130
1988	112	112	NA	106
1987	87	87	NA	87
Total	2,941	2,046	1,679	3,210

[a]Mandatory forecast requirement was introduced in 1991

Variable Definition

Absolute Earnings Errors

The mean value of total corporate earnings before tax is used as proxy of earnings forecast. Using earnings before tax in the analysis will eliminate other factors, such as raising cash for capital investment, earnings retention for capital investment, and stock distribution from paid in capital that might impact the accuracy of the analysis. Absolute earnings (before tax) forecast error is used to compare the relative accuracy of management and analysts' earnings forecasts.

Management's forecasts errors are calculated as follows:

$$MF_{m,i,t} = \frac{1}{N} \times \sum_{j=1}^{n} FE_{m,i,j,t}$$

$$MF1_{m,i,t} = \frac{1}{N} \times \sum_{j=1}^{n} FE1_{m,i,j,t}$$

$$MF2_{m,i,t} = \frac{1}{N} \times \sum_{j=1}^{n} FE2_{m,i,j,t}$$

$$AFE_{m,i,t} = \left|\left(MF_{m,i,t} - AE_{i,t}\right)/AE_{i,t}\right|$$

$$AFE1_{m,i,t} = \left|\left(MF1_{m,i,t} - AE_{i,t}\right)/AE_{i,t}\right|$$

$$AFE2_{m,i,t} = \left|\left(MF2_{m,i,t} - AE_{i,t}\right)/AE_{i,t}\right|$$

where,

$MF_{m,i,t}$: mean management's pretax earnings forecast for company i at time t

$MF1_{m,i,t}$: mean management's voluntary pretax earnings forecast for company i at time t

$MF2_{m,i,t}$: mean management's mandatory pretax earnings forecast for company i at time t

$FE_{m,i,j,t}$: management's jth pretax earnings forecast for company i at time t

$FE1_{m,i,j,t}$: voluntary management's jth pretax earnings forecast for company i at time t

$FE2_{m,i,j,t}$: mandatory management's jth pretax earnings forecast for company i in year t

$AFE_{m,i,t}$: absolute error of management's pretax earnings forecast for company i at time t

$AFE1_{m,i,t}$: absolute error of voluntary management's pretax earnings forecast for company i at time t

$AFE2_{m,i,t}$: absolute error of mandatory management's pretax earnings forecast for company i at time t

$AE_{i,t}$: actual pretax EPS for company i at time t

Analysts forecast errors are calculated as follows:

$$MF_{f,i,t} = \frac{1}{n} \times \sum_{j=1}^{n} FE_{f,i,j,t}$$

$$AFE_{f,i,t} = \left|\left(MF_{f,i,t} - AE_{i,t}\right)/AE_{i,t}\right|$$

where

$MF_{f,i,t}$: mean analysts' pretax earnings forecast for company at time t

$FE_{f,i,j,t}$: analysts' jth pre-tax earnings forecast for company i at time t;

$AFE_{f,i,t}$: absolute error of analyst's pre-tax earnings forecast for company i at time t;

$AE_{i,t}$: actual pre-tax EPS for company i at time t.

Company Size

Unlike other previous researchers who used market value of company's equity as a indication of size of a company, in this study, a company's last year's total revenue is used as the size of the company. The reason is because Taiwanese market is not efficient and investors are not informed fully with information they need during their investment decision-making process; speculations among individual investors are the main cause of the stock market volatility; thus, market value of company's equity cannot fully represent a company's real size.In order to better control the company size for our regression analysis, a logarithm of company's last year's total revenue is used as the following:

$$\mathrm{SIZE}_{i,t} = \ln\left(\mathrm{TA}_{i,t-1}\right)$$

where

$\mathrm{SIZE}_{i,t}$: the size of company i at time t;

$\mathrm{TA}_{i,t-1}$: total revenue of company i at time t−1.

Market Volatility

This study adapts what Pettengill, Glenn N, Sundaram, Sridhar, and Mathur and Ike used in their research to measure market volatility. Market volatility is measured as upmarket or down market by using market-adjusted return. This return is calculated as $R_m - R_f$, in which R_m is the last month's market return minus the last first month's market return divided by the first month's market return in a given year. R_f is the risk-free interest rate in the same year:

$$\mathrm{Return}_{(\mathrm{Market-adjusted})} = R_m - R_f$$

where

Upmarket if Return (Market-adjusted) >0

Down market if Return (Market-adjusted) <0

A dummy variable is used to identify market volatility. Market volatility is set to 1 if a year's Return $_{(\mathrm{Market-adjusted})}$ is greater than 0 and set to 0 otherwise. Table 9.3 reports the market volatility of Taiwan market.

Trading Volume Turnover

Trading volume turnover is defined as the value of a company's stock daily trading volume divided by the company's number of shares outstanding. To make this proxy better fit in the regression analysis, a logarithm is applied to the value and multiplies by 1,000:

$$AV_{i,t} = \frac{1}{N}\sum_{j=1}^{n} V_{i,j,t}$$

$$TR_{i,t} = \ln\left(1,000 \times \frac{AV_{i,t-1}}{CS_{i,t-1}}\right)$$

where

$V_{i,j,t}$: daily trading volume in day j at time t for company i
$AV_{i,t}$: mean daily trading volume at time t for company i
$CS_{i,t-1}$: number of shares outstanding at time t−1 for company i
$TR_{i,t}$: rate of trading volume turnover at time t for company i

Corporate Earnings Variance

In this research, we only consider the past 3 years' historical earnings surprises as a proxy of a company's corporate earnings variances. Thus, the corporate earnings variance is defined as the following:

$$CV_{i,t} = LN\left(\frac{\sigma(X)}{|\overline{X}|}\right)$$

$$\sigma(X) = \sqrt{\frac{\sum_{t=1}^{n}\left(X_t - \overline{X}\right)^2}{n-1}}$$

$$\overline{X} = \frac{1}{3}\sum_{t=1}^{3} Y_t$$

where

$CV_{i,t}$: corporate earnings variance at time t for company i
$\sigma(x)$: actual corporate earnings variance for company i
X_t: actual earnings at time t for company i
$\overline{X}$: mean EPS (before tax) for company i
Y_t: actual EPS at time t for company i

Type of Industry

There are two major ways to classify industries:

(i) "Industry classification of Republic of China" by State Council in 1987
(ii) Industry classification used by stock exchange house

In this research, we use the latter one to classify industries, and a variable I_j is set to represent nineteen different industries: cement, food, plastic, textile, electrical machinery, electronic equipment and cable, chemical, glass and ceramic, paper manufacturing, steel, rubber, auto, electronics, construction, transportation, travel, insurance, and others.

Proxy for Experience

According to research done by previous researchers, although earnings forecast accuracy is positively related to management and analysts' previous forecasting experiences, it is difficult to quantify the experiences. In this research, we argue that the accuracy of nth management and analyst's earnings forecast depends on their (n − 1)th forecasting experience. Therefore, the proxy of experience is defined as the following:

$$E_{m,i,t} = \sum_{j=1}^{n} E_{m,i,j,t-1}$$

$$E1_{m,i,t} = \sum_{j=1}^{n} E1_{m,i,j,t-1}$$

$$E2_{m,i,t} = \sum_{j=1}^{n} E2_{m,i,j,t-1}$$

$$E_{f,i,t} = \sum_{j=1}^{n} E_{f,i,j,t-1}$$

where

$E_{m,i,t}$: total number of times of management's earnings forecasting experience at time t for company i

$E1_{m,i,t}$: total number of times of voluntary management's earnings forecasting experience at time t for company i

$E2_{m,i,t}$: total number of times of mandatory management's earnings forecasting experience at time t for company i

$E_{f,i,t}$: total number of times of analysts' earnings forecasting experience at time t for company i

Regression Model

A multiple regression model is used to examine the effect of six factors: firm size, market volatility, trading volume turnover, corporate earnings variances, type of industry, and experience.

Regression model for management's forecast absolute error percentage:

$$\begin{aligned}\mathrm{AFE}_{m,i} = {} & a_0 + a_1(\mathrm{SIZE}_i) + a_2 I_{i,1} \\ & + a_3(\mathrm{TR}_i) + a_4(\mathrm{CV}_i) + a_5\left(E_{m,i}\right) \\ & + \sum_{k=6}^{23} a_k I_{i,k-4} + \varepsilon_i\end{aligned} \tag{9.1}$$

Regression model for voluntary management's forecast absolute error percentage:

$$\begin{aligned}\mathrm{AFE1}_{m,i} = {} & b_0 + b_1(\mathrm{SIZE}_i) + b_2 I_{i,1} \\ & + b_3(\mathrm{TR}_i) + b_4(\mathrm{CV}_i) + b_5\left(E1_{m,i}\right) \\ & + \sum_{k=6}^{23} b_k I_{i,k-4} + \varepsilon_i\end{aligned} \tag{9.2}$$

Regression model for mandatory management's forecast absolute error percentage:

$$\begin{aligned}\mathrm{AFE2}_{m,i} = {} & c_0 + c_1(\mathrm{SIZE}_i) + c_2 I_{i,1} \\ & + c_3(\mathrm{TR}_i) + c_4(\mathrm{CV}_i) + c_5\left(E2_{m,i}\right) \\ & + \sum_{k=6}^{23} c_k I_{i,k-4} + \varepsilon_i\end{aligned} \tag{9.3}$$

Regression model for analysts' forecast absolute error percentage:

$$\begin{aligned}\mathrm{AFE}_{f,i} = {} & d_0 + d_1(\mathrm{SIZE}_i) + d_2 I_{i,1} \\ & + d_3(\mathrm{TR}_i) + d_4(\mathrm{CV}_i) + d_5\left(E_{f,i}\right) \\ & + \sum_{k=6}^{23} d_k I_{i,k-4} + \varepsilon_i\end{aligned} \tag{9.4}$$

where

$\mathrm{AFE}_{m,i}$: absolute error percentage of management's forecast for company i

$\mathrm{AFE1}_{m,i}$: absolute error percentage of voluntary management's forecast for company i

$\mathrm{AFE2}_{m,i}$: absolute error percentage of mandatory management's forecast for company i

$\mathrm{AFE}_{f,i}$: absolute error percentage of analysts' forecast for company i

SIZE_i: size of company i

$I_{i,1}$: market volatility (1 if market is upmarket, 0 if market is down market)

TR_i: rate of trading volume turn over for company i
CV_i: corporate earnings variances for company i
$E_{m,i}$: management's earnings forecasting experience for company i
$E_{1m,i}$: voluntary management's earnings forecasting experience for company i
$E2_{m,i}$: mandatory management's earnings forecasting experience for company i
$E_{f,i}$: analyst earnings forecasting experience for company i
$I_{i,2-19}$: type of industry for company i

Wilcoxon Two-Sample Test

If the two-sample groups are related, Wilcoxon two-sample test will be used to determine the relative earnings forecast accuracy:

$$Z = \frac{W - E(W)}{\sqrt{V(W)}}$$

where

W1: rank sum of absolute error percentage for management's earnings forecasts
W2: rank sum of absolute error percentage for analysts' earnings forecasts
W: smaller value between W1 and W2
E(W): expected values of W-distribution
$\sqrt{V(W)}$: deviation of W-distribution

References

Bartley, J. W., & Cameron, A. B. (1991). Long-run earnings forecasts by managers and financial analysts. *Journal of Business Finance & Accounting, 18*, 21–41.

Bhushan, R. (1989). Firm characteristics and analyst following. *Journal of Accounting & Economics, 11*, 255–274.

Chia, Yuxuan (1994). *Examination of mandatory management earnings forecasts and the rate of trading turnover*. Unpublished thesis, Accounting Department, National Chengchi University.

Clement, B. M. (1999). Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter? *Journal of Accounting & Economics, 27*, 285–303.

Das, S., Carolyn, B., Levine, K. S. (1998). Earnings predictability and bias in analysts' earnings forecasts. *Accounting Review, 73*, 277–294.

Imhoff, E. A., & Pare, P. (1982). Analysis and comparison of earnings forecasts agents. *Journal of Accounting Research, 20*, 429–439.

Jaggi, B. (1980). Further evidence on the accuracy of management forecasts vis-à-vis analyst' forecasts. *The Accounting Review, 55*, 96–101.

Jiang, Jianquan (1993). *Factors associated with analyst's earnings forecast accuracy under different assumptions in Taiwan*. Unpublished thesis, School of Business, Taiwan University.

Kross, W., Ro, B., & Schreoder, D. (1990). Earnings expectations: The analysts information advantage. *The Accounting Review, 65*, 461–476.

Pettengill, G. N., Sundaram, S., & Mathur, I. (1995). The conditional relation between beta and returns, *Journal of Financial and Quantitative Analysis, 30*(1), 101–116

Su, Yongru (1996). *Comparison of relative accuracy of management and analysts earnings forecast*. Unpublished thesis, Accounting Department, Soochow University.

Xu, Xiubin (1990). *A examination of analysts earnings forecasts accuracy*. Unpublished thesis, Accounting Department, National Chengchi University.

Xu, Jinjuan (1992). *The effect of management earnings forecast on portfolio management*. Unpublished thesis, Accounting Department, National Chengchi University, Taiwan.

Yie, Meiling (1996). *Examination on factors in management and financial analyst earnings forecasting accuracy*. Unpublished thesis, Accounting Department, National Chung Cheng University.

Market-Based Accounting Research (MBAR) Models: A Test of ARIMAX Modeling

10

Anastasia Maggina

Contents

Abstract

The purpose of this study is to provide evidence drawn from publicly traded companies in Greece as far as some of the standard models of accounting earnings and returns relations mainly collected through the literature. Standard models such as earnings level and earnings changes have been investigated in this study. Models that fit better to the data drawn from companies listed on the Athens Stock Exchange have been selected employing autoregressive integrated moving average with exogenous variables (ARIMAX) models. Models I (price on earnings model), II (returns on change in earnings divided by beginning-of-period price and prior period), V (returns on change in earnings over opening market value), VII (returns deflated by lag of 2 years on earnings over opening market value), and IX (differenced-price model) have statistically significant coefficients of explanatory variables. In addition, model II (returns on change in earnings divided by beginning-of-period price and prior period with MSE (minimum squared error)

A. Maggina
Business Consultant/Research Scientist, Avlona, Attikis, Greece
e-mail: anastasiamaggina@yahoo.gr; a.maggina@yahoo.com

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_10,

loss function in ARIMAX (2,0,2)) is prevalent. These models take place with backward-looking information instead of forward-looking information that recent literature is assessed. Application of generalized autoregressive conditional heteroscedasticity (GARCH) models is suggested for further future research.

Keywords

Market-based accounting research • Price on earnings model • Earnings level • Earnings change • Return models • Autoregressive-moving average with exogenous variables • Minimum value of squared residuals (MSE loss function) • Unit root tests • Stationarity • Dickey-Fuller test

10.1 Introduction

The relationship between accounting earnings and stock prices has been discussed both in the accounting and financial literature. In a consideration of market-based accounting research, the association between returns and earnings has been very low due to (i) poor specification of the estimating equation, (ii) poor informational properties of reported earnings, (iii) inappropriate choice of the assumed proxy for expected earnings, and (iv) the availability of more timely sources of the value-relevant information in earnings statement (Strong and Walker 1993). This has been resolved by allowing for time-series and cross-sectional variation in the regression parameters, by including an earnings yield and partitioning all-inclusive earnings into pre-exceptional, exceptional, and extraordinary components. Yet, empirical evidence on returns-earnings association in certain applications is very strong (Easton and Harris 1991; Kothari and Zimmerman 1995). Standard empirical models, that is, price-earnings association, and returns-earnings association (earnings level and earnings changes) have been investigated in the USA and UK with no empirical evidence from at least the rest of Europe. The purpose of this paper is to develop an empirical background by investigating whether some of the standard models that have been collected through the literature would be relevant for evaluating accounting earnings/returns associations in a stock market being in transition from an emerging to a matured one. In other words, we purport to select those models that better fit to available data. The rationale of the Greek stock market response is one major objective, while the other is to make some comparisons with other findings presented in the literature.

The Athens stock exchange has been established on 1876. It was an emerging market some years ago and has run up to a mature one. It is a normal market with no surprises, and investments are not only a place for corporations to get financing; more importantly it is a place to create wealth. Registered companies are young, growth-oriented to long-established enterprises. The Greek stock market has played a great role in the economic development of the country in the last half of the twentieth century. Facts that have influenced the ASE is the inclusion of the country in the Economic and Monetary Union and the **crash** of 1999 that affected the life of many Greek families. The ASE operates in a country which has been in deep recession since 2008.

The paper is organized as follows: Literature is discussed in Sect. 10.2. Methodology and model building process are described in Sect. 10.3. Section 10.4 discusses the sample design. Section 10.5 presents empirical findings. Conclusions are summarized in last Sect. 10.6.

10.2 Review of the Literature

Common interest of both accounting and finance scholars and users of financial statements (primarily, financial analysts and investors) is to obtain concrete and increased knowledge of the association between accounting earnings and stock prices for a more or less predictive ability explanation. An evaluation of the informational content of, at least, the most basic accounting numbers contributes to the improvement of models. Among the various treatments of the relation between accounting earnings and stock prices, the seminal work of Ball and Brown (1968) is precedent. Their work indicates that accounting earnings and some of its components exhibit an information content drawn through stock prices. Later on, Brown et al. (1985) worked on the relative ability of the current innovation in annual earnings and the revision in next year's earnings forecasts to explain changes in stock prices. However, it has empirically been tested (Fama et al. 1969; Fama 1970). Successive changes in individual common stocks are very nearly independent because of an efficient market in which adjustments to new information are rapidly made (Mandelbrot 1966; Samuelson 1965). Beaver (1970) and Rosenberg and Marathe (1975) worked to discover financial statement components that are related to risk and, thus, predict expected stock returns. They maintain that financial statements constitute an assured source of information (and can be prepared in a short time period) and, consequently, the direction of 1-year-ahead earnings changes affects stock prices. Their results indicate that the summary measure robustly predicts future stock returns.

Beaver et al. (1980) developed a model known as BLM theory. They followed the traditional price-earnings relation and tested for the information content of prices with respect to future earnings. BLM regress percentage change in price which may contain information about future earnings not reflected in the current earnings. Lev and Ohlson (1982) describe price and return models as complementary. Given that stock prices are associated with accounting earnings, analysts forecast the reported accounting numbers, and as Brown et al. (1985) maintain, analysts use their 1-year-ahead forecasts to convey their expectations about permanent earnings. In all the above treatments of the subject matter referring to accounting earnings and stock prices, the more and more prevailing role of managerialism could be more or less emphatic. Trueman (1986) explained why managers would be willing to disclose forecasts of higher or lower than expected earnings. Undoubtedly, the disclosure of favorable accounting numbers or the disclosure of forecasts of favorable expected earnings contributes to an increase of the market value of the firm. Hirst et al. (2008) assert that managers often issue earnings forecasts to correct information asymmetry problems and, thus, influence their firm's stock price.

Lipe and Kormendi (1987) found that stock returns were considered as a function of the revisions in expectations of earnings. They show that the stock return reaction to earnings is a function of (1) the time-series properties of earnings, (2) the interest rate used to discount expected future earnings, and (3) the relative ability of earnings. Furthermore, Freeman (1987) maintains that the relation between security prices and firm-specific information is associated with the market value of a firm's common equity and concludes that the magnitude of the abnormal returns related to earnings is a decreasing function of firm size. Of much interest is a hypothesis tested by Freeman (1987) which implies that (i) the abnormal security returns related to accounting earnings occur earlier for large than for small firms and (ii) abnormal returns are lower for large firms and higher for small firms. The information content theorists consider size as an important conditioning variable when testing the information content of prices with respect to future earnings and contemporaneous price changes.

Beaver et al. (1987), Collins et al. (1987), Collins and Kothari (1989), and many others show that unexpected earnings for a year are correlated with returns from a prior year. Christie (1987) concludes that while return and price models are economically equivalent, return models are econometrically less problematic. Meanwhile, Holthausen and Verrecchia (1988) worked on price changes when information is announced. Landsman and Magliolo (1988) argue that price models are superior to return models for certain applications. According to Cornell and Landsman (1989), stock prices respond to earnings announcements. Alternatively stated, unexpected increases in earnings are associated with a rise in stock prices, and unexpected decreases in earnings are associated with a fall in stock prices. Relating unexpected accounting earnings and security prices aims to assess the information content of the latter (Collins and Kothari 1989). Conservative proponents support the view that financial statement ratios are the basic tools for evaluating and predicting accounting earnings and, consequently, security prices. Ou and Penman (1989) maintain that the relationship of financial statement characteristics to value is not apparent. Easton and Zmijewski (1989) and that of Board and Walker (1990) analyze a coefficient that measures the response of stock prices to accounting earnings coefficient. They measure the response of stock prices to accounting earnings announcements, and they empirically show that the higher this coefficient, the smaller the stock price changes.

Lipe (1990) worked on the relation between stock returns and accounting earnings, assuming that market observes current-period information other than earnings. In the process of relating stock prices to accounting earnings, (i) a coefficient which measures the stock-return response to a one dollar earnings changes as a function of both "predictability" and "persistence" of earnings, and (ii) the variance of stock price changes during the reporting of earnings has been tested in the literature (Cho and Jung 1991; etc.). Easton and Harris (1991) presented models relating earnings variables and security returns, concluding that both current earnings level and the earnings change variables play a role in the security valuation. They are in fact correlated. Strong and Walker (1993) used a panel regression approach to examine the association between annual stock price returns and reported earnings figures of industrial companies in the UK and proved through an analysis of earnings components that models of the relation between earnings and returns that focus exclusively on the deflated first difference of earnings are misspecified. Kahya et al. (1993) showed that earnings growth rates are predictable using past earnings growth rates, stock price returns are predictable using past earnings growth rates as well as stock price returns, and firm size has no incremental explanatory power with respect to equity prices.

Kothari and Zimmerman (1995) provide empirical results confirming that price models' earnings response coefficients are less biased and argue that the price specification suffers more from heteroscedasticity/misspecification problems than the return model. They conclude that in some research contexts, the combined use of both price and return models may be useful. Chen et al. (2001) obtain evidence of value relevance of accounting information in China based on a return and a price model according to both the pooled cross section and time-series regressions or the year-by-year regressions, and find that value relevance is higher for companies using only A shares to domestic investors despite of the lack of alternative information sources (i.e., earnings forecasts, financial analysts), the lack of a sufficient level of corporate governance, and the even recent new phenomenon of independent auditing in China. Lundholm and Myers (2002) examined how a firm's disclosures affect the mix of earnings information reflected in its annual stock return and found that increased disclosure activity "brings the future forward" into current stock returns. Chen et al. (2011) examined the properties of accounting numbers of listed firms in China by investigating the interplay between accounting earnings and stock prices. They found that core earnings (or operating income) have a greater association with contemporaneous stock returns than nonoperating revenues and expenses, and they also found that different types of firm ownership may have different impacts on the information content of earnings components.

10.3 Methodology and Models Used

Literature concerning accounting earnings and stock prices has been formulated in the framework of an earnings-based valuation model expanded with the dividend irrelevance proposition. These standard models which have also been tested by Easton and Harris (1991) and Zimmerman and Kothari (1995) are as follows:

I. Price on earnings

$$P_{i,j} = a + bA_{i,j} + e_{i,j}. \tag{10.1}$$

II. Returns on change in earnings divided by beginning-of-period price and prior period earnings divided by the price at the beginning of the return period

$$\left[\left(A_{i,j}/P_{i,j-1}\right)\right] = a_{i,j} + b_{i,j}\left(A_{i,j} - A_{i,j-1}\right)/P_{i,j-1} + \left(A_{i,j-1}/P_{i,j-1}\right)\right] \tag{10.2}$$

III. Returns on earnings over opening market value

$$\left[\left(P_{i,j} - P_{i,j-1}\right) + d_{i,j}\right]/P_{i,j-1} = a_{i,j} + b_{i,j}\left(A_{i,j}/P_{i,j-1}\right) + e_{i,j} \tag{10.3}$$

IV. Returns on prior earnings model over opening market value

$$\left[\left(P_{i,j} - P_{i,j-1}\right) + d_{i,j}\right]/P_{i,j-1} = a_{i,j} + b_{i,j}\left(A_{i,j-1}/P_{i,j-1}\right) + e_{i,j} \tag{10.4}$$

V. Returns on change in earnings over opening market value

$$\left[\left(P_{i,j} - P_{i,j-1}\right) + d_{i,j}\right]/P_{i,j-1} = a_{i,j} + b_{i,j}\left[\left(A_{i,j} - A_{i,j-1}\right)/P_{i,j-1}\right] + e_{i,j} \tag{10.5}$$

VI. Returns on change in earnings over opening market value and on earnings over opening market value

$$\begin{aligned}\left[\left(P_{i,j} - P_{i,j-1}\right) + d_{i,j}\right]/P_{i,j-1} = a &+ b_{1i,j}\left[\left(A_{i,j} - A_{i,j-1}\right)/P_{i,j-1}\right] \\ &+ b_{2i,j}\left[\left(A_{i,j}/P_{i,j-1}\right) + e_{i,j}\right]\end{aligned} \tag{10.6}$$

VII. Returns (deflated by lag of 2 years) on earnings over opening market value

$$\left[\left(P_{i,j} - P_{i,j-1}\right) + d_{i,j}\right]/P_{i,j-2} = a_{i,j} + b_{i,j}\left(A_{i,j}/P_{i,j-1}\right) + e_{i,j} \tag{10.7}$$

VIII. Return model regressed on earnings over opening market value

$$P_{i,j}/P_{i,j-1} = a + b_{1i,j}A_{i,j}/P_{i,j-1} \tag{10.8}$$

IX. Differenced-price model

$$P_{i,j} - P_{i,j-1} = A_{i,j} - A_{i,j-1} \tag{10.9}$$

where
$P_{i,j}$ = stock price (per share) of firm i in period j
$A_{i,j}$ = earnings per share of firm i in period j
$d_{i,j}$ = dividend per share of firm i in period j
a = a constant in a linear relationship(intercept parameter)
b_1,b_2 = a slope parameter or a coefficient in a linear regression
i = cross-selection item, j = time-series item

To be familiar and consistent with the existing literature, some requirements are stressed. For example, earnings per share divided by price at the beginning of the return period ($A_{i,j}/P_{i,j-1}$) refers to current earnings level variable. Change in earnings divided by beginning-of-period price refers to earnings change variable [($A_{i,j} - A_{i,j-1}$)/$P_{i,j-1}$]. Thus far, the models that have been selected to be tested express the following:

Model I: Expresses price as a multiple of earnings.

Model II: Expresses historical price-earnings ratios with an earnings change versus earnings levels explanation form.

Model III: From an earnings valuation perspective, earnings levels(divided by beginning-of-period price) will be associated with returns.
Model IV: Illustrates a linear relation between prior earnings divided by beginning-of-period price and security returns over that period.
Model V: Regresses annual security returns on change in earnings divided by beginning-of-period price.
Model VI: Expresses the contribution of change in earnings versus earnings levels in the explanation of stock returns.
Model VII: Returns (deflated by lag of 2 years) on earnings over opening market value.
Model VIII: Expresses returns on earnings level divided by beginning-of-period price.
Model IX: Expresses differenced-price model.

As in Easton and Harris (1991), the models under investigation have been based on either the book value valuation model or the earnings valuation model.

The book value valuation model indicates that

$$P_{ij} = BV_{ij} + u_{ij} \tag{10.10}$$

Taking first differences we have

$$\Delta P_{ij} = \Delta BV_{ij} + u_{ij} \tag{10.11}$$

But in general

$$\Delta BV_{ij} = A_{ij} - d_{ij} \tag{10.12}$$

Substituting (3) into (2), rearranging and dividing by P_{ij-1} yields
$(\Delta P_{ij} + d_{ij})/P_{ij-1} = A_{ij}/P_{ij-1} + u_{ij}$ (model III)
On the other hand,

$$P_{ij} = \rho A_{ij} + u_{ij} \tag{10.13}$$

Given the dividends irrelevance proposition, we have

$$P_{ij} + d_{ij} = \rho A_{ij} + u_{ij} \tag{10.14}$$

It follows that $(\Delta P_{ij} + d_{ij})/P_{ij-1} = \rho(\Delta A_{ij}/P_{ij-1}) + u_{ij}$ (model V)

ARIMAX (autoregressive integrated moving average with exogenous variables) as a suitable technique for nonstationarity time-series modeling is employed in this study.

10.4 Sample Selection

The whole population containing all Greek-listed companies in the Athens Stock Exchange is investigated in this study. The total number of companies amounts to 513 companies. The main source of data is the Athens Stock

Exchange Annual Yearbook, the annual statistical bulletin, and the Internet. Total number of companies refers to the time period 1974 up to 2005 (the most recently available data when writing the paper). The full sample (1974–2005) is separated in two samples (1974–1999 and 2000–2005) which correspond to two-time periods, that is, before the Euro currency and the Euro era. As in Kothari and Zimmerman (1995), to avoid any undue influence of extreme observations, the largest and the lowest 1 % of observations is excluded from the sample. EPS take positive or zero values. All firms have a December fiscal year-end. Annual earnings include those from discontinued operations and extraordinary items.

10.5 Empirical Findings

For each year, a nonconstant number of companies is available. We define the total number of companies at each year t as n_t. We would like to estimate the following nine models for the period from 1974 to 2005:
Model I

$$P_t = a + bA_t + e_t, \quad \textbf{for} \quad e_t \sim N\left(0, \sigma^2\right), \tag{10.15}$$

where $P_t = {n_t}^{-1}\sum_{i=1}^{n_t} P_{i,t}$ and $A_t = {n_t}^{-1}\sum_{i=1}^{n_t} A_{i,t}$.

Model II

$$\frac{A_t}{P_{t-1}} = a + b_1\frac{A_t - A_{t-1}}{P_{t-1}} + b_2\frac{A_{t-1}}{P_{t-1}} + e_t, \quad \textbf{for} \quad e_t \sim N\left(0, \sigma^2\right), \tag{10.16}$$

where $\frac{A_t}{P_{t-1}} = {n_t}^{-1}\sum_{i=1}^{n_t}\frac{A_{i,t}}{P_{i,t-1}}$, $\frac{A_t - A_{t-1}}{P_{t-1}} = {n_t}^{-1}\sum_{i=1}^{n_t}\frac{A_{i,t} - A_{i,t-1}}{P_{i,t-1}}$, and $\frac{A_{t-1}}{P_{t-1}} = {n_t}^{-1}\sum_{i=1}^{n_t}\frac{A_{i,t-1}}{P_{i,t-1}}$.

Model III

$$\frac{P_t - P_{t-1} + d_t}{P_{t-1}} = a + b_1\frac{A_t}{P_{t-1}} + e_t, \quad \textbf{for} \quad e_t \sim N\left(0, \sigma^2\right), \tag{10.17}$$

where $\frac{P_t - P_{t-1} + d_t}{P_{t-1}} = {n_t}^{-1}\sum_{i=1}^{n_t}\frac{P_{i,t} - P_{i,t-1} + d_{i,t}}{P_{i,t-1}}$ and $\frac{A_t}{P_{t-1}} = {n_t}^{-1}\sum_{i=1}^{n_t}\frac{A_{i,t}}{P_{i,t-1}}$.

Model IV

$$\frac{P_t - P_{t-1} + d_t}{P_{t-1}} = a + b_1 \frac{A_{t-1}}{P_{t-1}} + e_t, \ \ \textbf{for} \ \ e_t \sim N\left(0, \sigma^2\right), \tag{10.18}$$

where $\frac{P_t - P_{t-1} + d_t}{P_{t-1}} = {n_t}^{-1} \sum_{i=1}^{n_t} \frac{P_{i,t} - P_{i,t-1} + d_{i,t}}{P_{i,t-1}}$ and $\frac{A_{t-1}}{P_{t-1}} = {n_t}^{-1} \sum_{i=1}^{n_t} \frac{A_{i,t-1}}{P_{i,t-1}}$.

Model V

$$\frac{P_t - P_{t-1} + d_t}{P_{t-1}} = a + b_1 \frac{A_t - A_{t-1}}{P_{t-1}} + e_t, \ \ \textbf{for} \ \ e_t \sim N\left(0, \sigma^2\right), \tag{10.19}$$

where $\frac{P_t - P_{t-1} + d_t}{P_{t-1}} = {n_t}^{-1} \sum_{i=1}^{n_t} \frac{P_{i,t} - P_{i,t-1} + d_{i,t}}{P_{i,t-1}}$ and $\frac{A_t - A_{t-1}}{P_{t-1}} = {n_t}^{-1} \sum_{i=1}^{n_t} \frac{A_{i,t} - A_{i,t-1}}{P_{i,t-1}}$.

Model VI

$$\frac{P_t - P_{t-1} + d_t}{P_{t-1}} = a + b_1 \frac{A_t - A_{t-1}}{P_{t-1}} + b_2 \frac{A_t}{P_{t-1}} + e_t, \ \ \textbf{for} \ \ e_t \sim N\left(0, \sigma^2\right), \tag{10.20}$$

where $\frac{P_t - P_{t-1} + d_t}{P_{t-1}} = {n_t}^{-1} \sum_{i=1}^{n_t} \frac{P_{i,t} - P_{i,t-1} + d_{i,t}}{P_{i,t-1}}$, $\frac{A_t - A_{t-1}}{P_{t-1}} = {n_t}^{-1} \sum_{i=1}^{n_t} \frac{A_{i,t} - A_{i,t-1}}{P_{i,t-1}}$, and $\frac{A_t}{P_{t-1}} = {n_t}^{-1} \sum_{i=1}^{n_t} \frac{A_{i,t}}{P_{i,t-1}}$.

Model VII

$$\frac{P_t - P_{t-1} + d_t}{P_{t-2}} = a + b_1 \frac{A_t}{P_{t-1}} + e_t, \ \ \textbf{for} \ \ e_t \sim N\left(0, \sigma^2\right), \tag{10.21}$$

where $\frac{P_t - P_{t-1} + d_t}{P_{t-2}} = {n_t}^{-1} \sum_{i=1}^{n_t} \frac{P_{i,t} - P_{i,t-1} + d_{i,t}}{P_{i,t-2}}$ and $\frac{A_t}{P_{t-1}} = {n_t}^{-1} \sum_{i=1}^{n_t} \frac{A_{i,t}}{P_{i,t-1}}$.

Model VIII

$$\frac{P_t}{P_{t-1}} = a + b_1 \frac{A_t}{P_{t-1}} + e_t, \ \ \textbf{for} \ \ e_t \sim N\left(0, \sigma^2\right), \tag{10.22}$$

where $\frac{P_t}{P_{t-1}} = {n_t}^{-1} \sum_{i=1}^{n_t} \frac{P_{i,t}}{P_{i,t-1}}$ and $\frac{A_t}{P_{t-1}} = {n_t}^{-1} \sum_{i=1}^{n_t} \frac{A_{i,t}}{P_{i,t-1}}$.

Table 10.1 The ARIMAX (p,d,q) specification for each of the nine models and the relative MSE loss functions

Model I	ARIMAX(1,0,1)	2.766371
Model II	ARIMAX(2,0,2)	0.003763
Model III	ARIMAX(2,0,2)	0.741482
Model IV	ARIMAX(2,0,2)	0.719591
Model V	ARIMAX(0,1,1)	0.732002
Model VI	ARIMAX(2,0,2)	0.584564
Model VII	ARIMAX(2,1,1)	0.888239
Model VIII	ARIMAX(2,0,2)	0.762231
Model IX	ARIMAX(2,1,2)	2.245854

Model IX

$$(P_t - P_{t-1}) = a + b_1(A_t - A_{t-1}) + e_t, \quad \textbf{for} \quad e_t \sim N\left(0, \sigma^2\right), \tag{10.23}$$

where $(P_t - P_{t-1}) = {n_t}^{-1}\sum_{i=1}^{n_t}\left(P_{i,t} - P_{i,t-1}\right)$ and $(A_t - A_{t-1}) = {n_t}^{-1}\sum_{i=1}^{n_t}\left(A_{i,t} - A_{i,t-1}\right)$.

The nine aforementioned models are characterized by autocorrelated and heteroscedastic residuals. Thus, the models are expanded in the ARIMAX (autoregressive integrated moving average) framework (for details about ARIMAX modeling, the interested reader is referred to Box and Jenkins (1976)) in order to model the autocorrelated residuals. Moreover, we take into consideration White's (1980) heteroscedasticity-consistent covariance matrix estimator which provides correct estimates of the coefficient covariances in the presence of heteroscedasticity of unknown form.

The ARIMAX (p,d,q) models are estimated in the following form:

$$\begin{gathered}(1 - L)^d y_t = \mathbf{X}_t\beta + e_t \\ \left(1 - \sum_{i=1}^{p} c_i L^i\right) e_t = \left(1 + \sum_{i=1}^{q} d_i L^i\right)\varepsilon_t \\ \varepsilon_t \sim N(0, \sigma^2),\end{gathered} \tag{10.24}$$

where y_t is the dependent variable, $\mathbf{X}_t$ is the vector of explanatory variables, L is the lag operator, and β is a vector of parameters to be estimated. c_i, for $i = 1,\ldots,p$, and d_i, for $i = 1,\ldots,q$, are also parameters to be estimated.

For each of the nine models, the ARIMAX (p,d,q) specification is estimated for $p = 0,1,2$, $d = 0,1,2$ and $q = 0,1,2$. Therefore, for each model, 27 ARIMAX specifications are estimated. The ARIMAX (p,d,q) specification with the minimum value of squared residuals (MSE loss function) is selected as the most appropriate. The following Table presents the selected specifications for each model (Table 10.1).

Hence, the estimated models are the following:

Model I

$$\begin{aligned} P_t &= 8.5 - 5.1A_t + e_t \\ e_t &= 1.05e_{t-1} + \varepsilon_t - 1.46\varepsilon_{t-1} \\ \varepsilon_t &\sim N(0, \sigma^2). \end{aligned} \tag{10.25}$$

Model II

$$\begin{aligned} \frac{A_t}{P_{t-1}} &= -0.004 + 0.99\frac{A_t - A_{t-1}}{P_{t-1}} + 1.04\frac{A_{t-1}}{P_{t-1}} + e_t \\ e_t &= 1.5e_{t-1}0.84e_{t-2} + \varepsilon_t - 1.14\varepsilon_{t-1} + 0.52\varepsilon_{t-2} \\ \varepsilon_t &\sim N(0, \sigma^2). \end{aligned} \tag{10.26}$$

Model III

$$\begin{aligned} \frac{P_t - P_{t-1} + d_t}{P_{t-1}} &= 0.28 - 0.21\frac{A_t}{P_{t-1}} + e_t \\ e_t &= 0.32e_{t-1} + 0.46e_{t-2} + \varepsilon_t + 0.92\varepsilon_{t-1} - 1.16\varepsilon_{t-2} \\ \varepsilon_t &\sim N(0, \sigma^2). \end{aligned} \tag{10.27}$$

Model IV

$$\begin{aligned} \frac{P_t - P_{t-1} + d_t}{P_{t-1}} &= -0.47 + 2.48\frac{A_{t-1}}{P_{t-1}} + e_t \\ e_t &= 0.21e_{t-1} + 0.36e_{t-2} + \varepsilon_t - 0.68\varepsilon_{t-1} - 1.42\varepsilon_{t-2} \\ \varepsilon_t &\sim N(0, \sigma^2). \end{aligned} \tag{10.28}$$

Model V

$$\begin{aligned} (1-L)\frac{P_t - P_{t-1} + d_t}{P_{t-1}} &= 0.055 - 0.5\frac{A_t - A_{t-1}}{P_{t-1}} + e_t \\ e_t &= \varepsilon_t - 1.52\varepsilon_{t-1} \\ \varepsilon_t &\sim N(0, \sigma^2). \end{aligned} \tag{10.29}$$

Model VI

$$\begin{aligned} \frac{P_t - P_{t-1} + d_t}{P_{t-1}} &= -0.33 - 2.6\frac{A_t - A_{t-1}}{P_{t-1}} + 2.4\frac{A_t}{P_{t-1}} + e_t \\ e_t &= 0.4e_{t-1} + 0.3e_{t-2} + \varepsilon_t 0.37\varepsilon_{t-1} - 1.9\varepsilon_{t-2} \\ \varepsilon_t &\sim N(0, \sigma^2). \end{aligned} \tag{10.30}$$

Model VII

$$\begin{aligned} (1-L)\frac{P_t - P_{t-1} + d_t}{P_{t-2}} &= 0.2 - 0.76\frac{A_t}{P_{t-1}} + e_t \\ e_t &= 0.1e_{t-1} + 0.3e_{t-2} + \varepsilon_t - 1.7\varepsilon_{t-1} \\ \varepsilon_t &\sim N(0, \sigma^2). \end{aligned} \tag{10.31}$$

Model VIII

$$\begin{gathered}\frac{P_t}{P_{t-1}} = 1.3 - 0.22\frac{A_t}{P_{t-1}} + e_t \\ e_t = 0.31e_{t-1} + 0.46e_{t-2} + \varepsilon_t + 0.94\varepsilon_{t-1} - 1.1\varepsilon_{t-2} \\ \varepsilon_t \sim N(0, \sigma^2).\end{gathered} \tag{10.32}$$

Model IX

$$\begin{gathered}(1 - L)(P_t - P_{t-1}) = 0.16 - 5.54(A_t - A_{t-1}) + e_t \\ e_t = -0.25e_{t-1} + 0.13e_{t-2} + \varepsilon_t - 1.2\varepsilon_{t-1} - 0.64\varepsilon_{t-2} \\ \varepsilon_t \sim N(0, \sigma^2).\end{gathered} \tag{10.33}$$

In the sequel, for the models being estimated, the estimates of the coefficients, their standard errors, the t-ratios, and their p-values are presented. According to the t-statistics, computed as the ratio of the coefficients to their standard errors, the coefficients of the explanatory variables are statistically significant in models I (price on earnings), II (returns on change in earnings divided by beginning-of-period price and prior period earnings divided by the price at the beginning of the return period), V (returns on change in earnings over opening market value), VII (returns (deflated by lag of 2 years) on earnings over opening market value, and IX (differenced-price model).

Thus, we conclude that these models explain the relationship between dependent and explanatory variables, whereas the models III (returns on earnings over opening market value), IV (returns on prior earnings model over opening market value), VI (returns on change in earnings over opening market value, and on earnings over opening market value), and VIII (return model regressed on earnings over opening market value) fail to explain any strong relationship for the variables under investigation.

Model I

	Coefficient	Std. error	t-statistic	Prob.
C	8.533220	14.06649	0.606635	0.5494
A	**−5.132325**	**1.427713**	**−3.594788**	**0.0013**
AR(1)	1.059546	0.183145	5.785297	0.0000
MA(1)	−1.461802	0.439241	−3.328021	0.0026

Model II

	Coefficient	Std. error	t-statistic	Prob.
C	−0.004505	0.001576	–2.857574	0.0092
A_A_1_TO_P_1	**0.999564**	**0.000720**	**1388.375**	**0.0000**
A_1_TO_P_1	**1.045894**	**0.006215**	**168.2873**	**0.0000**
AR(1)	1.586004	0.418002	3.794251	0.0010

(*continued*)

	Coefficient	Std. error	t-statistic	Prob.
AR(2)	–0.846287	0.326826	–2.589408	0.0167
MA(1)	–1.149986	0.488050	–2.356289	0.0278
MA(2)	0.523007	0.324461	1.611925	0.1212

Model III

	Coefficient	Std. error	t-statistic	Prob.
C	0.283162	0.469953	0.602534	0.5527
A_P_1	**−0.215456**	**0.245653**	**−0.877076**	**0.3895**
AR(1)	0.328559	0.487341	0.674188	0.5069
AR(2)	0.461880	0.150453	3.069922	0.0054
MA(1)	0.922833	0.503703	1.832097	0.0799
MA(2)	−1.168622	0.757600	−1.542532	0.1366

Model IV

	Coefficient	Std. error	t-statistic	Prob.
C	−0.474203	0.250825	−1.890572	0.0713
A_1_TO_P_1	**2.482914**	**2.237444**	**1.109710**	**0.2786**
AR(1)	0.210692	0.313382	0.672316	0.5081
AR(2)	0.363890	0.330713	1.100322	0.2826
MA(1)	−0.683117	0.316795	−2.156340	0.0418
MA(2)	−1.427611	0.387230	−3.686727	0.0012

Model V

	Coefficient	Std. error	t-statistic	Prob.
C	0.055179	0.014928	3.696284	0.0010
A_A_1_TO_P_1	**−0.508631**	**0.167841**	**−3.030438**	**0.0053**
MA(1)	−1.529472	0.264051	−5.792342	0.0000

Model VI

	Coefficient	Std. error	t-statistic	Prob.
C	−0.333214	0.401790	−0.829324	0.4158
A_A_1_TO_P_1	**−2.616280**	**1.897450**	**−1.378840**	**0.1818**
A_P_1	**2.428167**	**1.979060**	**1.226929**	**0.2328**
AR(1)	0.442328	0.194141	2.278386	0.0328
AR(2)	0.322828	0.140587	2.296282	0.0316
MA(1)	0.373195	0.331781	1.124822	0.2728
MA(2)	−1.954250	0.408247	−4.786934	0.0001

Model VII

	Coefficient	Std. error	t-statistic	Prob.
C	0.238146	0.139477	1.707424	0.1018
A_P_1	**−0.760945**	**0.282609**	**−2.692572**	**0.0133**
AR(1)	0.094551	0.236234	0.400242	0.6928
AR(2)	0.326603	0.207761	1.572016	0.1302
MA(1)	−1.739517	0.430607	−4.039685	0.0005

Model VIII

	Coefficient	Std. error	t-statistic	Prob.
C	1.336294	0.475802	2.808510	0.0100
A_P_1	**−0.226643**	**0.247267**	**−0.916591**	**0.3689**
AR(1)	0.313565	0.480801	0.652171	0.5208
AR(2)	0.462875	0.157081	2.946732	0.0072
MA(1)	0.937106	0.494468	1.895180	0.0707
MA(2)	−1.106481	0.744072	−1.487061	0.1506

Model IX

	Coefficient	Std. error	t-statistic	Prob.
C	0.169538	0.371999	0.455748	0.6530
A_MINUS_A_1	**−5.545373**	**0.888799**	**−6.239178**	**0.0000**
AR(1)	−0.254505	0.276560	−0.920251	0.3674
AR(2)	0.135901	0.243004	0.559251	0.5816
MA(1)	−1.192648	0.472557	−2.523820	0.0193
MA(2)	−0.640639	0.667506	−0.959750	0.3476

10.5.1 Unit Root Tests: Testing for Stationarity

We estimate the augmented Dickey and Fuller (1979) test in order to investigate the null hypothesis of a unit root or nonstationarity of the time series under investigation, i.e., y_t. The Dickey-Fuller test is carried out by estimating:

$$\Delta y_t = a_0 + a_1 y_{t-1} + a_2 t + \sum_{i=1}^{l} b_i \Delta y_{t-i} + \varepsilon_t \tag{10.34}$$

where $\varepsilon_t \sim N(0,\sigma^2)$, and Δ is the difference operator, or $\Delta = (1 - L)$. The lag order l of $\sum_{i=1}^{l} b_i \Delta y_{t-i}$ is selected based on the Schwarz (1978) information criterion. The null and the alternative hypotheses may be written as

$$\begin{aligned} H_0 &: a_1 = 0 \\ H_A &: a_1 \neq 0. \end{aligned} \tag{10.35}$$

A rejection of the null hypothesis indicates that the series y_t is stationary. The Eq. 10.20 is estimated in three different versions: (i) for $a_0 = a_2 = 0$, (ii) for $a_2 = 0$, and (iii) for $a_0 \neq 0$ and $a_2 \neq 0$. The variables we are interested in are

P_t in model I; A_t in model I; $\frac{A_t}{P_{t-1}}$ in models II, III, VI, VII, and VIII; $\frac{A_t - A_{t-1}}{P_{t-1}}$ in models II, V, VI; $\frac{A_{t-1}}{P_{t-1}}$ in models II and IV; $\frac{P_t - P_{t-1} + d_t}{P_{t-1}}$ in models III, IV, V, and VI; $\frac{P_t - P_{t-1} + d_t}{P_{t-2}}$ in model VII; $\frac{P_t}{P_{t-1}}$ in model VIII; $(P_t - P_{t-1})$ in model IX; and $(A_t - A_{t-1})$ in model IX.

The test statistics and the relative p-values from testing 21 for the Eq. 10.20 for $a_0 = a_2 = 0$ are the following:

Variable	Test statistic	p-value
P_t	−2.256635	0.0254
A_t	−3.414078	0.0013
$\frac{A_t}{P_{t-1}}$	−5.218131	0.00
$\frac{A_t - A_{t-1}}{P_{t-1}}$	−6.010163	0.00
$\frac{A_{t-1}}{P_{t-1}}$	−3.238033	0.00
$\frac{P_t - P_{t-1} + d_t}{P_{t-1}}$	−4.363369	0.00
$\frac{P_t - P_{t-1} + d_t}{P_{t-2}}$	−5.920676	0.00
$\frac{P_t}{P_{t-1}}$	−2.698176	0.0087
$(P_t - P_{t-1})$	−7.122593	0.00
$(A_t - A_{t-1})$	−7.294227	0.00

The test statistics and the relative p-values from testing 21 for the Eq. 10.20 for $a_2 = 0$ are the following:

Variable	Test statistic	p-value
P_t	−3.625190	0.0109
A_t	−5.722834	0.00
$\frac{A_t}{P_{t-1}}$	−5.587381	0.00
$\frac{A_t - A_{t-1}}{P_{t-1}}$	−6.107705	0.00
$\frac{A_{t-1}}{P_{t-1}}$	−5.774024	0.00
$\frac{P_t - P_{t-1} + d_t}{P_{t-1}}$	−4.399238	0.0016
$\frac{P_t - P_{t-1} + d_t}{P_{t-2}}$	−6.043400	0.00
$\frac{P_t}{P_{t-1}}$	−4.345041	0.0018
$(P_t - P_{t-1})$	−6.998199	0.00
$(A_t - A_{t-1})$	−7.157803	0.00

The test statistics and the relative p-values from testing 21 for the Eq. 10.20 for $a_0 \neq 0$ and $a_2 \neq 0$ are the following:

Variable	Test statistic	p-value
P_t	−3.565939	0.0497
A_t	−5.698735	0.0003
$\frac{A_t}{P_{t-1}}$	−5.667882	0.0004
$\frac{A_t - A_{t-1}}{P_{t-1}}$	−6.213167	0.0001
$\frac{A_{t-1}}{P_{t-1}}$	−5.673596	0.0003
$\frac{P_t - P_{t-1} + d_t}{P_{t-1}}$	−4.332965	0.0092
$\frac{P_t - P_{t-1} + d_t}{P_{t-2}}$	−5.931949	0.0002
$\frac{P_t}{P_{t-1}}$	−4.290153	0.0102
$(P_t - P_{t-1})$	−6.875067	0.00
$(A_t - A_{t-1})$	−7.024534	0.00

In all the cases, the p-values are less than a 5 % level of significance. Therefore, the null hypothesis is rejected at any case. Hence, the series are defined, by the augmented Dickey-Fuller tests, to be stationary.

10.5.2 Forecasting Dependent Variables

For the following models, the dependent variable is predicted for the year 2006, after adding in the dataset the values of the explanatory variables for the year 2006:
Model I

$$\begin{gathered} P_t = 8.5 - 5.1A_t + e_t \\ e_t = 1.05e_{t-1} + \varepsilon_t - 1.46\varepsilon_{t-1} \\ \varepsilon_t \sim N(0, \sigma^2). \end{gathered} \tag{10.36}$$

Model II

$$\begin{gathered} \frac{A_t}{P_{t-1}} = -0.004 + 0.99\frac{A_t - A_{t-1}}{P_{t-1}} + 1.04\frac{A_{t-1}}{P_{t-1}} + e_t \\ e_t = 1.5e_{t-1}0.84e_{t-2} + \varepsilon_t - 1.14\varepsilon_{t-1} + 0.52\varepsilon_{t-2} \\ \varepsilon_t \sim N(0, \sigma^2). \end{gathered} \tag{10.37}$$

Model III

$$\begin{gathered} \frac{P_t - P_{t-1} + d_t}{P_{t-1}} = 0.28 - 0.21\frac{A_t}{P_{t-1}} + e_t \\ e_t = 0.32e_{t-1} + 0.46e_{t-2} + \varepsilon_t + 0.92\varepsilon_{t-1} - 1.16\varepsilon_{t-2} \\ \varepsilon_t \sim N(0, \sigma^2). \end{gathered} \tag{10.38}$$

Table 10.2 The 1-year-ahead forecasts (year 2006) of the dependent variables for the nine models

Number of model	Best model	Forecast of the dependent variable
Model I	ARIMAX(1,0,1)	−1.066527
Model II	ARIMAX(2,0,2)	0.010287
Model III	ARIMAX(2,0,2)	−0.094187
Model IV	ARIMAX(2,0,2)	−0.236392
Model V	ARIMAX(0,1,1)	−0.715401
Model VI	ARIMAX(2,0,2)	0.059651
Model VII	ARIMAX(2,1,1)	0.372343
Model VIII	ARIMAX(2,0,2)	0.833744
Model IX	ARIMAX(2,1,2)	−4.755135

Model IV

$$\begin{gathered}\frac{P_t - P_{t-1} + d_t}{P_{t-1}} = -0.47 + 2.48\frac{A_{t-1}}{P_{t-1}} + e_t \\ e_t = 0.21e_{t-1} + 0.36e_{t-2} + \varepsilon_t - 0.68\varepsilon_{t-1} - 1.42\varepsilon_{t-2} \\ \varepsilon_t \sim N(0, \sigma^2).\end{gathered} \tag{10.39}$$

Model V

$$\begin{gathered}(1-L)\frac{P_t - P_{t-1} + d_t}{P_{t-1}} = 0.055 - 0.5\frac{A_t - A_{t-1}}{P_{t-1}} + e_t \\ e_t = \varepsilon_t - 1.52\varepsilon_{t-1} \\ \varepsilon_t \sim N(0, \sigma^2).\end{gathered} \tag{10.40}$$

Model VI

$$\begin{gathered}\frac{P_t - P_{t-1} + d_t}{P_{t-1}} = -0.33 - 2.6\frac{A_t - A_{t-1}}{P_{t-1}} + 2.4\frac{A_t}{P_{t-1}} + e_t \\ e_t = 0.4e_{t-1} + 0.3e_{t-2} + \varepsilon_t 0.37\varepsilon_{t-1} - 1.9\varepsilon_{t-2} \\ \varepsilon_t \sim N(0, \sigma^2).\end{gathered} \tag{10.41}$$

Model VII

$$\begin{gathered}(1-L)\frac{P_t - P_{t-1} + d_t}{P_{t-2}} = 0.2 - 0.76\frac{A_t}{P_{t-1}} + e_t \\ e_t = 0.1e_{t-1} + 0.3e_{t-2} + \varepsilon_t - 1.7\varepsilon_{t-1} \\ \varepsilon_t \sim N(0, \sigma^2).\end{gathered} \tag{10.42}$$

Model VIII

$$\begin{gathered}\frac{P_t}{P_{t-1}} = 1.3 - 0.22\frac{A_t}{P_{t-1}} + e_t \\ e_t = 0.31e_{t-1} + 0.46e_{t-2} + \varepsilon_t + 0.94\varepsilon_{t-1} - 1.1\varepsilon_{t-2} \\ \varepsilon_t \sim N(0, \sigma^2).\end{gathered} \tag{10.43}$$

Model IX

$$\begin{gathered}(1-L)(P_t-P_{t-1})=0.16-5.54(A_t-A_{t-1})+e_t\\ e_t=-0.25e_{t-1}+0.13e_{t-2}+\varepsilon_t-1.2\varepsilon_{t-1}-0.64\varepsilon_{t-2}\\ \varepsilon_t\sim N(0,\sigma^2).\end{gathered}\tag{10.44}$$

See Table 10.2

10.6 Conclusions and Suggestions for Further Future Research

This study denotes whether some standard models of market-based accounting research can explain stock returns in a different country than the USA and UK. In other words, in the Greek stock market, there is a memory of earnings in stock returns. Models I (price on earnings), II (returns on change in earnings divided by beginning-of-period price and prior period earnings divided by the price at the beginning of the return period), V (returns on change in earnings over opening market value), VII (returns deflated by lag of 2 years on earnings over opening market value), and IX (differenced-price model) have statistically significant coefficients of explanatory variables. In addition, model II with MSE (minimum squared error) loss function in ARIMAX (2,0,2) is prevalent. ARIMAX (2,0,2) is a representation of the ARIMAX (p,d,q) with two lags in autoregressive, after being differenced 0 times, with two lags in residuals.

Unlike the US market, in the Athens Stock Exchange, there is a far lower ratio of professional analysts per registered company, and management earnings forecasts are only recently and very rarely made publicly available which as fact makes earnings change and earnings level as determinant factor in stock returns. Further analysis in earnings components and revenues from sales may explain more satisfactorily the returns and earnings association. This is in the due course. The issuance of a law on 1985 (Presidential Decree 360/1985) for the publication of semiannual reports and the issuance of a law (2533/1997) for the publication of quarterly reports seem to have no effect on the main accounting moments employed in this study. Even the institutional changes such as liberalization of the auditing profession, corporate tax cuts, and the change in currency seem to have no effect on the models under investigation. The different business environment that is going to be formed with forecasted financial statements and the IAS(IFRS) may make the model selection a different task. Besides application of generalized autoregressive conditional heteroscedasticity (GARCH) models is suggested for further future research.

References

Ball, R., & Brown, P. (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research, 6*, 159–178.

Beaver, W. (1970). The time series behavior of earnings variables, empirical research in accounting. Selected studies. *Supplement to Journal of Accounting Research*, 62–99.

Beaver, W., Lambert, R., & Morse, D. (1980). The information content of security prices. *Journal of Accounting and Economics, 2*, 3–28.
Board, J.L.G., & Walker, M. (1990). Intertemporal and cross sectional variation in the association between unexpected accounting rates of return and abnormal returns. *Journal of Accounting Research*, 182–192.
Box, G. E. P., & Jenkins, G. M. (1976). *Time series analysis: Forecasting and control, revised edition*. Oakland: Holden-Day.
Brown, P., Foster, G., & Noreen, E. (1985). *Security analyst multiyear earnings forecasts and the capital market*. Monograph, American Accounting Association, *11*, 1–45.
Chen, C., Chen, S., & Su, X. (2001). Is accounting information value-relevant in the emerging Chinese stock market? *Journal of International Accounting, Auditing and Taxation, 10*, 1–22.
Chen, G., Firth, M., & Ning Gao, D. (2011). The information content of earnings components: Evidence from the Chinese stock market. *European Accounting Review, 20*, 669–692.
Cho, J. Y., & Jung, K. (1991). Earnings response coefficients: A synthesis of theory and empirical evidence. *Journal of Accounting Literature, 10*, 85–116.
Christie, A. (1987). On cross-sectional analysis in accounting research. *Journal of Accounting and Economics, 9*, 231–258.
Collins, D., & Kothari, S. (1989). An analysis of international and cross-sectional determinants of earnings response coefficients. *Journal of Accounting and Economics, 11*, 143–181.
Collins, D., Kothari, S., & Rayburn, J. (1987). Firm size and the information content of prices with respect to earnings. *Journal of Accounting and Economics, 9*, 111–138.
Cornell, B., & Landsman, W. (1989). Security price response to quarterly earnings announcements and analyst forecast revisions. *The Accounting Review, 64*, 680–692.
Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association, 74*, 427–431.
Easton, P., & Harris, T. (1991). Earnings as an explanatory variable for returns. *Journal of Accounting Research, 29*, 19–36.
Easton, P., & Zmijewski, M. (1989). Cross-sectional variation in the stock market response to accounting earnings announcements. *Journal of Accounting and Economics, 29*, 19–36.
Fama, E. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance, 25*(2), 383–417.
Fama, E., Fisher, L., Jensen M., & Roll, R. (1969). The adjustment of stock prices to new information. *International Economic Review, 10*(1), 1–21.
Freeman, R. (1987). The association between accounting earnings and stock returns for large and small firms. *Journal of Accounting and Economics, 9*, 195–228.
Greene, W.H. (2007). *Econometric analysis*. Prentice Hall, New Jersey.
Hirst, E., Koonce, L., & Venkataraman, S. (2008). Management earnings forecasts: A review and framework. *Accounting Horizons, 22*, 315–338.
Holthausen, R. W., & Verrecchia, R. E. (1988). The effect of informedness and consensus on price and volume behavior. *The Accounting Review*, 63, 191–208.
Kahya, E., Maggina, A., & Theodossiou, P. (1993). The information content of accounting earnings in Greece: An empirical investigation. *Journal of Multinational Financial Management, 3*, 143–156.
Kormendi, R., & Lipe, R. (1987). Earnings innovations, earnings persistence and stock returns. *Journal of Business, 60*, 323–345.
Kothari, S. P., & Zimmerman, J. L. (1995). Price and return models. *Journal of Accounting and Economics, 20*, 155–192.
Landsman, W. R., & Magliolo, J. (1988). Cross-sectional capital market research and model specification. *The Accounting Review, 64*, 586–604.
Lev, B., & Ohlson, J. A. (1982). Market based empirical research in accounting: A review, interpretations, and extensions. *Supplement to Journal of Accounting Research, 20*, 249–322.
Lipe, R. (1990). The relation between stock returns and accounting earnings given alternative information. *The Accounting Review, 65*, 49–71.

Lundholm, R., & Myers, L. (2002). Bringing the future forward: The effect of disclosure on the returns-earnings relation. *Journal of Accounting Research, 40*, 809–839.
Mandelbrot, B. (1966). Forecasts of future prices, unbiased markets and martingale models. *Journal of Business, 39*, 242–255.
Ou, J., & Penman, S. (1989). Financial statement analysis and the prediction of stock returns. *Journal of Accounting and Economics, 11*, 295–329.
Rosenberg, B., & Marathe, V. (1975). The prediction of investment risk: Systematic and residual risk. In *Proceedings of the seminar on the analysis of security prices* (pp. 185–226). Chicago: Center for Research in Security Prices, University of Chicago.
Samuelson, P. (1965). Proof that properly anticipated prices fluctuate randomly. *Industrial Management Review, 31*, 41–49.
Schwarz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*, 461–464.
Strong, N., & Walker, M. (1993). The explanatory power of earnings for stock returns. *The Accounting Review, 68*, 385–399.
Trueman, B. (1986). Why do managers voluntarily release earnings forecasts. *Journal of Accounting and Economics, 8*, 53–71.
Wei, W.W. (2006). Time series analysis: Univariate and multivariate methods. Pearson Education.
White, H. (1980). A heteroskedasticity-consistent covariance matrix and a direct test for heteroskedasticity. *Econometrica, 48*, 817–838.

An Assessment of Copula Functions Approach in Conjunction with Factor Model in Portfolio Credit Risk Management

11

Lie-Jane Kao, Po-Cheng Wu, and Cheng-Few Lee

Contents

Abstract

In credit risk modeling, factor models, either static or dynamic, are often used to account for correlated defaults among a set of financial assets. Within the realm of factor models, default dependence is due to a set of common systemic factors. Conditional on these common factors, defaults are independent. The benefit of a factor model is straightforward coupling with a copula function to give an analytic formulation of the joint distribution of default times. However, factor models fail to account for the contagion mechanism of defaults in which a firm's

L.-J. Kao (✉) • P.-C. Wu
Department of Finance and Banking, Kainan University, Taoyuan, ROC, Taiwan
e-mail: ljkao@mail.knu.edu.tw; pcwu@mail.knu.edu.tw

C.-F. Lee
Department of Finance and Economics, Rutgers Business School, Rutgers, The State University of New Jersey, Piscataway, NJ, USA

Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
e-mail: cflee@business.rutgers.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_11,

default risk increases due to their commercial or financial counterparties' defaults. This study considers a mixture of the dynamic factor model of Duffee (*Review of Financial Studies 12*, 197–226, 1999) and a contagious effect in the specification of a Hawkes process, a class of counting processes which allows intensities to depend on the timing of previous events (Hawkes. *Biometrika 58*(1), 83–90, 1971). Using the mixture factor-contagious-effect model, Monte Carlo simulation is performed to generate default times of two hypothesized firms.

The goodness-of-fit of the joint distributions based on the most often used copula functions in literature including the normal, *t*-, Clayton, Frank, and Gumbel copula, respectively, is assessed against the simulated default times. It is demonstrated that as the contagious effect increases, the goodness-of-fit of the joint distribution functions based on copula functions decreases, which highlights the deficiency of the copula function approach.

Keywords

Static factor model • Dynamic factor model • Correlated defaults • Contagious effect • Hawkes process-Monte • Carlo simulation • Normal copula • *t*-copula • Clayton copula • Frank copula • Gumbel copula

11.1 Introduction

An understanding of correlated defaults of the underlying assets is fundamental to portfolio management and the pricing of credit derivatives such as a CDS contract, a CDO contract, or a basket default swap contract (Hull and White 2001; Zhou 2001; Das et al. 2006). There exist, however, different ways of introducing correlations when modeling assets' defaults. In Schönbucher (2003), a good modeling framework for correlated defaults should include (1) being able to produce default correlations of a realistic magnitude, (2) being able to model the timing of defaults, (3) being capable of reproducing default clustering periods, and (4) easily calibrated and implemented by keeping the number of parameters under control without growing dramatically with the number of assets.

In responding to these properties, two types of factor models coupled with a set of small number of common systemic factors have been followed in literature to account for default correlations. The first type of factor model is static in the sense that the modeling of default correlations among a set of I assets is based on the creditworthiness indices $X_1, \ldots, X_I$ at a specific time point T, which is analogous to Merton's structural model (1974) in which an asset defaults when its creditworthiness index falls below some threshold at a specific time point. The static factor model has been widely used for the computation of joint default and loss probability distribution in a portfolio (Vasicek 1997; Koyuoglu and Hickman 1998; Finger 2000; Schönbucher 2003). Portfolio credit risk models fit within this framework include CreditMetrics (1997) of JP Morgan, Credit Risk$^+$ of Credit Suisse Financial Products (1997), and KMV's portfolio manager (Kealhofer 1995).

In the static factor model, default dependence among a set of I assets is introduced by the correlations between the creditworthiness $X_1, \ldots, X_I$ and a set of common systemic economic factors. Conditional on these common systemic factors, the creditworthiness $X_1, \ldots, X_I$ are independent. Depending on the distributions imposing on these common systemic factors, a specific type of copula function can be used to obtain an analytical expression for the joint distribution of the default times for the I assets (Li 1999, 2000; Schönbucher 2003; Gregory and Laurent 2005).

The second type of factor models uses a more consistent framework to model correlated defaults via assets' default intensities that describe the evolution of instantaneous default probabilities dynamically over time. In parallel to the static factor model, the dynamic factor model generates default dependency among assets through a set of common factors. Conditioned to the realization of these common factors, the assets' defaults are independent. However, being different from the static factor model in which common factors affect assets' creditworthiness at a specific time horizon, the common factors in dynamic factor models affect assets' default intensities dynamically over time (Duffee 1999; Yu 2005; Driessen 2005; Elizalde 2005).

The benefit of a factor model, either static or dynamic, is its straightforward conjunction with a copula function to give an analytic formulation of the joint distribution of default times (Li 1999, 2000; Schönbucher and Schubert 2001; Schönbucher 2003; Gregory and Laurent 2005). In this way, model parameters can be readily inferred from the common systemic factors and individual assets' default times. The main drawback of a factor model, however, is that the likelihood of an asset's default does not change due to the defaults of any other assets, i.e., the default of one asset does not trigger the defaults of other related assets. While this might be adequate for production firms, it may be inadequate for studying the default risk of a financial institution with large positions in a few assets whose default can trigger the failure of other financial institutions (Lucas 1995; Jarrow and Yu 2001; Nagpal and Bahar 2001; Das et al. 2006). Under such circumstance, correlated defaults are due to counterparties' default risk. For this reason, the contagious default models in which the default intensity of an asset depends on the status (default/not default) of the other assets are proposed (Davis and Lo 2001; Jarrow and Yu 2001). These contagious default models allow extra-default dependence to be introduced compared to the factor models.

This study uses a mixture of a dynamic factor model by Duffee (1999) with a contagious effect as a benchmark model to generate simulated default times of two assets. Various copula functions that are often used in credit risk modeling, including the normal copula, t-copula, Clayton copula, Frank copula, and Gumbel copula, are employed to calibrate the simulated default times. As the contagious effect increases, the simulated default times show that the goodness-of-fit of the aforementioned copula functions decreases.

The paper is organized as follows. Section 11.2 introduces the static factor model and its link to the copula function approach. Section 11.3 introduces the dynamic factor model in a point process. Section 11.4 introduces the contagious effect model based on mutually exciting point processes. The theoretical foundations of the copula function approach are given in Sect. 11.5. A simulation study is given in Sect. 11.6. Section 11.7 concludes.

11.2 Dependence Structure in Static Factor Model

In a static factor model, the creditworthiness indices $X_1, \ldots, X_I$ of I assets at times $0 < t_1, \ldots, t_I \leq T$ are explained by 1 or a set of common systemic risk factors $Y_1, \ldots, Y_K$ in the form

$$X_i = \sum_{j=1}^{K} \rho_{ij} Y_j + \rho_{i(K+1)} \varepsilon_i$$

where ρ_{ij} determines the relative importance of asset i to the common factor Y_j and $\rho_{i(K+1)}$ the relative weight of idiosyncratic factor ε_i for asset i. To the extent that the creditworthiness indices are related to the common systemic factors $Y_1, \ldots, Y_K$, the likelihood of joint default events across assets varies accordingly. Here the common factors $Y_1, \ldots, Y_K$ and the idiosyncratic factors $\varepsilon_1, \ldots, \varepsilon_I$ are independent. In the following, for simplicity, we will consider a single factor Y and the ith asset's creditworthiness index:

$$X_i = \rho_i Y + \sqrt{1 - \rho_i^2}\, \varepsilon_i. \tag{11.1}$$

Equation 11.1 is the one-factor model adopted in CreditMetrics (Gupton et al. 1997). Since the default environment is completely determined by creditworthiness indices $X_1, \ldots, X_I$ at times $t_1, \ldots, t_I$, respectively, the factor model is static in that the dynamics of the firms' credit quality that evolves during the time horizon $[0, T]$ is ignored.

Suppose default of the ith asset occurs if the creditworthiness index X_i falls below a certain threshold u_i at time t_i and the marginal distribution functions

$$F_1(t_1) = \Pr(\tau_1 \leq t_1), \ldots, F_I(t_I) = \Pr(\tau_I \leq t_I)$$

of the default times $\tau_1, \ldots, \tau_I$ are given. The threshold u_i and the joint distribution of the default times $\tau_1, \ldots, \tau_I$ satisfy $F_i(t_i) = \Pr(X_i \leq u_i)$ and $\mathrm{F}(t_1, \ldots, t_I) = \Pr(X_1 \leq u_1, \ldots, X_I \leq u_I)$, respectively. By noting that conditional on the common factor Y, the creditworthiness indices $X_1, \ldots, X_I$ are independent and therefore

$$\Pr(X_1 \leq U_1, \ldots, X_I \leq U_I) = \int \left[\prod_{i=1}^{I} P(X_i \leq u_i | y) \right] f(y) dy.$$

If assuming the latent common factors Y and idiosyncratic factors $\varepsilon_1, \ldots, \varepsilon_I$ are standard normally distributed, then the creditworthiness indices X_i are also standard normally distributed. Therefore, the conditional probability

$$P(X_i \leq u_i|y) = \Phi\left(\frac{\Phi^{-1}(F_i(t_i)) - \rho_i y}{\sqrt{1-\rho_i^2}}\right)$$

And the joint distribution $F(t_1, \ldots, t_I)$ of default times $\tau_1, \ldots, \tau_I$ is linked to a one-factor normal copula function in the form

$$F(t_1,\ldots,t_I) = \int\left[\prod_{i=1}^{I}\Phi\left(\frac{\Phi^{-1}(F_i(t_i)) - \rho_i y}{\sqrt{1-\rho_i^2}}\right)\right]\phi(y)dy \tag{11.2}$$

where ϕ is the standard normal density function. Alternatively, one can assume that the latent common factor Y is Gamma distributed with parameter $1/\theta$ and let

$$P(X_i \leq u_i|y) = \exp\left(-y\varphi^{-1}(F_i(t_i))\right)$$

where $\varphi^{-1}(s) = s^{-\theta} - 1$ and φ is the generator. Then the joint distribution $F(t_1, \ldots, t_I)$ is linked to a one-factor Clayton copula function in the form

$$F(t_1,\ldots,t_I) = \int\left[\prod_{i=1}^{I}\exp\left(-y\varphi^{-1}(F_i(t_i))\right)\right]\phi(y)dy. \tag{11.3}$$

The Clayton copula function and its generator φ are given in Definition 2.7 of Appendix 2.

11.3 Dependence Structure in Dynamic Factor Model Using Intensity Function

In the dynamic factor model, as in the reduced-form model put forward by Jarrow and Turnbull (1995), Lando (1998), and Duffie and Singleton (1999a), default is treated as an unpredictable jump of a firm's value and the default time τ is treated as the time of the unpredictable jump. Often the unpredictable jump or default is considered as triggered by an exogenous event that occurs with an instantaneous likelihood specified by the intensity function of the first jump in a point process. The formal definition of a point process and properties of its intensity function are given in Definitions 1.1–1.2 of Appendix 1.

In the dynamic factor model, the intensity function of a default depends on certain exogenously determined stochastic common systemic factors X_t which induce correlated defaults among assets (Duffee 1999; Driessen 2005; Elizalde 2005; Yu 2005). In contrast to the static factor model, the evolution of these common factors over time is modeled, and conditional on the evolution of these common factors X_t, defaults are independent. The formulation of the default intensity as such is referred as a doubly Poisson point process (Lando 1998). In Definition 1.3, the properties of a point process and a (doubly) Poisson process are given. An example of a dynamic factor model is given in Duffee (1999), in which the default intensity of asset i is

$$\lambda_{i,t} = a_{0,i} + a_{1,i}\left(X_{1,t} - \overline{X}_1\right) + a_{2,i}\left(X_{2,t} - \overline{X}_2\right) + \lambda^*_{i,t}\psi$$

where $a_{0,i}$, $a_{1,i}$, and $a_{2,i}$ are constants and $X_{1,t}$ and $X_{2,t}$ are two latent factors interpreted as the slope and level of the default-free yield curve, i.e., the risk-free interest rate

$$r_t = a_r + X_{1,t} + X_{2,t}.$$

The asset's specific intensity $\lambda^*_{i,t}$, independent across assets, is assumed to obey a mean-reversion process

$$d\lambda^*_{i,t} = \kappa_i\left(\theta_i - \lambda^*_{i,t}\right)dt + \sigma_i^{\lambda}\sqrt{\lambda^*_{i,t}}\,dW_{i,t} \tag{11.4}$$

where $W_{1,t}, \ldots, W_{\mathrm{I},t}$ are independent Brownian motions. To introduce more default correlation, Duffie and Singleton (1999b) incorporate joint as well as idiosyncratic jumps in the default intensity $\lambda_{i,t}$.

In Das et al. (2007), the hypothesis whether default events can be modeled as a doubly Poisson point process that solely depends on "exogenous" factors is tested. Based on a time series of US corporate defaults, they strongly rejected the hypothesis that defaults can be modeled as a doubly Poisson point process. Instead of a factor model, in the following, the contagious model in which default status of other firms will affect the default intensity of the underlying asset will be considered.

11.4 Contagious Model: Mutually Exciting Intensity Function

Examples of contagious models include the infectious default model by Davis and Lo (2001) in which the intensity function follows a piecewise deterministic Markov process and the propensity model by Jarrow and Yu (2001). In the propensity model, firms are divided into K primary and I-K-1 secondary firms: the default intensities $\lambda_{1,t}, \ldots, \lambda_{K,t}$ of primary firms are determined by some exogenously common factors, while those of the secondary firms depend on the status (default or not) of the primary firms but not the other way around (asymmetric dependence). Specifically, the default intensity of the ith secondary firm is given by

$$\lambda_{i,t} = \overline{\lambda}_{i,t} + \sum_{j=1}^{K} a_{i,t}^{j} 1_{\{\tau_j \leq t\}} \tag{11.5}$$

where $K + 1 \leq i \leq I$. Here $\overline{\lambda}_{i,t}$ represents the part of secondary firm i's hazard rate independent of the default status of other firms.

Recently, Hawkes process, a class of counting processes which allow intensities to depend on the timing of previous events (Hawkes 1971), had been adopted to model the aggregate default intensity of a portfolio (Azizpour and Giesecke 2008; Errais et al. 2010). In Errais et al. (2010), the cumulated default intensity of n firms is specified by

$$\lambda(t) = \lambda_0 + \sum_{j=1}^{n} a e^{-\beta(t-\tau_j)} 1_{\{\tau_j \leq t\}}. \tag{11.6}$$

In Lando and Nielsen (2010), a total of 2,557 firms with an average of 1,142 and a minimum of 1,007 firms at any time throughout the sample period from January 1982 to December 2005 are used to calibrate the goodness-of-fit of an extended dynamic factor model that includes exogenously determined common factors as well as a contagion effect in terms of a Hawkes process. During the sampling period, 370 firms default and a contagious effect is concluded.

In this study, we simulate the default times of two firms using a mixture model based on a dynamic factor model of Duffee (1999) together with a contagious effect in the specification of a Hawkes process. Various commonly used copula functions are calibrated against simulated default times to demonstrate the copula functions in the modeling of credit default risk.

11.5 Copula Functions

Copulas, introduced by Sklar (1959), have been extensively applied in areas such as actuarial science using survival data. It was adopted by Li (2000) and Gregory and Laurent (2005) to the application in modeling the joint distribution of the default times $\tau_1, \ldots, \tau_I$ of a set of I firms. According to Sklar (1959), for any continuous joint distribution $F(t_1, \ldots, t_I)$, there exists a uniquely determined I-dimensional copula $C_F(u_1, \ldots, u_I)$, where $u_i = F(t_i)$, $1 \leq i \leq I$, such that for all $(t_1, \ldots, t_I)$ in R^I,

$$F(t_1, \ldots, t_I) = C_F(F_1(t_1), \ldots, F_I(t_I)). \tag{11.7}$$

In Definition 2.1, the formal definition of a copula is given. The theorem by Sklar (1959) is given in Appendix 2.

According to Sklar (1959), as the existence of the copula function C_F is guaranteed, the joint distribution F of default times $\tau_1, \ldots, \tau_I$ can be obtained via Eq. 11.7 given the marginal distributions $F_1, \ldots, F_I$ of individual default times. However, the analytic formulation of the specific copula function C_F is usually unknown and even intractable. For this reason, some copula functions are chosen in an ad hoc way in credit risk modeling. The most often used copulas are the normal copulas (Li 2000; Frey et al. 2001; Gregory and Laurent 2005). However, normal copula presents no tail dependency and has been criticized for not assigning enough probability for the occurrence of extreme events.

To account for tail dependency, the double-*t*-copula or a member of the Archimedean-type copulas can be used. The *t*-copula is radically symmetric in that its lower and upper tail dependence is the same. In case lower tail dependence is desired, the Clayton copula should be used, while a Gumbel copula should be used for upper tail dependence. However, the Gumbel copula does not allow for negative dependence. To compare these copulas, three indices to measure the dependence structure between two random variables introduced by the copula function are given in Definitions 2.3–2.4, respectively. The three dependence measures include the global dependent measure, i.e., Kendall's tau, and two tail dependent measures, i.e., the upper/lower tail dependence coefficients.

Instead of tail dependence, this study focused on the dependent structure of the firms' default times due to contagious effects. Schönbucher and Schubert (2001) study the dynamics of default intensities and show that a Clayton copula, a member of the Archimedean copula family, is related to the contagious models of Davis and Lo (2001) and Jarrow and Yu (2001). In the following simulation study based on a mixture of the dynamic factor model of Duffee (1999) and a contagious effect specified by a Hawkes process, the aforementioned copula functions are calibrated against the simulated default times generated by the mixture model to test the goodness-of-fit of the copulas when contagious effect is present.

11.6 Simulation Study

A simulation based on a mixture of the dynamic factor model of Duffee (1999) incorporated with contagious effect specified by a Hawkes process is performed here. In Duffee (1999), a firm's default intensity is determined by two factors $X_{1,t}$ and $X_{2,t}$, where $X_{1,t}$ and $X_{2,t}$ are the two latent components of the risk-free interest rate $r_t = a_r + X_{1,t} + X_{2,t}$. The latent components $X_{j,t}$, $j = 1, 2$, obeys the mean-reversion process:

$$dX_{j,t} = \eta_j\left(\mu_j - X_{j,t}\right)dt + \sigma_j^X\sqrt{X_{j,t}}dW_{j,t}^X \tag{11.8}$$

where $W_{1,t}^X$ and $W_{2,t}^X$ are two independent Wiener processes. To account for counterparty default risk, however, a modification is made on the default intensity in Eq. 11.8 by including the default status of the other firm(s). For the two-firm case, the default intensities are, respectively,

Table 11.1 Parameter specification for mixture model

Panel a:			
i	κ_i	θ_i	σ_j^λ
1	0.023	0.0036	0.051
2	0.600	0.1407	0.104
Panel b:			
j	a_r	$\eta_j\mu_j$	σ_j^X
1	−10.474	1.003	0.0134
2	−10.032	0.060	0.0449
Panel c:			
i	$a_{0,i}$	$a_{1,i}$	$a_{2,i}$
1	0.0132	–0.142	0.001
2	0.0196	0.001	0.062

This table reports the parameter values of the proposed mixture model. In Panel a, the parameter values in the mean-reversion process in Eq. 11.4 for the two firms ($i = 1,2$), respectively, are given. In Panel b, parameter values of the latent components $X_{j,t}$ ($j = 1,2$) in Eq. 11.8 are given. In Panel c, parameter values of default intensities $\lambda_i(i = 1,2)$ in Eqs. 11.9 and 11.10, respectively, are given

$$\lambda_{1,t} = a_{0,1} + a_{1,1}\left(X_{1,t} - \overline{X}_1\right) + a_{2,1}\left(X_{2,t} - \overline{X}_2\right) + \lambda_{1,t}^* + \alpha e^{-\beta(t-\tau_2)} 1_{\{\tau_2 \le t\}} \tag{11.9}$$

$$\lambda_{2,t} = a_{0,2} + a_{1,2}\left(X_{1,t} - \overline{X}_1\right) + a_{2,2}\left(X_{2,t} - \overline{X}_2\right) + \lambda_{2,t}^* + \alpha e^{-\beta(t-\tau_1)} 1_{\{\tau_1 \le t\}} \tag{11.10}$$

where τ_1 and τ_2 are the default times of the 1st and 2nd firm, respectively. The firm's specific default intensities $\lambda_{1,t}^*$ and $\lambda_{2,t}^*$ are assumed to follow the mean-reversion processes in Eq. 11.4.

The values of the parameters κ_i, θ_i, η_j, μ_j, $a_{0,i}$, $a_{1,i}$, $a_{2,i}$, σ_j^X, $1 \le i \le 2$, $1 \le j \le 2$ are from Duffee (1999), in which month-end prices of noncallable corporate bonds from January 1985 to December 1995 across 161 firms, with majority of investment grade bonds, are used to calibrate default intensity process. In Table 11.1, the parameter values used are given. Two cases with different contagious effects are considered in the study. In the first case, the parameter α that specifies the contagious effect is set to 0.1, while the α value is 0.25 in the second case. For each case, 10,000 pairs of simulated default times for the two firms are generated.

Five different copula functions are considered to fit the empirical joint distribution of the default times τ_1 and τ_2. They are the normal, *t*-, Clayton, Frank, and Gumbel copula, respectively. For the normal and *t*-copula, the parameter ρ is set to the empirical correlation coefficient of the simulated default times. Minimum mean squares error is used to obtain the estimated parameter values of the degree of freedom v in *t*-copula and θ in Clayton, Frank, and Gumbel copula, respectively.

In Table 11.2, the estimated parameters for the five copula functions and the corresponding sum squares of errors (SSE) are given. As can be seen in Table 11.2,

Table 11.2 Parameter estimation of copula functions

Case I: $\alpha = 0.1$		
Copula type	SSE	Parameter estimation
Normal	0.2394	$\rho = 0.2137$
t-copula	0.1650	$\rho = 0.2137$, d.f. $= 10$
Clayton	0.6043	$\theta = 0.51$
Frank	0.1695	$\theta = 1.70$
Gumbel	0.1591	$\theta = 1.15$
Case II: $\alpha = 0.25$		
Copula type	SSE	Parameter estimation
Normal	0.3867	$\rho = 0.4687$
t-copula	0.2213	$\rho = 0.4687$, d.f. $= 3$
Clayton	1.4439	$\theta = 1.51$
Frank	0.4525	$\theta = 3.75$
Gumbel	0.1431	$\theta = 1.45$

This table reports the estimated parameters of the normal, t-, Clayton, Frank, and Gumbel copulas. For the parameter ρ, empirical correlation of the simulated default times is used. SSE is used to obtain the estimates of the degree of freedom (d.f.) for the t-copula and the parameter θ for the Clayton, Frank, and Gumbel copula, respectively

the normal copula is outperformed by the t-, Frank, and Gumbel copulas when $\alpha = 0.10$, while outperformed by the t- and Gumbel copulas only when $\alpha = 0.25$. It can also be seen that as the contagious effect parameter α increases from 0.10 to 0.25, except the Gumbel copula, the SSEs also increase significantly for the normal, t-, Clayton, Frank, and Gumbel copulas. In both cases, i.e., $\alpha = 0.10$ and $\alpha = 0.25$, the Gumbel copula performs the best while the Clayton copula performs the worst among the five copula approximations.

In Figs. 11.1 and 11.2, the contour plots of the differences between the empirical and copula-based joint distribution functions when the contagious effect parameter $\alpha = 0.10$ and $\alpha = 0.25$, respectively, are given. For the normal copula, as can be seen in Fig. 11.1 (panel a), larger deviations from the empirical joint distribution occur when the default times τ_1 and τ_2 are in the range from 5 to 10 years when $\alpha = 0.10$. When $\alpha = 0.25$, as illustrated in Fig. 11.2 (panel a), larger deviations occur in the range from 5 to 15 years. For the t-copula, larger deviations from the empirical joint distribution occur when the default times τ_1 and τ_2 are in the range from 5 to 10 years when $\alpha = 0.10$. However, in panel b of Fig. 11.2 where $\alpha = 0.25$, in addition to the range between 5 and 10 years, the t-copula fails to approximate the empirical joint distribution well for small default times τ_1 and τ_2. This implies that as the contagious effect increases, the t-copula does not explain the lower tail dependence well. In panel d of Figs. 11.1 and 11.2, when using the Gumbel copula, larger deviations occur when the default times τ_1 and τ_2 are in the range from 2 to 7 years. This indicates that the Gumbel copula does not approximate well for smaller default times.

Taken together, the goodness-of-fit of the normal, t-, Clayton, Frank, and Gumbel copula, respectively, decreases as the contagious effect parameter α increases. This phenomenon is more apparent for the Frank copula. The contour plots of the

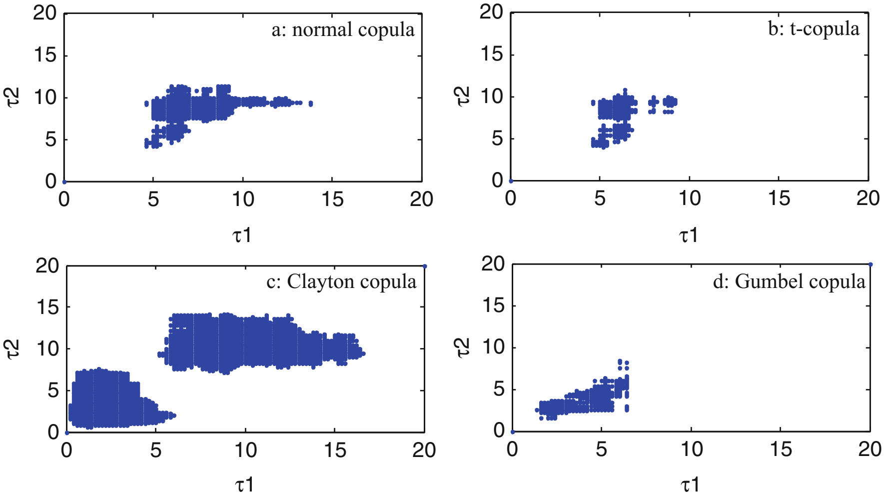


Fig. 11.1 Contours of Differences between Empirical Joint Distribution and Copula Approximation. Note Panels a-d show the contours of differences with absolute values exceeding 0.01 for normal, t-Clayton, and Gumbel copula, respectively. The contagious effect parameter a is 0.1

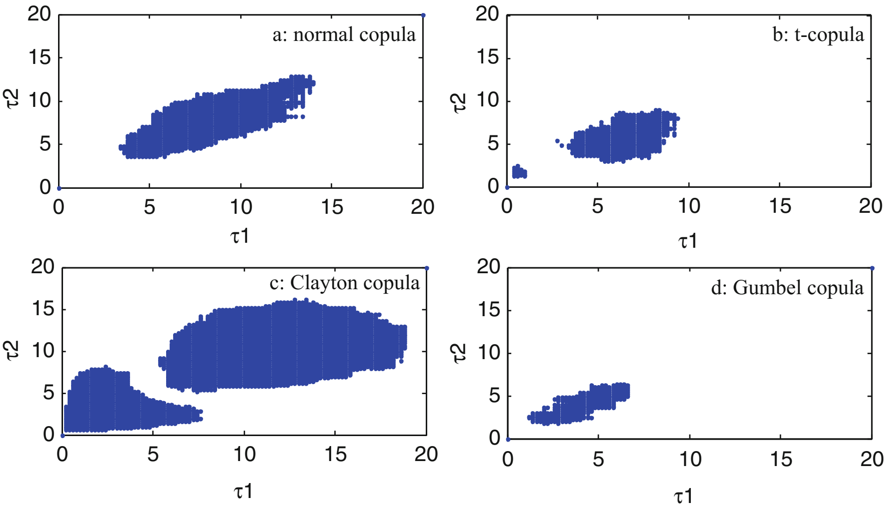


Fig. 11.2 Contours of Differences between Empirical Joint Distribution and its Copula Approximation. Note Panels a-d show the contours of differences with absolute values exceeding 0.01 for normal, t-Clayton, and Gumbel copula, respectively. The contagious effect parameter a is 0.25

differences between the empirical and Frank copula-based joint distribution functions when the contagious effect parameter $\alpha = 0.10$ and $\alpha = 0.25$, respectively, are given in Panel a–b of Fig. 11.3. As can be seen in Fig. 11.3, when $\alpha = 0.10$, the Frank copula fails to approximate the empirical distribution well only for small default

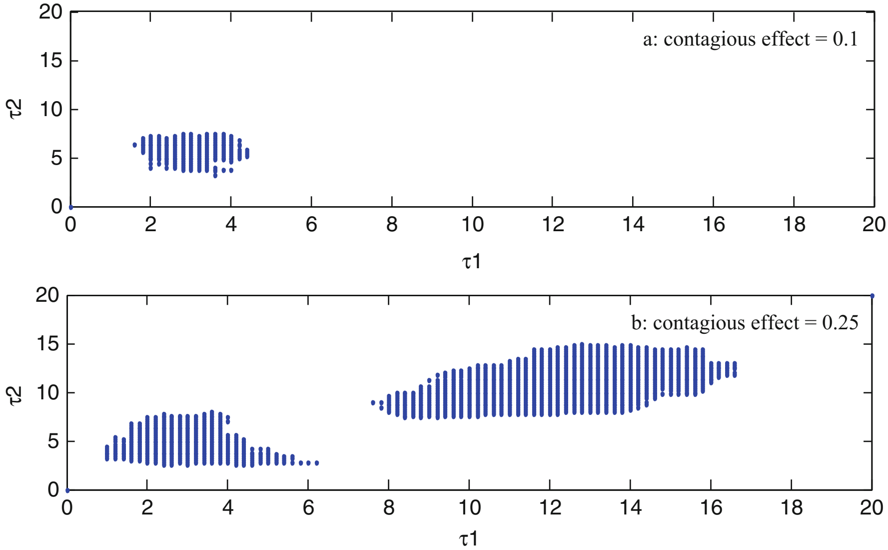


Fig. 11.3 Contours of Differences between Empirical Joint Distribution and its Frank Copula Approximation. Note Panels a–d show the contours of differences with absolute values exceeding 0.01 for normal, t-Clayton, and Gumbel copula, respectively. The contagious effect parameter a is 0.1 and 0.25

times τ_1 and τ_2. As the contagious effect parameter $\alpha = 0.25$, however, the Frank copula fails to approximate the empirical distribution well not only for small default times τ_1 and τ_2 but also larger τ_1 and τ_2. This can also be seen in Table 11.2, where the SSE increases from 0.1695 to 0.4525 as α increases from 0.10 to 0.25.

11.7 Conclusion

As a firm's default determines the economic opportunities available, it is very likely that the status of firms' defaults will affect one another. That is, a firm's default might have contagious effect on other firms' defaults. As a result, one often observes firms' default simultaneously. This study demonstrates that the copula functions from literature do not necessarily, and most unlikely, match the joint distribution F of the correlated default times $\tau_1, \ldots, \tau_I$ of a set of I firms as the contagious effect is increasing.

Five of the most often used copula functions in literature, namely, the normal, *t*-, Clayton, Frank, and Gumbel copulas, are under study. Among the five copula functions, the Clayton copula performs the worst, whereas the Gumbel copula performs the best. Except the Gumbel copula, the goodness-of-fit of the other four copulas decreases as the contagious effect increases. This suggests further advanced statistical tool for the modeling of contagious defaults in a more consistent and accurate way is in need.

Appendix 1: Point Process and Its Intensity

Let (Ω, H, P) be a probability space, where P is a physical measure. A formal definition of a point process is given below.

Definition 1.1 A point process N with stopping times $\tau_1, \tau_2, \ldots . \in [0,T]$ is a counting measure on the probability space (Ω, H, P) in that for any Borel subset $E \subseteq [0,T]$, the counting measure $N(E)$ represents the number of time points $\tau_1, \tau_2, \ldots$, in E.

Let $\{H_t: t \in [0, T]\}$ be a filtration on (Ω, P) so that H_t contains the accumulated information generated by a point process N till time t. The notion of the compensator of the point process N is given below.

Definition 1.2 Define the compensator A of the point process N as the unique random measure on (Ω, P) such that:

i. A is H_t-predictable.

ii. For every nonnegative H_t-predictable process H,

$$E\left(\int_0^T HdN\right) = E\left(\int_0^T HdA\right).$$

The characterization of the compensator A is essential to the statistical inference of the point process N in that the compensated process $M(t) = N(t) - A(t)$, $t \in [0,T]$ is a martingale under P. In this case a positive, H_t-predictable process $\lambda(t)$ exists so that

$$A(t) = \int_0^t \lambda(s)ds.$$

Then $\lambda(t)$ is called the intensity process of the point process N. The H_t-predictable intensity has the x dt, the probability one of the stopping time τ_k that occurs during $(t\text{-}dt,t]$ is

$$P[N(t-dt,t] = 1|H_t-] = \lambda(t)dt + \mathrm{o}(dt).$$

Definition 1.3 If the compensator process $A(t)$ is continuous and deterministic, then the point process $N(t)$ is a Poisson process. A doubly stochastic Poisson process $M(t)$ is a point process in which the intensity process $\lambda(t)$ is F_t-predictable, where F_t contains the accumulated information generated by the point process M till time t and the entire trajectory $\{\lambda(s): s > 0\}$. Conditioning on one realization of the intensity process $\{\lambda(s): s > 0\}$, the point process $M(t)$ is an inhomogeneous Poisson process.

Appendix 2: Copula Functions

Basically, a copula function C links the joint distribution $F(y_1, \ldots, y_I)$ of I multivariate random variables $Y_1, \ldots, Y_I$ to their univariate marginal distributions $F_1(y_1), \ldots, F_I(y_I)$. A formal definition of a copula function is given as follows.

Definition 2.1 A function C: $[0, 1]^N \rightarrow [0, 1]$ is a copula if there are uniform random variables $U_1, \ldots, U_N$ taking values in [0, 1] such that C is their joint distribution function that satisfies:

1. $C(1, \ldots, 1, u_i, 1, \ldots, 1) = u_i$ for all $i = 1, \ldots, N$ and $u_i \in [0, 1]$.
2. For all $\boldsymbol{u} \in [0, 1]^N$, $C(\boldsymbol{u}) = 0$ if at least one coordinate $u_i = 0$, $i = 1, \ldots, N$.
3. For all $\boldsymbol{u} \in [0, 1]^N$, $\boldsymbol{v} \in [0, 1]^N$ with $u_i \leq v_i$, $i = 1, \ldots, N$, the C volume of the hypercube

$$\sum_{i_1=1}^{2} \cdots \sum_{i=1}^{2} (-1)^{i_1+\ldots+i_N} C(w_{i_1}, \ldots, w_{i_N}) \geq 0$$

where $w_{i_k} = u_k$ if $i_k = 1$ else $w_{i_k} = v_k$.

The theoretical groundwork of applying a copula function is based on Sklar's theorem (Sklar 1959) in the following:

Theorem 2.2 Let $\tau_1, \ldots, \tau_I$ be random variables with marginal distribution functions $F_1, \ldots, F_I$ and joint distribution function F. Then there exists an I-dimensional copula C such that

$$F(t_1, \ldots, t_I) = C(u_1, \ldots, u_I) = C(F_1(t_1), \ldots, F_I(t_I))$$

for all $(t_1, \ldots, t_I)$ in R^I, where $u_i = F(t_i)$, $1 \leq i \leq I$. Moreover, if each F_i is continuous, then the copula C is unique.

In order to calibrate a copula, indices that measure the dependence structure introduced by the choice of the copula function are used. We focus on three dependence measures that depend only on the copula function, not in the marginal distributions. Among the three dependence measures, Kendall's *tau* is the measure of global dependence, while upper/lower tail dependence coefficients are two local measures of dependence.

Definition 2.3 Let F be the joint distribution of two random variables X_1 and X_2 with marginal F_1 and F_2, respectively. Kendall's *tau* is the probability of concordance minus the probability of discordance. Specifically, if (X_1, X_2) and (X_1', X_2') are two realizations of F, then Kendall's *tau* is defined as

$$P\left[\left(X_1 - X_1'\right)\left(X_2 - X_2'\right) > 0\right] - P\left[\left(X_1 - X_1'\right)\left(X_2 - X_2'\right) < 0\right] \qquad (11.11)$$

If C is the copula of the joint distribution F, i.e., $C(u_1, u_2) = C(F_1(x_1), F_2(x_2))$, Kendall's *tau* is

$$tau = 4\int_0^1\int_0^1 C(u,v)dC(u,v) - 1.$$

Definition 2.4 Let $F(X_1, X_2)$ be the joint distribution of two random variables X_1 and X_2, with marginal F_1 and F_2. The coefficients of upper and lower tail dependence are defined, respectively, as

$$\lambda_U = \lim_{u\uparrow 1} \Pr\left[X_1 > F_1^{-1}(u) \middle| X_2 > F_2^{-1}(u)\right]$$
$$\lambda_L = \lim_{u\downarrow 0} \Pr\left[X_1 < F_1^{-1}(u) \middle| X_2 < F_2^{-1}(u)\right].$$

Definition 2.5. (Normal Copula) The I-dimensional normal copula is expressed as

$$C(u_1, \ldots, u_{\boldsymbol{I}}) = \Phi_\Sigma^I\left(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_I)\right),$$

where Σ is a positive-definite correlation matrix, Φ_Σ^I is the distribution function of an I-dimensional multivariate normal random vector with correlation matrix Σ, and Φ^{-1} is the inverse of the distribution function of a standard normal random variable. The density function c of $C(u_1, \ldots, u_{\boldsymbol{I}})$ is

$$c(u_1, \ldots, u_{\boldsymbol{I}}) = \frac{1}{|\Sigma|^{1/2}} \exp\left(\frac{1}{2}\boldsymbol{\zeta}'\left(\Sigma^{-1} - \mathrm{I}_{I\times I}\right)\boldsymbol{\zeta}\right)$$

where $\boldsymbol{\zeta}' = (\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_I))$, and $\mathrm{I}_{I\times I}$ is the unity matrix. For a normal copula, the relationship between the linear correlation coefficient ρ and Kendall's *tau* is

$$\rho = \sin(2\pi \times tau). \tag{11.12}$$

On the other hand, the coefficients of upper and lower tail dependence are both zero, i.e., a normal copula is tail independent.

Definition 2.6 (*t*-Copula) The I-dimensional t-copula is expressed as

$$C(u_1, \ldots, u_{\boldsymbol{I}}) = t_{v,\Sigma}^I\left(t_v^{-1}(u_1), \ldots, t_v^{-1}(u_I)\right)$$

where $t_v{}^{-1}$ denotes the inverse of the distribution function of a univariate t-student random variable with v degrees of freedom and $t_{v,\Sigma}^I$ denotes the distribution function of a multivariate t-distribution with v degrees of freedom and positive-definite dispersion matrix Σ. Its density function is

$$\frac{\left(\Gamma\left(\frac{v}{2}\right)\right)^{I-1}\Gamma\left(\frac{v+I}{2}\right)\left(1+\frac{\boldsymbol{\zeta}'\Sigma^{-1}\boldsymbol{\zeta}}{v}\right)^{-\frac{v+I}{2}}}{\left(\Gamma\left(\frac{v+1}{2}\right)\right)^{I}|\Sigma|^{-\frac{1}{2}}\,{}_{i=1}^{I}\left(1+\frac{\zeta_i^2}{v}\right)^{-\frac{(v+1)}{2}}}$$

where $\boldsymbol{\zeta}' = (t_v^{-1}(u_1), \ldots, t_v^{-1}(u_I))$. For a t-copula, the relationship between the linear correlation coefficient ρ and Kendall's *tau* is the same as Eq. 11.12. The coefficients of upper and lower tail dependence are

$$\lambda_U = \lambda_L = 2 - 2t_v\left(\frac{(v+1)(1-\rho)}{1+\rho}\right)^{1/2}.$$

Definition 2.7 (Archimedean Copula) The I-dimensional Archimedean copula can be expressed as

$$C(u_1, \ldots, u_I) = \phi^{-1}\left(\sum_{i=1}^{I} \phi(u_i)\right) \tag{11.13}$$

where the generator ϕ is a continuous strictly decreasing function from [0, 1] to $[0, \infty]$ satisfying $\phi(0) = \infty$ and $\phi(1) = 0$. In particular, when the generator

$$\phi(u) = \frac{u^{-\theta} - 1}{\theta}, \theta \geq 0$$

then Eq. 11.13 is called a Clayton copula. When the generator

$$\phi(u) = -\ln\left(\frac{e^{-\theta u} - 1}{e^{-\theta} - 1}\right), \theta \neq 0$$

Equation 11.13 is called a Frank copula. When the generator $\phi(u) = (-\ln u)^{\theta}$, $\theta \geq 1$, Eq. 11.13 is a Gumbel copula. In particular, the density function of a Clayton copula is

$$c(u_1, \ldots, u_I) = \left(1 - I + \sum_{i=1}^{I} u_i^{-\theta}\right)^{-I-1/\theta} \prod_{i=1}^{I}\left[u_i^{-\theta-1}((i-1)\theta + 1)\right]. \tag{11.14}$$

For a Clayton copula, Kendall's *tau* and tail dependency measures are

$$tau = 1 + 4\int_0^1 \frac{u^{\theta} - 1}{\theta u^{\theta-1}} du = \frac{\theta}{\theta - 2}$$

$$\lambda_U = 2 - 2\lim_{v \to 0} \frac{\phi'(2v)}{\phi'(v)} = 0$$

$$\lambda_L = 2\lim_{v \to 0} \frac{\phi'(2v)}{\phi'(v)} = 2$$

where the function $\varphi(u) = \phi^{-1}(v)$ is the inverse of the generator ϕ.

References

Azizpour, S., & Giesecke, K. (2008). *Self-exciting corporate defaults: Contagion vs. Frailty*. Working paper, Stanford University.

Credit Suisse Financial Products. (1997). CreditRisk^{+}: A credit risk management framework. New York.

CreditMetrics. (1997). The benchmark for understanding credit risk. Technical document. New York: JP Morgan.

Das, S. R., Freed, L., Geng, G., & Kapadia, N. (2006). Correlated default risk. *Journal of Fixed Income, 16*(2), 7–32.

Das, S., Duffie, D., Kapadia, N., & Saita, L. (2007). Common failings: How corporate defaults are correlated. *Journal of Finance, 62*, 93–117.

Davis, M., & Lo, V. (2001). Infectious defaults. *Quantitative Finance, 1*, 382–387.

Driessen, J. (2005). Is default event risk priced in corporate bonds? *Review of Financial Studies, 18*, 165–195.

Duffee, G. R. (1999). Estimating the price of default risk. *Review of Financial Studies, 12*, 197–226.

Duffie, D., & Singleton, K. J. (1999a). Modeling term structures of defaultable bonds. *Review of Financial Studies Special, 12*(4), 687–720.

Duffie, D., Singleton, K. J. (1999b). Simulating correlated defaults. Working paper, Graduate School of Business, Stanford University.

Elizalde, A. (2005). Do we need to worry about credit risk correlation? *Journal of Fixed Income, 15*(3), 42–59.

Errais, E., Giesecke, K., & Goldberg, L. R. (2010). Affine point processes and portfolio credit risk. *SIAM Journal on Financial Mathematics, 1*, 642–665.

Finger, C. (2000). A comparison of stochastic default rate models. *Risk Metrics Journal, 1*(2), 49–73.

Frey, R., McNeil, A., & Nyfeler, M. (2001). Copulas and credit models. *RISK*, 111–114.

Gregory, J., & Laurent, J. P. (2005). Basket default swaps, CDO's and factor copulas. *Journal of Risk, 7*(4), 103–22.

Gupton, G., Finger, C., & Bhatia, M. (1997). *Credit metrics technical document*. New York: Morgan Guaranty Trust.

Hawkes, A. G. (1971). Spectra of some self-exciting and mutually exciting point processes. *Biometrika, 58*(1), 83–90.

Hull, J., & White, A. (2001). Valuing credit default swaps II: Modeling default correlations. *Journal of Derivatives, 8*(3), 12–22.

Jarrow, R., & Turnbull, S. (1995). Pricing derivatives on financial securities subject to credit risk. *Journal of Finance, 50*, 53–86.

Jarrow, R., & Yu, F. (2001). Counterparty risk and the pricing of defaultable securities. *The Journal of Finance, 56*, 1765–1799.

Kealhofer, S. (1995). Managing default risk in portfolios of derivatives. In: David M. Rowe (ed.) *Derivative Credit Risk: Advances in Measurement and Management*. London: Risk Publications.

Koyuoglu, H. U., & Hickman, A. (1998). Reconcilable differences. *Risk, 11*(10), 56–62.

Lando, D. (1998). On cox processes and credit risky securities. *Review of Derivative Research, 2*, 99–120.

Lando, D., & Nielsen, M. S. (2010). Correlation in corporate default: Contagion or conditional independence? *Journal of Financial Intermediation, 19*(3), 355–372.

Li, D. X. (1999). On default correlation: A copula function approach. Working paper 99–07, The Risk Metrics Group.

Li, D. X. (2000). On default correlation: A copula approach. *Journal of Fixed Income, 9*, 43–54.

Lucas, D. (1995). Default correlation and credit analysis. *Journal of Fixed Income, 9*(4), 76–87.

Merton, R. (1974). On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance, 29*, 449–470.

Nagpal, K., & Bahar, R. (2001). Measuring default correlation. *Risk, 14*(3), 129–132.

Schönbucher, P. J. (2003). Credit derivatives pricing models. Wiley Finance, UK.
Schönbucher, P., & Schubert, M. (2001). Copula dependent default risk in intensity models. Working paper, Department of Statistics, Bonn University.
Sklar, A. (1959). Fonctions de repartition a *n* dimensions et leurs marges. *Public Institute Statistic University Paris, 8*, 229–231.
Vasicek, O. A. (1997). The loan loss distribution. Technical report, KMV Corporation.
Yu, F. (2005). Correlated default in reduced-form models. *Journal of Investment Management, 3*(4), 16–27.
Zhou, C. (2001). An analysis of default correlations and multiple defaults. *Review of Financial Studies, 14*(2), 555–576.

Assessing Importance of Time-Series Versus Cross-Sectional Changes in Panel Data: A Study of International Variations in Ex-Ante Equity Premia and Financial Architecture

12

Raj Aggarwal and John W. Goodell

Contents

R. Aggarwal (✉)
University of Akron, Akron, OH, USA
e-mail: aggarwa@uakron.edu

J.W. Goodell
College of Business Administration, University of Akron, Akron, OH, USA
e-mail: johngoo@uakron.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_12,

Abstract

In the study of economic and financial panel data, it is often important to differentiate between time series and cross-sectional effects. We present two estimation procedures that can do so and illustrate their application by examining international variations in expected equity premia and financial architecture where a number of variables vary across time but not cross-sectionally, while other variables vary cross-sectionally but not across time. Using two different estimation procedures, we find a preference for market financing to be negatively associated with the size of expected premia. However, we also find that US corporate bond spreads negatively determine financial architecture according to the first procedure but not according to the second estimation as US corporate bond spreads change value each year but have the same value across countries. Similarly some measures that change across countries but do not change across time, such as cultural dimensions as well as the index of measures against self-dealing, are significant determinants of financial architecture according second estimation but not according to the first estimation. Our results show that using these two estimation procedures together can assess time series versus cross-sectional variations in panel data. This research should be of considerable interest to empirical researchers.

We illustrate with simultaneous-equation modeling. Following a Hausman test to determine whether to report fixed or random-effects estimates, we first report random-effects estimates based on the estimation procedure of Baltagi (Baltagi 1981; Baltagi and Li 1995; Baltagi and Li 1994). We consider that the error component two-stage least squares (EC2SLS) estimator of Baltagi and Li (1995) is more efficient than the generalized two-stage least squares (G2SLS) estimator of Balestra and Varadharajan-Krishnakumar (1987). For our second estimation procedure, for comparative purposes we use the dynamic panel modeling estimates recommended by Blundell and Bond (1998). We employ the model of Blundell and Bond (1998), as these authors argue that their estimator is more appropriate than the Arellano and Bond (1991) model for smaller time periods relative to the size of the panels. We also use this two-step procedure and use as an independent variable the first lag of the dependent variable, reporting robust standard errors of Windmeijer (2005). Thus, our two different panel estimation techniques place differing emphases on cross-sectional and time series effects, with the Baltagi-Li estimator emphasizing cross-sectional effects and the Blundell-Bond estimator emphasizing time series effects.

Keywords

Panel data estimates • Time series and cross-sectional effects • Econometrics • Financial institutions • Banks • Financial markets • Comparative financial systems • Legal traditions • Uncertainty avoidance • Trust • Property rights • Error component two-stage least squares (EC2SLS) • Generalized two-stage least squares (G2SLS)

12.1 Introduction

Recently, there has been a veritable explosion of studies in a wide variety of areas that employ panel data econometric methodology. Analysis of panel data allows independent variables to vary both cross-sectionally and across time, while panel data econometrics correct for cross-correlations between time series and cross-sectional error terms. In recent years there has been the introduction of a number of new refinements in analyzing panel data (Petersen 2009; Wooldridge 2010), all maintaining the goal of accounting for cross-correlation between time series and cross-sectional error terms when assessing coefficient significance.

However, in the study of economic and financial panel data, it is often important to assess the differential impact of time series versus cross-sectional effects; but panel data techniques are unclear how this may be accomplished. In other words, panel data methodologies typically do not inform us fully regarding which effect (time series or cross-sectional) is more important or more dominant within particular data sets or contexts. In this chapter we employ two contrasting estimation procedures, which, respectively, emphasize cross-sectional versus time series differences, to clarify the impacts of these two influences. We undertake this comparison of econometric methods within a finance-related context which takes into account possible endogeneity.

We illustrate with simultaneous-equation modeling (outlined below). Following a Hausman test to determine whether to report fixed or random-effects estimates, we first report random-effects estimates based on the estimation procedure of Baltagi (1981; Baltagi and Li 1994, 1995). We consider that the error component two-stage least squares (EC2SLS) estimator of Baltagi and Li (1995) is more efficient than the generalized two-stage least squares (G2SLS) estimator of Balestra and Varadharajan-Krishnakumar (1987).

For our second estimation procedure, for comparative purposes we use the dynamic panel modeling estimates recommended by Blundell and Bond (1998). We employ the model of Blundell and Bond (1998), as these authors argue that their estimator is more appropriate than the Arellano and Bond (1991) model for smaller time periods relative to the size of the panels. We also use the two-step procedure, reporting robust standard errors of Windmeijer (2005). Within this modeling we use as an independent variable the first lag of the dependent variable. Thus, our two different panel estimation techniques place differing emphases on cross-sectional and time series effects, with the Baltagi-Li estimator emphasizing cross-sectional effects and the Blundell-Bond estimator emphasizing time series effects.

Regarding the context of this econometric study, recent research suggests that both national variations in the structures of financial intermediation and equity premia are determined by many similar socioeconomic factors so that national variations in equity premia can be expected to influence national variations in the structure of financial intermediation and vice versa. But, until very recently prior literatures in these two areas have ignored each other and the resulting possible endogeneity problems. Recent exceptions include Aggarwal and Goodell (2011a, b), both of which examine the role of financial architecture in determining nations' equity premia. In this chapter we use this economic context to illustrate two differing estimation procedures.

We find that a number of independent variables used in this study significantly determine financial architecture across nations and across time. While a number of our results are consistent across differing estimation procedures, we also document a number of results that differ according to the two estimation procedures. After controlling for outliers and serial correlation in further robustness checks, we find that some significant differences remain, differences that help assess the relative influence of time series versus cross-sectional variations.

We document, using the two different estimation procedures, a preference for market financing to be negatively associated with the size of expected premia, However, after controlling for multicollinearity and serial correlation in robustness checks, we also find that US corporate bond spreads negatively determine financial architecture according to the first procedure but not according to the second estimation as US corporate bond spreads change value each year but have the same value across countries. Similarly some measures that change across countries but do not change across time, such as cultural dimensions as well as the index of measures against self-dealing, are significant determinants of financial architecture according to the second estimation but not according to the first estimation. We conclude from our presented example that the two estimation procedures can produce results with different emphases with regard to cross-sectional and time series effects.

12.2 Literature

12.2.1 Panel Data Estimation Procedures: Time Series Versus Cross-Sectional Effects

In recent years, the development of panel data econometrics has facilitated a large increase in scholarship where panel data models are applicable. This is particularly the case in international finance where data can be described across countries and time as well as across industries and time. The development of panel data methods has followed from the introduction by Hansen (1982) of Generalized Method of Moments (GMM). GMM, including the use of instrumental variables, allows the implementation of consistent estimations based on conditional expectations which are inconsistent with the use of earlier methods such as ordinary least squares regression.

Dynamic effects can render the fixed-effects estimator of panel models biased and inconsistent, especially for data covering finite and short time periods. Among alternative estimators that control for persistence is the system Generalized Method of Moments (GMM) estimator proposed by Blundell and Bond (1998). This procedure addresses econometric problems such as regressor endogeneity, measurement error, and weak instruments while controlling for time-invariant, country-specific effects such as distance or common language. Arellano and Bond (1991) suggest transforming the model, either in the first differences or in orthogonal deviations, to eliminate the fixed effects and to estimate it by using the two-step GMM estimator. The second and higher lags of the endogenous variable in levels are suitable instruments to overcome the estimation problem. However, when data are highly persistent, Blundell and Bond (1998) argue that this procedure can be improved by using the system – GMM estimation – which supplements the equations in the first differences with equations in levels. For the equations in the first differences, the instruments used are the lagged levels, and for the equations in levels, the instruments are the lagged differences.

In addition, the use of simultaneous-equation modeling on the same data set allows us to assess the results using the Blundell and Bond estimation procedure against an alternative. Following a Hausman test to determine whether to report fixed or random-effects estimates, we report random-effects estimates based on the estimation procedure of Baltagi (1981; Baltagi and Li 1994, 1995). We believe that the error component two-stage least squares (EC2SLS) estimator of Baltagi and Li (1995) is more efficient than the generalized two-stage least squares (G2SLS) estimator of Balestra and Varadharajan-Krishnakumar (1987) because of a broader set of transformations of the instruments.

It is very useful to examine the differences obtained with these two estimation procedures. These two panel data estimation procedures have different emphases with regard to cross-sectional versus time series effects. Because of the nature of their construction, while the Baltagi-Li estimator emphasizes cross-sectional effects, the Blundell-Bond estimator emphasizes time series effects.

12.2.2 Financial Architecture, Transactions Costs, and Risks

The channeling of funds from savers to investors, or financial intermediation, is a necessary function in all countries and is generally undertaken primarily through financial institutions and/or through financial markets. Either financing channel must resolve the issues of asymmetric information, adverse selection, and agency costs involved in financing contracts that cover the monitoring and collection of funds provided by savers to investors. Given that all optimal contracts are incomplete, the efficacy and efficiency of overcoming contracting costs depends on the nature of "hold-up" costs in a country, i.e., the ability and willingness of the contracting parties to try and take advantage of each other.

This ability and willingness to take advantage of the other party in incomplete contracts depend not only on industrial structure and the legal environment (reflecting the relative power of the contracting agents and legal constraints on their behavior) but also on ethical and other informal conventions that depend on social and cultural values. As these differ from country to country and given that institutions and markets differ in how they enforce incomplete contracts, financial institutions may be optimal in some combinations of ethical, cultural, and social conditions, while financial markets may be optimal in other conditions, and financial institutions may be favored in some countries, while financial markets are favored in other countries. Recent research notes that national preferences for market financing increase with political stability, societal openness, economic inequality, and equity market concentration and decrease with regulatory quality and ambiguity aversion (e.g., Modigliani and Perotti 2000; Ergungor 2004; Kwok and Tadesse 2006; Aggarwal and Goodell 2009).

Of course, as noted by Coase (1960), in a theoretically ideal and perfect financial system, it would make no difference whether financial intermediation was privately done through banks or publicly through markets. However, in reality many factors must be considered. According to North (1990), the costliness of information needed for measurement and enforcement of exchanges creates "transaction costs." Transaction costs involve costs of defining property rights and costs of enforcing contracts – including costs of information. "Transformation costs" are the costs associated with using technology and the efficiency of factor and product markets and are reflected in transactions costs. Whether institutions lower or raise overall transactions costs has to do in part with the ability of participants to be informed and to understand the nature of the particular institutional environment. This includes not just understanding the nature of contracts and their enforceability but also the temperament and motivations of other participants.

Additional transactions costs may also be associated with market transactions. As noted by Williamson (1988) and others more recently (e.g., Aggarwal and Zhao 2009), Transaction Cost Economics (TCE) suggests that when the costs of market exchange are sufficiently high, firms can obtain cheaper financing through some other means. The alternative to market financing is typically through some sort of a prescribed arrangement, such as a bank loan or, more broadly, through a prescribed transfer of resources through a horizontal or vertical network. Hart (1995, 2001) recognize that the primary transaction costs of market exchanges stem from the uncertainties of contracts. From the point of view of the equity investor, obtaining reliable information about firms is innately costly and, to some degree, fallible. These costs will be shared with the supplier of equity, causing equity financing to be more costly for the firm. This view is supported by Bhattacharya and Thakor (1993) who suggest that a unifying thread among a great number of papers on banking is that "intermediation is a response to the inability of market-mediated mechanisms to efficiently resolve informational problems."

Modigliani and Perotti (2000) theorize that when societies' enforcement regimes are not adequate, bank financing is favored. In this instance the binding of transactions becomes more private than public and reflects longer-term reputations and relationships between the parties, such as those between firms and their banks. Modigliani and Perotti (2000) suggest that when the rights of minority (or outside) investors are not adequate, less equity investment will be available for new enterprises (also see Myers 1977). According to Modigliani and Perotti (2000), in such societies, there will be more bank lendings instead of financing with public equity.

Modigliani and Perotti (2000) also suggest that banks, because of an emphasis on collateral, are less likely or able to differentiate firms with good future prospects versus those with poor future prospects. Alternatively, markets with good governance are better able to distinguish between these types of firms, a view supported by recent literature.[1] However, Rajan (1992) notes that a higher emphasis on

[1]For instance, Shirai (2004) reports that, because of improvements in official oversight for the period 1997–2001, Indian capital markets improved significantly in being able to differentiate high-quality firms from low-quality firms.

collateral can also lead to such an informational advantage for banks such that they can charge excessively high interest rates which can weaken economic development. Underlying Rajan and Zingales (1998) is the notion that when reliable information about firms is too difficult or costly for the general public, banks provide delegated monitoring (see Diamond 1984).

As this brief review of the relevant literature indicates, the international determinants of financial architecture must include national characteristics such as quality of investor protection, cultural and legal variables, and the equity premium. However, there may be an endogeneity problem as financial architecture seems to depend on a number of similar variables and the two variables, financial architecture and equity premia, may influence each other. Our research design includes a resolution of this issue.

12.2.3 Equity Premium as Measure of Equity Risk

We can expect variations in the ex ante equity premia across countries. Equity premia can be expected to reflect the price of risk in equity investments. Depending on national characteristics such as the nature of their institutional structures and their levels of financial development, countries may differ with regard to both the risk involved in equity investments and in the price of such risk. One way to think about this extra risk and its price is to think about how the supply and demand for equity investments may differ across countries, especially as most countries have less than perfect capital markets.

Ibbotson et al. (2006) suggest that because of many obstacles and limitations, the supply and demand for equity in markets may not respond to market forces as would be expected from a theoretical view of efficient markets.[2] For example, the supply of equity may be restricted as bureaucratic rules and regulations may deter the formation and market listing of corporate shares. Similarly, the demand for equity may be limited due to uncertain property rights and the unreliability of public information on potential investments. As the equity premium is the price of equity risk, it is determined by the balance between supply and demand for equity. In order to understand the nature and size of equity premia, it is important to account for the nature of equity demand and supply in actual, imperfect, markets.

While there are no perfect measures of national supply of equity, a number of variables could be used as indirect indicators. For example, stock market capitalization as a ratio of GDP is one such widely used measure. But this ratio by itself is an inadequate proxy because this measure will rise with a rise in valuation (as well as with new listings) and valuation is itself affected by supply. Other arguable measures of equity supply would include the number of shares

[2]Ibbotson et al. (2006) actually start with somewhat of an alternative view to the commonly held notion that prices in capital markets are set by the supply and demand for capital. Instead, they focus on the viewpoint of the supplier of capital (an investor) and suggest that there is a supply and demand for returns and that it is returns that are priced in the marketplace.

listed per unit of population or the number of new shares listed. A more extensive evaluation of how developed an equity market is might include R^2, the degree to which individual stocks move synchronously with the overall equity market in that country (Morck et al. 2000), or the degree to which market capitalization is concentrated in a few firms. Turner (2003) associates the development of nations' private bond markets with the quality of their local investor bases. Regulatory restrictions and lack of accounting standards can inhibit bond trading by institutional investors. So if the quality of local bond markets in some measure reflects the narrowness of the local investor base, then the quality of the bond market might partially and indirectly determine the nonpecuniary (e.g., for control rights) demand for equity.

Similarly, the demand for equity returns is likely to be influenced by a great variety of factors that influence the risk level of equity and society's perceptions, tolerance, or price of equity risk. The nature of legal protection for investors, disclosure requirements, the level of social trust that a particular society believes can be placed in strangers, and the political stability of a country certainly are some factors that come to mind. However, it is also reasonable to suppose that there is correlation among many of the social, cultural, legal, and governance factors that might affect the demand for equity. Ibbotson et al. (2006) suggest that the demand for equity return is potentially also affected by concern for real returns as opposed to nominal returns. Further, Moerman and van Dijk (2010) document evidence that inflation risk is priced in international asset returns – so our investigation of the demand for equity returns also ought to control for inflation variability.

The supply and demand for equity in a country are also likely to be affected by its financial architecture, i.e., the relative importance of the banking sector versus financial markets in a country. It is now well recognized that some countries like Japan and Germany are bank oriented, while other countries, like the Anglo-Saxon countries, depend more on financial markets. For example, it can be expected that equity premia are likely to be lower in countries with well-developed financial and equity markets with a less restricted, and so greater, supply of equity than in bank-oriented countries.

12.3 Methodology

As discussed briefly above, it seems that both national variations in the structures of financial intermediation and equity premia are determined by many similar socio-economic factors so that national variations in equity premia can be expected to influence national variations in the structure of financial intermediation and vice versa. But, until very recently, prior literatures in these two areas have ignored each other and the resulting possible endogeneity problem. Recent exceptions include Aggarwal and Goodell (2011a, b), both of which papers examine the role of financial architecture in partial determining nations' equity premia. This chapter overcomes this limitation of prior research and examines the determinants of international variations in financial architecture, accounting for the relationship between financing architecture and ex ante equity risk premia with different estimation techniques. One estimation technique uses simultaneous-equation estimates that include the equity premium as an instrumental variable along with a number of other relevant institutional, governance, and cultural factors. Another estimation technique uses a dynamic panel data modeling which includes as an independent variable the ex ante equity premium as a predetermined variable. Specifically, this chapter examines the nature of international variations in national financial architecture with a special emphasis on exploring the role of ex ante equity premia.

12.3.1 Estimating the Equity Premium

This chapter uses the abnormal earnings growth (AEG) model to estimate the ex ante equity premium. This model is chosen not just because of its efficacy at predicting ex post values but rather also because it is regarded as a fundamental model upon which other previously used models are based (Penman 2005). In addition, we use AEG estimations rather than alternative estimations in order to avoid assumptions of future payout ratios. Further, the AEG model avoids the problem with some estimation procedures when composite market-to-book ratios for particular country/years are less than one.

Our estimates of ex ante equity premiums are based on data from the period 1996–2006. Consensus of all available individual earnings forecasts are obtained from Institutional Brokers Estimate System (I/B/E/S) of Thomson Financial, as well as data for actual earnings per share, dividends per share, share prices, and the number of shares outstanding. In the interests of consistency, forecasts are collected as of April each year. Most firms end their fiscal years in December, and, typically, values are reported within 90 days of the fiscal year end. If a firm has I/B/E/S earnings forecasts for +1 and +2 years and a 5-year growth forecast, it is retained in the sample. Firm-level data are then aggregated for each year. This paper uses national treasury bill rates as a measure of the risk-free rate. This data is obtained from *International Financial Statistics* of the International Monetary Fund (IMF).

This chapter estimates that abnormal earnings grow at a constant rate after 5 years at a rate equal to the rate of expected nominal GDP growth. Expected GDP growth is modeled according to the following equation:

$$\text{expNomGDPgrowth} = (1 + \text{expInflation}) * (1 + \text{expRealGDPgrowth}) - 1 \tag{12.1}$$

Expected inflation is modeled as the arithmetic average of the current, preceding, and subsequent years' inflation as reported by IFS. Expected real GDP growth is modeled as the arithmetic average of the current, preceding and subsequent years' real GDP growth as reported by *World Development Indicators*. If national treasury bill rates are not available, a 1-year money market rate or a similar country-specific short-term rate is used.

This chapter uses the following AEG model:

$$p_0 = \frac{eps_1}{k}\left(\frac{g_{st} + k\frac{d_1}{eps_1} - g_{lt}}{k - g_{lt}}\right) \tag{12.2}$$

where eps_1 is the next-period forecasted earnings per share (when all firms aggregated, this becomes next year's forecasted earnings), p_0 is the price which when all firms are aggregated becomes aggregate market values, d_1 are the dividends per share which become aggregate dividends, g_{lt} is the long-term growth rate, and g_{st} is the short-term growth rate. We proxy the short-term rate as the geometric average of growth in earnings forecasts up to the fifth forecast. The long-term growth rate is proxied as the rate of expected nominal GDP growth. We do not incorporate the 5-year growth forecast from I/B/E/S in the estimate of the short-term growth rate because of concerns that, by using an average of the 5-year growth forecast from I/B/E/S, a bias will be introduced as a consequence of equal weighting across years.

12.3.2 Estimating the Cross-Border Determinants of Financial Architecture

Besides the use of expected equity premia as an instrumental variable, we also use a number of other independent variables. We control for inflation variability (INFLATION_VOLATILITY) using the variance of the preceding 5 years as reported by IFS. To capture time variation in risk appetite, we include the annual difference from Moody's Baa corporate bond yield and 10-year Treasury yield as reported by the *St Louis Federal Reserve* for the middle of April for each year. STOCK_VOLATILITY is the annualized standard deviation of monthly equity returns of the respective Morgan Stanley Country Index. To account for relative firm size and equity market concentration, we establish a Herfindahl index for each country for each year (CONCENTRATION). A value of CONCENTRATION close to 1 would suggest that most market capitalization for a particular country in a given year is due to a small number of firms. Market capitalization data is obtained from I/B/E/S. Based on Aggarwal and Goodell (2009), we speculate in advance of empirical findings that country/years with high market capitalization concentration will be more market based.

We control for regional differences by including as an independent variable a dummy variable and REGION_EUROPE that receive "1" if the country is Europe. We consider that the European region has historically had a unique relationship with banking (see Rajan and Zingales 2003). We also control for cross-national differences in wealth and wealth inequality by including real GDP and the Gini coefficient from *World Development Indicators*.

In order to assess the cultural impact on financial architecture, we include as independent variables four cultural dimensions of Hofstede (2001):

uncertainty avoidance (UAI), power distance (PDI), masculinity (MAS), and individualism (IDV). Kwok and Tadesse (2006) and others have found that national culture is an important partial determinant of nations' preference for markets versus banks.

We control for regulatory quality using a governance indicator from Djankov et al. (2008) his is the comprehensive index of cross-national differences in protection of investors from self-dealing, the anti-self-dealing index (ANTI_SELF_DEAL). Fligstein (2001) suggests that market participants primarily value stable worlds and that markets are state-directed societal solutions to competition. More efficient regulation and control of corruption will improve transparency and so lower costs of resolving the asymmetry of information inherent with markets. Aggarwal and Goodell (2009), Kwok and Tadesse (2006), and Ergungor (2004) have found that governance impacts financing preferences.

12.3.3 Estimation Methodologies

In order to use the estimator of Baltagi and Li (1995), we design a set of simultaneous equations. With regard to our modeling of simultaneous equations, we suggest that the risk factors and costs generally involved with resolving asymmetric information in markets is reflected in the equity premium. Therefore we model financial architecture as partially determined by the respective country/year equity premium as well as political factors. Our empirical estimation models are based on this set of equations:

$$\mathrm{FIN\_ARCH}_{it} = \alpha_{it} + \sum \beta_{1it} * \mathrm{EQ\_PREM}_{it} + \beta_{2it}\mathrm{X2}_{it} + e_{it} \tag{12.3}$$

$$\mathrm{EQ\_PREM}_{it} = \alpha_{it} + \beta_{1it}\mathrm{FIN\_ARCH}_{it} + \sum \beta_{2it}\mathrm{X1}_{it} + e_{it} \tag{12.4}$$

In Eqs. 12.3 and 12.4, FIN_ARCH is domestic stock market capitalization divided by domestic assets of deposit money banks; this measure is constructed from the data from Beck et al. (2000). EQ_PREM is our estimate of the equity premium, and X1 and X2 represent a number of independent variables, including risk, political, and social factors. Alternatively, as noted earlier, we also use, for the same respective sets of independent variables, the results of Blundell and Bond (1998) dynamic panel estimation procedures.

12.3.4 Data

We analyze data for 41 countries over a recent 11-year period using panel data methods. Specifically, we assess data for the 11-year period, 1996–2006, from

41 countries: Argentina, Australia, Austria, Belgium, Brazil, Canada, Chile, Czech Republic, Denmark, Finland, France, Germany, Greece, India, Indonesia, Ireland, Italy, Hong Kong, Hungary, Israel, Japan, Korea, Malaysia, Mexico, New Zealand, the Netherlands, Norway, Philippines, Pakistan, Peru, Poland, Portugal, Russia, Singapore, South Africa, Spain, Sweden, Thailand, Turkey, the United Kingdom, and the United States.[3] The countries we include here cover much breadth across regions, cultures, legal origins, and difference in national wealth. In selecting independent variables we have tried and managed to avoid excess correlation among them. We estimate models that focus on structural variables and then add in turn a regional variable and then cultural, governance, and security protection variables. All models have variance inflation factors (VIF) of less than 10 for all regressors indicating that any multicollinearity is unlikely to be a problem. Nevertheless, in subsequent robustness checks, we also address other specific correlations among particular pairs of independent variables.

12.4 Results

12.4.1 Descriptive Statistics

Table 12.3 presents the means, and standard deviations of the variables used in our models. Together these independent variables reflect the factors we described earlier that may affect cross-national differences in financing predilections. Column 4 shows the standard deviation divided by the mean. This column suggests that the dummy variable for European region, and our Herfindahl for market concentration are the most variable of the independent variables (1.06, 0.96), while FIN_ARCH, and EQUITY PREMIUM are similarly variable (0.73, 0.40) (Table 12.3).

12.4.2 Results of Baltagi-Li EC2SLS Modeling and Blundell-Bond Modeling

Table 12.4 shows the results of panel regressions using the system of Eqs. 12.3 and 12.4, with our estimate of financial architecture as a dependent variable and our estimate of the equity premium as an instrumented variable. The estimation is carried out using the random-effects EC2SLS estimator proposed by Baltagi and Li (1992, 1995), and Baltagi (2001). Baltagi (2001) suggests the EC2SLS estimator is more

[3]We begin our period of study in 1996 in order to include our measure of market concentration which is a Herfindahl index we construct using data from I/B/E/S. There is insufficient data from I/B/E/S for many countries prior to 1996. Another important reason, however, is that the measures we use for political stability, control of corruption, and regulatory from Kaufmann et al. (2008) are only available from 1996. Generally, our sample is restricted to those country/years which have sufficient data reported by I/B/E/S and are included by Kaufmann et al. (2008). We stop in 2006 to avoid the effects of the global financial crises and recession that started in 2007.

Table 12.1 Mean financial architecture

Country	Mean	Country	Mean
South Africa	2.47	Greece	0.69
The United States	2.35	Israel	0.64
Finland	2.00	Korea	0.60
Hong Kong	1.93	Indonesia	0.59
Sweden	1.72	Spain	0.58
Chile	1.52	Brazil	0.58
Singapore	1.47	Norway	0.57
Russia	1.40	Hungary	0.57
Switzerland	1.36	Ireland	0.57
Argentina	1.31	Belgium	0.56
Malaysia	1.20	New Zealand	0.51
The United Kingdom	1.15	Pakistan	0.49
Australia	1.11	Italy	0.48
Philippines	0.99	Poland	0.45
Canada	0.92	Thailand	0.41
India	0.87	Czech Republic	0.41
Mexico	0.87	Japan	0.40
The Netherlands	0.75	Germany	0.33
Denmark	0.72	Portugal	0.31
France	0.72	Austria	0.18
Turkey	0.71		

This table lists the mean national financial architecture for 41 countries for 1996–2006. Financial architecture is domestic stock market capitalization divided by domestic assets of deposit money banks (this measure is constructed from the data from Beck et al. (2000))

efficient than the usual Balestra and Varadharajan-Krishnakumar (1987) G2SLS estimator because of a broader set of transformations of the instruments. Alternatively Table 12.5 shows, for the same respective sets of independent variables as Table 12.4, the results of Blundell and Bond (1998) dynamic panel estimates.

Table 12.4 reports the results of EC2SLS random-effects estimates, while Table 12.5 reports Blundell and Bond (1998) estimates. Model 1 restricts the independent variables to just the equity premium. In this model EQ_PREM is negatively significant at 5 % in both Tables 12.4 and 12.5. Both estimation procedures suggest in Model 1 an association of lower equity premia with more market-oriented countries.

Model 2 adds to the independent variables that describe the volatility of the financial and economic environment: INFLATION_VOLATILITY and STOCK_VOLATILITY. The equity premium is again negatively significant, at 1 % in Table 12.4 and 10 % in Table 12.5. Equity premium is again negatively significant according to both estimation procedures. INFLATION_VOLATILITY is not significant with regard to either estimation procedure, while STOCK_VOLATILITY is negatively significant in both Tables 12.4 and 12.5.

Table 12.2 Mean equity premia

Country	Mean	Country	Mean
Brazil	−4.24	Finland	7.24
Turkey	1.61	Australia	7.33
Argentina	1.71	Hong Kong	7.37
Hungary	3.08	Germany	7.51
Mexico	3.62	Austria	7.65
Poland	3.75	Canada	7.65
Norway	3.82	France	7.84
India	4.44	Japan	7.85
Indonesia	5.19	Chile	8.03
South Africa	5.54	Greece	8.35
Belgium	6.01	Switzerland	8.38
Denmark	6.08	Portugal	8.38
The United Kingdom	6.24	The Netherlands	8.41
Italy	6.27	Malaysia	8.62
Israel	6.38	Thailand	8.85
Philippines	6.42	Spain	9.07
New Zealand	6.57	Korea	9.27
Singapore	6.64	Pakistan	10.61
Czech Republic	6.84	Ireland	10.65
Sweden	6.84	Russia	39.16
The United States	7.02		

This table lists the mean expected national mean equity premia for 41 countries for 1996–2006. This chapter uses the following AEG model

$$p_0 = \frac{eps_1}{k}\left(\frac{g_{st} + k\frac{d_1}{eps_1} - g_{lt}}{k - g_{lt}}\right)$$

Where eps_1 is the next-period forecasted earnings per share (when all firms aggregated, this becomes next year's forecasted earnings), p_0 is the price which when all firms are aggregated becomes aggregate market values, d_1 are the dividends per share which becomes aggregate dividends, g_{lt} is the long-term growth rate, and g_{st} is short-term growth rate. We proxy the short-term rate as the geometric average of growth in earnings forecasts up to the fifth forecast. The long-term growth rate is proxied as rate of expected nominal GDP growth

Model 3 adds to the other three independent variables CONCENTRATION, REGION_EUROPE, and LN_GDP. This results in CONCENTRATION being positively significant with regard to both estimation procedures. Equity premium is again negatively significant in both tables. A difference in the two tables is that in Table 12.4 REGION_EUROPE is negatively significant, while in Table 12.5 this variable is not significant. STOCK_VOLATILITY is negatively significant in both Tables 12.4 and 12.5, although only at 10 % in Table 12.4.

Model 4 add to the independent variables the Hofstede cultural dimensions UAI, PDI, MAS, and IDV. Model 4 also adds SPREAD, the spread between US corporate bonds and US 10-year Treasuries. This results in differences between

Table 12.3 Summary of data sources used in this study

1	2	3	4	5
Variable	Mean	Standard deviation	Stdev/ mean	Source
FIN_ARCH	0.92	0.67	0.73	Beck et al. (2000)
EQUITY_PREMIUM	7.28	2.91	0.40	Equity premium estimated from the data from I/B/E/S and treasury rate data from *International Financial Statistics*
INFLATION_VOLATILITY	8873.06	296.65	0.03	Variance of preceding 5 years of inflation
STOCK_VOLATILITY	24.49	14.77	0.60	The annualized standard deviation of monthly equity returns of respective Morgan Stanley Country Index
CONCENTRATION	0.18	0.17	0.96	Herfindahl index created for this chapter based on number of firms
REGION_EUROPE	0.47	0.50	1.06	A dummy variable that is assigned "1" if the country is in Europe and "0" otherwise
LN_GDP	9.12	1.24	0.14	PPP GDP per capita from *World Development Indicators*
UAI	62.76	24.88	0.40	Uncertainty Avoidance, Hofstede (2001)
PDI	53.74	21.72	0.40	Power Distance, Hofstede (2001)
MAS	51.22	19.44	0.38	Masculinity, Hofstede (2001)
IDV	51.52	23.83	0.46	Individualism, Hofstede (2001)
SPREAD	2.15	0.53	0.25	Difference from Moody's Baa corporate bond yield and 10-year Treasury yield *St Louis Federal Reserve*
ANTI_SELF_DEAL	0.52	0.24	0.47	Djankov et al. (2008)

Means, standard deviations, and sources of variables used in statistical estimates reported in succeeding tables, 1996–2006

Tables 12.4 and 12.5. SPREAD is significantly negative in Table 12.5 while not significant in Table 12.4. Three cultural variables, PDI, MAS, and IDV, are significant in Table 12.4 while not significant in Table 12.5. In Table 12.4, UAI and MAS are negatively significant at 1 % and 5 %, respectively. In Table 12.4, IDV and PDI are both positively significant at 1 %. However, UAI, MAS, IDV, and PDI are not significant in Table 12.5. In other results, both tables show LN_GDP positively significant, suggesting equity premia are larger in wealthier countries.

Model 5 adds to the independent variables of Model 4, the comprehensive index of anti-self-dealing (ANTI_SELF_DEAL) of Djankov et al. (2008). These result, in Table 12.4, with regulatory efficiency, as reflected in measures against self-dealing being positively significant. In Table 12.5, however, this variable is not significant.

Table 12.4 Cross-national determinants of financial architecture: static panel estimation

Dependent variable: FIN_ARCH	Model				
	1	2	3	4	5
INTERCEPT	2.73*** (0.001)	2.87*** (0.000)	0.75*** (0.173)	–0.68 (0.232)	–0.73 (0.255)
EQUITY_PREMIUM	–0.25** (0.023)	–0.19*** (0.003)	–0.03* (0.065)	–0.01 (0.385)	–0.03** (0.034)
INFLATION_VOLATILITY		–0.00 (0.302)	–0.00 (0.109)	–0.00 (0.136)	–0.00 (0.132)
STOCK_VOLATILITY		–0.03*** (0.008)	–0.01* (0.061)	–0.00 (0.280)	–0.00* (0.064)
CONCENTRATION			0.58*** (0.007)	0.80*** (0.000)	0.86*** (0.000)
REGION_EUROPE			–0.30*** (0.007)	–0.37*** (0.000)	–0.20 (0.140)
LN_GDP			0.00 (0.204)	0.13*** (0.005)	0.10* (0.067)
UAI				–0.01*** (0.001)	–0.00 (0.110)
PDI				0.01*** (0.000)	0.01*** (0.000)
MAS				–0.00** (0.039)	–0.00** (0.049)
IDV				0.01*** (0.000)	0.01*** (0.004)
SPREAD				–0.00 (0.968)	0.04 (0.465)
ANTI_SELF_DEAL					0.46* (0.071)
Observations/groups	397/41	397/41	396/41	395/41	395/41
Hausman	1.05 (0.305)	1.69 (0.429)	3.35 (0.646)	4.75 (0.447)	3.65 (0.723)
R-square (within, between, overall)	(0.01,0.04,0.00)	(0.01,0.03,0.00)	(0.03,0.00,0.01)	(0.06,0.23,0.20)	(0.04, 0.15,0.12)
Wald chi square	5.18** (0.023)	9.69** (0.021)	24.30*** (0.000)	74.21*** (0.000)	62.87*** (0.000)
Baltagi Wu LBI	0.89	0.88	0.95	0.97	0.97

This table reports results of tobit regressions for 41 countries, for 1996–2006. FIN_ARCH is stock market capitalization divided by domestic assets of deposit money banks from Beck, Demirguc-Kunt, Levine (2000); EQUITY PREMIUM is ex ante equity premium estimations from the data from I/B/E/S; INFLATION_VOLATILITY is variance of preceding 5 years of inflation *International Financial Statistics*; STOCK_VOLATILITY is the annualized standard deviation of monthly equity returns from respective Morgan Stanley Country Index; CONCENTRATION is a Herfindahl index of equity market concentration created for this chapter; REGION_EUROPE is a dummy variable that is assigned "1" if market is in Europe and "0" otherwise; LN_GDP is natural log of real GDP per capita; UAI, PDI MAS, and IDV are cultural dimensions of Hofstede (2001); SREAD is the difference between Moody's Baa corporate bond yield and 10-year Treasury yield from *St Louis Federal Reserve*; ANTI_SELF_DEAL is the index of measures against self-dealing from Djankov et al. (2008)

Variance inflation factors are less than 10 for all variables and models. P values in parentheses. *** Significance at 1 % level, ** significance at 5 % level, * significance at 10 % level random-effects, EC2SLS estimates reported

Table 12.5 Cross-national determinants of financial architecture: dynamic panel estimation

Dependent variable: financial architecture	Model				
	1	2	3	4	5
INTERCEPT	0.17^{***} (0.000)	0.36^{***} (0.000)	1.11^{**} (0.046)	–5.77 (0.240)	–6.48 (0.244)
LAG1_FIN_ARCH	0.85^{***} (0.000)	0.88^{***} (0.000)	0.85^{***} (0.000)	0.77^{***} (0.000)	0.75^{***} (0.000)
EQUITY_PREMIUM	-0.00^{**} (0.021)	-0.00^{*} (0.063)	-0.00^{***} (0.010)	-0.00^{**} (0.024)	-0.00^{**} (0.036)
INFLATION_VOLATILITY		–0.00 (0.822)	–0.00 (0.922)	–0.00 (0.593)	–0.00 (0.692)
STOCK_VOLATILITY		-0.01^{***} (0.000)	-0.01^{***} (0.000)	-0.01^{***} (0.000)	-0.01^{***} (0.000)
CONCENTRATION			0.97^{***} (0.000)	0.96^{***} (0.000)	1.10^{***} (0.000)
REGION_EUROPE			0.11 (0.391)	0.33 (0.786)	0.12 (0.936)
LN_GDP			–0.09 (0.158)	0.55^{**} (0.025)	0.73^{**} (0.028)
UAI				0.02 (0.405)	0.03 (0.380)
PDI				0.01 (0.658)	0.01 (0.793)
MAS				0.01 (0.688)	0.01 (0.661)
IDV				–0.02 (0.510)	–0.03 (0.464)
SPREAD				-0.09^{**} (0.013)	-0.08^{**} (0.025)
ANTI_SELF_DEAL					–1.43 (0.629)
Observations/groups	366/41	366/41	365/41	365/41	365/41
Wald chi square	1453.14^{***} (0.000)	358.33^{***} (0.000)	268.08^{***} (0.000)	318.68^{***} (0.000)	244.79^{***} (0.000)
AR tests	–0.04(0.965)	0.14(0.889)	–0.31(0.757)	0.22(0.825)	0.23(0.820)
	–2.11(0.035)	–2.07(0.038)	–2.26(0.024)	2.30(0.022)	2.31(0.021)
	–0.80(0.422)	–1.13(0.257)	–1.22(0.222)	1.01(0.314)	1.02(0.308)
	2.18(0.029)	2.24(0.025)	2.12(0.034)	2.36(0.018)	2.29(0.022)

This table reports results of tobit regressions for 41 countries, for 1996–2006. FIN_ARCH is stock market capitalization divided by domestic assets of deposit money banks from Beck, Demirguc-Kunt, Levine (2000), EQUITY PREMIUM is ex ante equity premium estimations from the data from I/B/E/S; INFLATION_VOLATILITY is variance of preceding 5 years of inflation from *International Financial Statistics*; STOCK_VOLATILITY is the annualized standard deviation of monthly equity returns of respective Morgan Stanley Country Index; CONCENTRATION is a Herfindahl index of equity market concentration created for this chapter; REGION_EUROPE is a dummy variable that is assigned "1" if market is in Europe and "0" otherwise; LN_GDP is natural log of real GDP per capita; UAI, PDI MAS, and IDV are cultural dimensions of Hofstede (2001); SREAD is the difference between Moody's Baa corporate bond yield and 10-year Treasury yield from *St Louis Federal Reserve*; ANTI_SELF_DEAL is the index of measures against self-dealing from Djankov et al. (2008)

Variance inflation factors are less than 10 for all variables and models. P values in parentheses. *** Significance at 1 % level, ** significance at 5 % level, * significance at 10 % level Blundell-Bond estimates reported

12.4.3 Additional Econometric Details

In our results for Table 12.4, we consistently report random-effects estimates. We do this because the results of Hausman tests (Hausman 1978) suggest that random-effects estimates may be used. Reading across the models in Table 12.4, we see that Hausman tests for the models are consistently insignificant. Therefore we do not reject the NULL hypothesis that the estimates are inconsistent (Hausman 1978). In the Appendix we also report the Pearson correlation coefficients (Appendix 1) and the variance inflation factors (Appendix 2) for our five models in Tables 12.4 and 12.5. Examining Appendix 1, we see that the largest correlation among our independent variables is –0.67 between IDV and PDI. The next largest correlation is between PDI and LN_GDP (–0.63). Since PDI is a partner in the two largest correlations in our sample, we exclude this independent variable in robustness tests described in Table 12.6.

In Table 12.4, we report the locally invariant test statistic of Baltagi and Wu (1999) (LBI). The LBI for our models in Table 12.4 are all not close to 2. This may suggest that serial correlation is a problem for these models. However, while commonly used, the significances of LBI are difficult to interpret.

Because the dynamic panel estimator of Blundell and Bond (1998) assumes that there is no autocorrelation in the idiosyncratic errors, we also report in Table 12.5 Arellano-Bond test statistics. The Arellano-Bond test for zero correlation in first-differenced errors is significant for differences of one lag. However, because the first difference of white noise is necessarily autocorrelated, it is only necessary to be concerned with second and higher autocorrelations. Examining the Arellano-Bond test across the models in Table 12.5, we see that the values for lags 2 and 4 are also consistently significant. In subsequent robustness tests described below, we also control for other lags in order to control for serial correlation.

12.4.4 Discussion of Initial Results

With the exception of Model 4 in Table 12.4, EQUITY_PREMIUM is negatively significant in every model of both Tables 12.4 and 12.5. Both modelings point toward a negative association of the size of the expected equity premium with financial architecture (the ratio of stock market size to banking size). It is also intuitively reasonable to expect a lower equity premium in more market-oriented nations. This is consistent with the results of Aggarwal and Goodell (2011a, b).

With the exception of Model 4 in Table 12.4, STOCK_VOLATILITY is negatively significant in every model of both Tables 12.4 and 12.5. CONCENTRATION is positively significant in every model in which it is present in both Tables 12.4 and 12.5. REGION_EUROPE is negatively significant in Models 3 and 4 in Table 12.4, but is not significant in any models in Table 12.5. LN_GPD is positively significant in Models 4 and 5 of both Tables 12.4 and 12.5. SPREAD is negatively significant in Models 4 and 5 in Table 12.5, but is not significant in any models in Table 12.4. ANTI_SELF_DEAL is positively

Table 12.6 Cross-national determinants of financial architecture: dynamic panel estimation controlling for three lags

Dependent variable: financial architecture	Model				
	1	2	3	4	5
INTERCEPT	0.23^{**} (0.041)	0.36^{***} (0.001)	0.86^{*} (0.093)	1.71 (0.946)	–2.76 (0.655)
LAG1_FIN_ARCH	1.34^{***} (0.000)	1.36^{***} (0.000)	1.31^{***} (0.000)	1.20^{***} (0.000)	1.11^{***} (0.000)
LAG2_FIN_ARCH	-0.80^{**} (0.019)	-0.76^{**} (0.012)	-0.70^{***} (0.000)	-0.65^{***} (0.000)	-0.62^{***} (0.000)
LAG3_FIN_ARCH	0.26 (0.121)	0.25^{*} (0.069)	0.26^{***} (0.000)	0.23^{***} (0.002)	0.19^{**} (0.030)
EQUITY_PREMIUM	-0.00^{**} (0.014)	–0.00 (0.147)	-0.00^{***} (0.002)	-0.00^{***} (0.000)	-0.00^{***} (0.000)
INFLATION_VOLATILITY		0.00 (0.684)	0.00 (0.708)	0.00 (0.437)	0.00 (0.723)
STOCK_VOLATILITY		-0.01^{***} (0.000)	-0.01^{***} (0.000)	-0.01^{*} (0.052)	–0.01 (0.180)
CONCENTRATION			0.91^{***} (0.000)	0.97^{***} (0.001)	1.28^{***} (0.000)
REGION_EUROPE			0.08 (0.629)	0.33 (0.490)	–0.31 (0.766)
LN_GDP			–0.06 (0.248)	0.04 (0.834)	0.45 (0.366)
UAI				0.00 (0.989)	–0.00 (0.935)
PDI				0.01 (0.637)	0.02 (0.454)
MAS				–0.01 (0.300)	–0.02 (0.348)
IDV				–0.00 (0.905)	–0.00 (0.983)
SPREAD				-0.12^{**} (0.013)	-0.09^{**} (0.025)

(*continued*)

Table 12.6 (continued)

Dependent variable: financial architecture	Model				
	1	2	3	4	5
ANTI_SELF_DEAL					–2.02 (0.227)
Observations/groups	291/41	291/41	290/41	290/41	290/41
Wald chi square	644.37*** (0.000)	1026.17*** (0.000)	1631.51*** (0.000)	569.15*** (0.000)	244.79*** (0.000)
AR tests	–1.04(0.298)	–1.09(0.277)	–1.74(0.082)*	–1.90(0.057)*	–1.84(0.066)*
	–0.57(0.566)	–0.77(0.442)	–0.71(0.475)	–1.03(0.301)	–0.80(0.422)
	0.204(0.839)	0.42(0.674)	0.00(0.999)	–0.30(0.766)	0.05(0.963)
	1.37(0.170)	1.37(0.171)	1.06(0.289)	1.50(0.133)	1.28(0.201)

This table reports results of tobit regressions for 41 countries, for 1996–2006. FIN_ARCH is stock market capitalization divided by domestic assets of deposit money banks from from Beck, Demirguc-Kunt, Levine (2000), EQUITY PREMIUM is ex ante equity premium estimations from the data from I/B/E/S; INFLATION_VOLATILITY is variance of preceding 5 years of inflation from *International Financial Statistics*; STOCK_VOLATILITY is the annualized standard deviation of monthly equity returns of respective Morgan Stanley Country Index; CONCENTRATION is a Herfindahl index of equity market concentration created for this chapter REGION_EUROPE is a dummy variable that is assigned "1" if market is in Europe and "0" otherwise; LN_GDP is natural log of real GDP per capita;; UAI, PDI MAS, and IDV are cultural dimensions of Hofstede (2001); SREAD is the difference between Moody's Baa corporate bond yield and 10-year Treasury yield from *St Louis Federal Reserve*; ANTI_SELF_DEAL is the index of measures against self-dealing from Djankov et al. (2008)

Variance inflation factors are less than 10 for all variables and models. P values in parentheses. Random-effects, EC2SLS estimates reported for Model 1.
*** Significance at 1 % level, ** significance at 5 % level, * significance at 10 % level Blundell-Bond estimates reported, controlling for three lags

significant (at 10 %) in Model 5 in Table 12.4 but is not significant in Table 12.5. Many of the cultural dimensions of Hofstede (2001) are significant in Table 12.4. For instance, PDI, IDV, and MAS are significant in every model in which they are present, while UAI is significant in Model 4. However, none of these cultural variables are significant in Table 12.5.

Summarizing these differences, both static and dynamic estimation methods find evidence of a negative association of financial architecture with the expected equity premium. This suggests that investors demand less return for holding equity in more market-oriented countries. Both static and dynamic estimation methods find evidence of a negative association of financial architecture with stock volatility. This suggests that less volatile markets are associated with being more market-oriented. Both static and dynamic estimation methods find evidence of a positive association of financial architecture with equity market concentration. This suggests that, despite the widely dispersed, market orientation of the United States, overall, more market-oriented countries have greater concentration of equity ownership into fewer firms. Both static and dynamic estimation methods also find evidence of a positive association of financial architecture with national wealth, suggesting that wealthier countries are more market oriented.

While the two estimation procedures used in Tables 12.4 and 12.5 yield very similar results with respect to a number of independent variables, some differences also exist. For instance, SPREAD is significantly negative in Table 12.5 but not significant in Table 12.4. SPREAD changes value each year but is the same across countries. We consider that the difference between Tables 12.4 and 12.5 with respect to SPREAD may be due to a difference in emphasis between these tables with regard to cross-sectional versus time series effects, with the Baltagi-Li estimator emphasizing cross-sectional effects and the Blundell-Bond estimator emphasizing time series effects. Similarly the cultural dimensions of Hofstede (2001) – as well as ANTI_SELF_DEALING – do not change over our time period. However these variables vary widely across countries. And so the result that these variables are significant in Table 12.4 but not significant in Table 12.5 is consistent with the notion that Table 12.4 emphasizes cross-sectional differences, while Table 12.5 emphasizes time series differences.

12.4.5 Robustness Checks

As noted above, because the dynamic panel estimator of Blundell and Bond (1998) assumes that there is no autocorrelation in the idiosyncratic errors, we also report in Table 12.5 Arellano-Bond test statistics. The Arellano-Bond test for zero correlation in first-differenced errors is significant for differences of one lag. However, because the first difference of white noise is necessarily autocorrelated, it is only necessary to be concerned with second and higher autocorrelations that are higher than 1 for the models in Table 12.5. However, examining the Arellano-Bond test across the models in Table 12.5, the values for lags 2 and 4 are also consistently significant. Therefore, in Table 12.6, we also control for the first three lags of our dependent variable.

Additionally, in order to avoid multicollinearity, we also employ, in Table 12.7, a two-stage model for some variables in order to lower multicollinearity. We note that the Pearson correlation coefficients between IDV and PDI as reported in Appendix 1 is –0.67. Further, the correlation between PDI and LN_GDP is –0.63. We first regress the respective independent variable (either IDV or PDI) on our measure of wealth along with some other cultural variables. Specifically, we regress each variable on wealth (LN_GDP). We then use the residuals from these regressions as our independent variables in Eq. 12.3:

$$\mathrm{IDV}_{it} = \alpha_i + \sum \beta_1 * LN\_GDP_{it} + e_{it} \tag{12.5}$$

$$\mathrm{PDI}_{it} = \alpha_i + \sum \beta_1 * LN\_GDP_{it} + e_{it} \tag{12.6}$$

We then include the residuals from Eqs. 12.5 and 12.6 as substitute independent variables for IDV and PDI, respectively. We refer to these variables as RESID_IDV and RESID_PDI, respectively. This procedure lowers the correlation between IDV (RESID_IDV) and PDI (RESID_PDI) to zero. The correlation between RESID_IDV and RESID_PDI is now lowered to –0.45.

12.4.6 Discussion of Robustness Tests

The models in Table 12.6 are the same as those in Tables 12.4 and 12.5, with the exception of including three lags of the dependent variable. The results of Table 12.6 substantially corroborate the results of Tables 12.5 and 12.4. EQUITY__PREMIUM is negatively significant in every model. CONCENTRATION is positively significant. STOCK_VOLATILITY is generally negatively significant, as in Tables 12.4 and 12.5. However, unlike the previous tables this variable is not significant in Model 5. Like Table 12.5, SPREAD is negatively significant. Overall the results of Table 12.6 suggest that the results of Table 12.5 are not driven by serial correlation in higher lags and/or excessive multicollinearity.

Table 12.7 shows the results of the most comprehensive model (Model 5 in Tables 12.4, 12.5, and 12.6) using three lags of the dependent variable and replacing IDV and PDI with RESID_IDV and RESID_PDI. As seen in Table 12.7, controlling for correlation between PDI and IDV by orthogonalizing both to LN_GDP results in little change to the estimates for our most comprehensive model. Model 1 in Table 12.7 is very similar to Model 5 in Table 12.4. Similarly, Model 2 in Table 12.7 is very similar to Model 5 in Table 12.5.

12.4.7 Discussion of Differences in the Results of the Two Different Modelings

As noted above, while the two estimation procedures used in this study yield very similar results with respect to a number of independent variables, some

Table 12.7 Cross-national determinants of financial architecture: comparison of static and dynamic panel modeling

Dependent variable: financial architecture	Model	
	1	2
INTERCEPT	0.88* (0.088)	0.86 (0.774)
LAG1_FIN_ARCH		1.11*** (0.000)
LAG2_FIN_ARCH		–0.62*** (0.000)
LAG3_FIN_ARCH		0.19** (0.030)
EQUITY_PREMIUM	–0.03** (0.021)	–0.00***(0.000)
INFLATION_VOLATILITY	–0.00 (0.133)	0.00 (0.723)
STOCK_VOLATILITY	–0.01** (0.050)	–0.01 (0.180)
CONCENTRATION	0.91*** (0.000)	1.28*** (0.000)
REGION_EUROPE	–0.17 (0.215)	–0.31 (0.766)
LN_GDP	0.04 (0.419)	0.18 (0.486)
UAI	–0.00 (0.199)	–0.00 (0.935)
RESID_PDI	0.01*** (0.000)	0.02 (0.454)
MAS	–0.00* (0.054)	–0.02 (0.348)
RESID_IDV	0.01*** (0.007)	–0.00 (0.983)
SPREAD	0.04 (0.405)	–0.09** (0.025)
ANTI_SELF_DEAL	0.46* (0.071)	–2.02 (0.227)
Observations/groups	395/41	290/41
Wald chi square	1453.14*** (0.000)	850.11*** (0.000)
Hausman	3.47 (0.838)	
LBI	0.97	
R-square	(0.04, 0.14,0.11)	
AR tests		–1.84(0.066)
		–0.80(0.422)
		0.05(0.963)
		1.28(0.201)

This table reports results of tobit regressions for 41 countries, for 1996–2006. FIN_ARCH is stock market capitalization divided by domestic assets of deposit money banks from Beck, Demirguc-Kunt, Levine (2000), EQUITY PREMIUM is ex ante equity premium estimations from the data from I/B/E/S; INFLATION_VOLATILITY is variance of preceding 5 years of inflation from *International Financial Statistics*; STOCK_VOLATILITY is the annualized standard deviation of monthly equity returns of respective Morgan Stanley Country Index; CONCENTRATION is a Herfindahl index of equity market concentration created for this chapter; REGION_EUROPE is a dummy variable that is assigned "1" if market is in Europe and "0" otherwise; LN_GDP is natural log of real GDP per capita; UAI, PDI MAS, and IDV are cultural dimensions of Hofstede (2001); SREAD is difference between Moody's Baa corporate bond yield and 10 year Treasury yield from *St Louis Federal Reserve*; ANTI_SELF_DEAL is the index of measures against self-dealing from Djankov et al. (2008)

Variance inflation factors are less than 10 for all variables and models. P values in parentheses. Random-effects, EC2SLS estimates reported for Model 1. Blundell-Bond dynamic panel estimates (Modified to control for three lags) reported in Model 2. *** Significance at 1 % level, ** significance at 5 % level, * significance at 10 % level differences in the results using Blundell-Bond dynamic panel modeling and Baltagi-Li static panel modeling persist. For instance, the variable SPREAD which has a time series aspect but no cross-sectional aspect is significantly negative in the estimation of Blundell-Bond but not significant using Baltagi-Li EC2SLS estimator. We consider that the difference between these two estimators with respect to SPREAD may be due to a difference in emphasis between these estimators with regard to cross-sectional versus time series effects, with the Baltagi-Li estimator emphasizing cross-sectional effects and the Blundell-Bond estimator emphasizing time series effects. Similarly the cultural dimensions of Hofstede (2001) and our index of measures against self-dealing have no time series aspect but these variables vary widely across countries. These variables are significant using Baltagi-Li estimation but not significant using Blundell-Bond estimation. These differences in the results using differing estimation procedures are consistent with the Baltagi-Li estimator emphasizing cross-sectional differences, while Blundell-Bond estimator places greater emphasis on time series differences, due to the fact that the Blundell-Bond modeling includes as an independent variable the one-period lag in the dependent variable and so controls for fixed effects across one-period lags.

12.5 Conclusions

There has been a veritable explosion of studies in a wide variety of areas that employ panel data econometric methodology in recent years. Analysis of panel data allows independent variables to vary both cross-sectionally and across time, while panel data econometrics correct for cross-correlations between time series and cross-sectional error terms. In addition, there has now been the introduction of a number of new refinements in analyzing panel data (Petersen 2009; Wooldridge 2010), all maintaining the goal of accounting for cross-correlation between time series and cross-sectional error terms when assessing coefficient significance.

However, in the study of economic and financial panel data, it is often important to assess the differential impact of time series versus cross-sectional effects; but panel data techniques are unclear how this may be accomplished. In other words, panel data methodologies typically do not inform us fully regarding which effect (time series or cross-sectional) is more important or more dominant within particular data sets or contexts. In this chapter we employ two contrasting estimation procedures, which, respectively, emphasize cross-sectional versus time series differences, to clarify the impacts of these two influences. We undertake this comparison of econometric methods within a finance-related context which takes into account possible endogeneity.

In this chapter we show that a number of independent variables are significant in partially determining financial architecture using either or both of the estimation procedures of Baltagi and Li (1992) and Blundell and Bond (1998). Overall, we find a negative association of the equity premium with financial architecture. A smaller equity premium inclines societies toward being more market oriented. A Herfindahl Index for equity market concentration is positively significant. Generally, when market capitalization is concentrated in fewer firms, societies are more market oriented. We also find that stock volatility is generally negatively significant. This suggests that stock volatility inclines societies away from markets and toward banks.

However, of greater interest for this study, we find that some of our independent variables impact cross-national differences in financial architecture according to Baltagi and Li (1992) estimation, while other independent variables impact financial architecture according to Blundell and Bond (1998) estimation. While the two estimation procedures yield very similar results with respect to a number of independent variables, some differences exist also.

US Corporate bond spreads negatively determine financial architecture according to Blundell and Bond (1998) estimation but not according to Baltagi and Li (1992) estimation. US Corporate bond spreads change value each year but have the same value across countries. Similarly some measures that change across countries but do not change across time, such as the cultural dimensions of Hofstede (2001) as well as the index of measures against self-dealing, are significant determinants of financial architecture according to Baltagi-Li estimation but not according to Blundell-Bond estimation. This is consistent with different estimation techniques placing differing emphasis on cross-sectional and time series effects, with the Baltagi-Li estimator emphasizing cross-sectional effects and the Blundell-Bond estimator emphasizing time series effects.

These are critical findings that have important implications for the use of panel data estimation procedures, especially where it is important to differentiate between time series and cross-sectional influences. Our results show that using the two panel estimation procedures used here together can better differentiate between time series and cross-sectional variations in panel data. Thus, our results should be of much interest to scholars especially in finance and economics.

Appendix 1: Pearson Correlation Coefficients

		1	2	3	4	5	6	7	8	9	10	11	12	13
1	FIN_ARCH	1												
2	EQUITY_PREMIUM	0.03	1											
3	INFLATION_VOLATILITY	−0.07	−0.16	1										
4	MARKET_VOLATILITY	−0.05	−0.10	0.11	1									
5	CONCENTRATION	0.03	0.00	0.04	0.30	1								
6	REGION_EUROPE	−0.19	0.12	−0.10	−0.14	0.18	1							
7	LN_GDP	0.08	0.00	−0.09	−0.42	−0.02	0.38	1						
8	UAI	−0.32	0.07	0.06	0.13	0.15	0.18	−0.12	1					
9	PDI	0.04	0.08	0.07	0.27	−0.18	−0.35	−0.63	0.18	1				
10	MAS	−0.10	−0.07	−0.01	−0.10	−0.21	−0.16	−0.06	0.15	0.06	1			

(*continued*)

		1	2	3	4	5	6	7	8	9	10	11	12	13
11	IDV	0.10	−0.05	−0.07	−0.37	0.00	0.44	0.51	−0.17	−0.67	0.12	1		
12	SPREAD	−0.05	0.14	−0.08	0.06	−0.10	0.02	0.00	0.02	−0.01	0.01	0.01	1	
13	ANTI_SELF_DEAL	0.29	0.05	−0.09	−0.09	−0.20	−0.48	0.05	−0.49	0.02	0.00	−0.07	−0.01	1

Appendix 2: Estimation Procedure of Baltagi-Li EC2SLS

Following Baltagi (1981), consider the jth structural equation:

$$\begin{aligned} y_j &= Y_j\alpha_j + X_j\beta_j + u_j = Z_j\delta_j + u_j, j = 1, 2, \ldots, M, \\ y_j &= NT \times 1; Y_j = NT \times \left(M_j - 1\right); X_j = NT \times K_j; \\ Z_j &= \left(Y_j X_j\right); \delta_j' = \left(\alpha_j', \beta_j'\right); n_j = M_j + K_j - 1 \end{aligned} \tag{12.7}$$

The additive error components structure is given by

$$u_j = Z_u u_j + Z_\lambda \lambda_j + v_j, \mathrm{j} = 1, 2, \ldots, \mathrm{M}, \tag{12.8}$$

where $Z_u = I_N \bigotimes e_T$, $\mathrm{Z}_\lambda = e_N \bigotimes I_T$

I_N and I_T are identifying matrices of order N and T, respectively. e_N and e_N are vectors of ones of order N and T, respectively.

$u_j' = (u_{1j}, u_{2j}, \ldots, u_{Nj})$, $\lambda_j' = (\lambda_{1j}, \lambda_{2j}, \ldots, \lambda_{Tj})$, and $v_j' = (v_{1j}, v_{2j}, \ldots, v_{NTj})$ are random vectors with means of zero and the following covariance matrix:

$$E\begin{pmatrix} u_j \\ \lambda_j \\ v_j \end{pmatrix}\left(u_1'\lambda_1'v_1'\right) = \begin{bmatrix} \sigma^2_{u_{jl}} I_N & 0 & 0 \\ 0 & \sigma^2_{\lambda_{jl}} I_T & 0 \\ 0 & 0 & \sigma^2_{v_{jl}} I_{TN} \end{bmatrix} \tag{12.9}$$

for j and l = 1,2,. . . .M. As noted by Baltagi (1981), this implies that the covariance matrix between jth and lth structural equation is

$$\Sigma_{jl} = E\left(u_j u_l'\right) = \sigma^2_{u_{jl}} A + \sigma^2_{\lambda_{jl}} B + \sigma^2_{v_{jl}} I_{NT}, \tag{12.10}$$

where $A = I_N \bigotimes e_T e'_T$, $B = e_N e'_N \bigotimes I_T$

It follows that

$$\Sigma_{ij} = \sigma^2_{3_{jl}} \frac{J_{NT}}{NT} + \sigma^2_{1_{jl}} \left(\frac{A}{T} - \frac{J_{NT}}{NT}\right) + \sigma^2_{2_{jl}} \left(\frac{B}{N} - \frac{J_{NT}}{NT}\right) + \sigma^2_{v_{jl}} Q \tag{12.11}$$

where

$Q = I_{NT} - \frac{A}{T} - \frac{B}{N} + \frac{J_{NT}}{NT}$ and $J_{NT} = e_{NT}e'_{NT}$

also

$\sigma^2_{3_{jl}} = \sigma^2_{v_{jl}} + N\sigma^2_{\lambda_{jl}} + T\sigma^2_{u_{jl}}$ and $\sigma^2_{1_{jl}} = \sigma^2_{v_{jl}} + T\sigma^2_{u_{jl}}$ as well as

$$\sigma^2_{2_{jl}} = \sigma^2_{v_{jl}} + N\sigma^2_{\lambda_{jl}}$$

$\sigma^2_{v_{jl}}, \sigma^2_{1_{jl}}, \sigma^2_{2_{jl}}, \sigma^2_{1_{jl}}$ are the characteristic roots of $\sum_{ij}$ of multiplicity (N−1)(T−1), N−1, T−1, and 1, respectively (Baltagi 1981; Nerlove 1971).

If the rank condition $K \geq K_1 + M_1 - 1$ holds for Eq. 12.7, then we can apply transformation Q_h to Eq. 12.7 to get

$$y_1^{(h)} = Y_1^{(h)}\alpha_1 + X_1^{(h)}\beta_1 + u_1^{(h)} = Z_1^{(h)}\delta_1 + u_1^{(h)} \tag{12.12}$$

$y_1^{(h)} = Q_h y_1, \ldots,; u_1^{(h)} = Q_h u_1$ for h = 1,2,3.

As noted by Baltagi (1981), $u_1^{(h)}$ has a covariance matrix which is a scalar times an identity matrix.

However, Baltagi (1981) applies a 2SLS procedure to Eq. 12.12 in order to correct for simultaneity:

$$X^{h'}y_1^h = X^{h'}Z_1^h\delta_1 + X^{h'}u_1^h \tag{12.13}$$

The resulting estimator of δ_1 is

$$\hat{\delta}^{(h)}_{1,2SLS} = \left[Z_1^{(h)'}P_{\chi^{(h)}}Z_1^h\right]^{-1}\left[Z_1^{(h)'}P_{\chi^{(h)}}y_1^h\right] \tag{12.14}$$

where

$$P_{\chi^h} = X^{(h)}X^{(h)}\left(X^{(h)'}X^{(h)}\right)^{-1}X^{(h)'}$$

The variance components are then estimated by

$$\sigma_{11}^{(h)2} = \left(y_1^{(h)} - Z_1^{(h)}\hat{\delta}^{(h)}_{1,2SLS}\right)'\left(y_1^{(h)} - Z_1^{(h)}\hat{\delta}^{(h)}_{1,2SLS}\right)/n(h) \tag{12.15}$$

As noted by Baltagi (1981), since δ_1 is common to all three transformations of Eq. 12.13, we can apply an Aitken estimation to the following sets of equations (Baltagi 1981):

$$\begin{pmatrix} X^{(1)'}y_1^{(1)} \\ X^{(2)'}y_1^{(2)} \\ X^{(3)'}y_1^{(3)} \end{pmatrix} = \begin{pmatrix} X^{(1)'}Z_1^{(1)} \\ X^{(2)'}Z_1^{(2)} \\ X^{(3)'}Z_1^{(3)} \end{pmatrix}\delta_1 + \begin{pmatrix} X^{(1)'}u_1^{(1)} \\ X^{(2)'}u_1^{(2)} \\ X^{(3)'}u_1^{(3)} \end{pmatrix} \tag{12.16}$$

This results in the following estimator:

$$\hat{\delta}_{1,GLS} = \left(\sum_{h=1}^{3} \left[Z_1^{(h)'} P_{\chi^h} Z_1^{(h)} / \sigma_{11}^{(h)2} \right] \right)^{-1} \left(\sum_{h=1}^{3} \left[Z_1^{(h)'} P_{\chi^h} y_1^{(h)} / \sigma_{11}^{(h)2} \right] \right) \tag{12.17}$$

Substituting the estimate of $\hat{\sigma}_{11}^{(h)2}$ from Eq. 12.15 into Eq. 12.17 yields

$$\hat{\delta}_{1,EC2SLS} = \left(\sum_{h=1}^{3} \left[Z_1^{(h)'} P_{\chi^h} Z_1^{(h)} / \hat{\sigma}_{11}^{(h)2} \right] \right)^{-1} \left(\sum_{h=1}^{3} \left[Z_1^{(h)'} P_{\chi^h} y_1^{(h)} / \hat{\sigma}_{11}^{(h)2} \right] \right)$$

Appendix 3: Estimation Procedure of Blundell and Bond (1998) System GMM Estimation

We consider the first-order autoregressive panel data model:

$$y_{it} = \alpha y_{i,t-1} + u_{it} \tag{12.18}$$

In this case, $u_{it} = \eta_i + v_{it}$
where i = 1, . . . N and t = 2, ... T
As described by Bun and Windmeijer (2010), following Blundell and Bond (1998), we assume that η_i and v_{it} have the error components structure

$$E(\eta_i) = 0; E(v_{it}) = 0; E(v_{it}\eta_i) = 0 \tag{12.19}$$

Additionally

$$E(v_{it}v_{is}) = 0 \text{ for } t \neq s \tag{12.20}$$

Additionally the initial condition satisfies

$$E(y_{i1}v_{it}) = 0. \tag{12.21}$$

It then follows that the following (T – 1)(T – 2)/2 linear moments conditions are valid:

$$E\left(y_i^{t-2}\Delta u_{it}\right) \text{ t} = 3, \ldots, \text{T}, \tag{12.22}$$

where

$$y_i^{t-2} = (y_{i1}, y_{i2}, \ldots, y_{it-2})' \text{ and } \Delta u_{it} = u_{it} - u_{it-1} = \Delta y_{it} - \alpha \Delta y_{it-1}$$

Defining

$$Z_{di} = \begin{bmatrix} y_{i1} & 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & y_{i1} & y_{i2} & \dots & 0 & \dots & 0 \\ . & . & . & \dots & . & \dots & . \\ 0 & 0 & 0 & \dots & y_{i1} & \dots & y_{iT-2} \end{bmatrix}; \Delta u_i = \begin{bmatrix} u_{i3} \\ u_{i4} \\ . \\ . \\ . \\ \Delta u_{iT} \end{bmatrix}$$

This leads to a more efficient representing of the moment conditions:

$$E\left(Z'_{di}\Delta u_i\right) = 0 \tag{12.23}$$

Following Arellano and Bover (1995), the GMM estimator of α is given by

$$\alpha_d = \frac{\Delta y'_{-1} Z_d W_N^{-1} Z'_d \Delta y}{\Delta y'_{-1} Z_d W_N^{-1} Z'_d \Delta y_{-1}}$$

where $\Delta y = (\Delta y'_1, \Delta y^1_2 \dots \Delta y'_N)'$, $\Delta y_i = (\Delta y_{i3} \Delta y_{i4}, \dots, \Delta y_{iT})'$, Δy_{-1} the lagged version of Δy, and $Z_d = (Z'_{d1}, Z'_{d2}, \dots, Z'_{dN})$ ', and W_N is a weight determining the efficiency of the GMM estimator.

α_d is a GMM estimator in the differenced model. As noted by Bun and Windmeijer (2010), this estimator is referred to as the difference GMM estimator.

However, following Blundell and Bond (1998), who follow on Arellano and Bover (1995), we also assume additional conditions:

$$E(\eta_i \Delta y_{i2}) = 0 \tag{12.24}$$

This condition holds when the process is mean stationary:

$$y_{i1} = \frac{\eta_i}{1-\alpha} + \varepsilon_i \tag{12.25}$$

Additionally, $E(\varepsilon_i) = E(\varepsilon_i \eta_i) = 0$.

Further, if Eqs. 12.19, 12.20, 12.21, and 12.24 hold, then further $(T-1)(T-2)/2$ linear moments conditions are valid:

$$E\left(u_{it}\Delta y_i^{t-1}\right) = 0 \quad \text{for } t = 3, \dots, T, \tag{12.26}$$

where

$$y_i^{t-1} = (\Delta y_{i2}, \Delta y_{i3}, \dots, \Delta y_{it-1})'$$

Defining

$$Z_{di} = \begin{bmatrix} \varDelta y_{\mathrm{i}2} & 0 & 0 & \dots & 0 & \dots & 0 \\ 0 & \varDelta y_{\mathrm{i}2} & \varDelta y_{\mathrm{i}3} & \dots & 0 & \dots & 0 \\ . & . & . & \dots & . & \dots & . \\ 0 & 0 & 0 & \dots & \varDelta y_{i2} & \dots & \varDelta y_{iT-1} \end{bmatrix}; \Delta u_i = \begin{bmatrix} u_{i3} \\ u_{i4} \\ . \\ . \\ . \\ \Delta u_{iT} \end{bmatrix}$$

The above moment conditions can be written as

$$E\left(Z_{li}^{'} u_i\right) = 0 \tag{12.27}$$

The GMM estimator based on these estimates is then given by

$$\alpha_l = \frac{y_{-1}^{'} Z_l W_N^{-1} Z_l^{'} y}{y_{-1}^{'} Z_l W_N^{-1} Z_l^{'} y_{-1}}$$

α_l is a GMM estimator in the levels model. This is referred to as the levels' GMM estimator (Bun and Windmeijer 2010).

The full set of linear moments under assumptions (Eqs. 12.19, 12.20, 12.21, and 12.24) is

$$E\left(y_i^{t-2} \Delta u_{it}\right) = 0 \tag{12.28}$$

$$E\left(u_{it} \Delta y_{i,t-1}\right) = 0$$

for $t = 3, \dots, \mathrm{T}$

Following Bun and Windmeijer (2010), this can be resolved as

$$E\left(Z_{si}^{'} p_i\right) = 0, \tag{12.29}$$

where

$$Z_{si} = \begin{bmatrix} \mathrm{Z}_{di} & 0 & \dots & 0 \\ 0 & \varDelta y_{\mathrm{i}2} & \dots & 0 \\ . & . & . & . \\ 0 & 0 & \dots & \varDelta y_{iT} \end{bmatrix}; p_i = \begin{bmatrix} \Delta u_i \\ u_i \end{bmatrix}$$

As noted by Wind, the GMM estimator based on these conditions is

$$\alpha_s = \frac{q_{-1}^{'} Z_s W_N^{-1} Z_s^{'} q}{q_{-1}^{'} Z_s W_N^{-1} Z_s^{'} q_{-1}}$$

where

$$q_i = \left(\Delta y_i^{'} y_i^{'}\right)^{'}$$

α_s is the system GMM estimator of Blundell and Bond (1998).

Acknowledgments Authors are grateful to John Hunter, Maurice Peat, Min Qi, an anonymous reviewer, and others for their kind advice on earlier versions but remain responsible for the contents including any remaining errors.

References

Aggarwal, R., & Goodell, J. W. (2009). Markets and institutions in financial intermediation: National characteristics as determinants. *Journal of Banking and Finance, 33*, 1770–1780.
Aggarwal, R., & Goodell, J. W. (2011a). International variations in expected equity premia: Role of financial architecture and governance. *Journal of Banking and Finance, 35*, 3090–3100.
Aggarwal, R., & Goodell, J. W. (2011b). Variations in emerging market equity premia: Impact of financial architecture. *Journal of the Asia Pacific Economy, 16*, 313–330.
Aggarwal, R., & Zhao, S. (2009). The diversification discount puzzle: Empirical evidence for a transactions cost resolution. *Financial Review, 44*, 113–135.
Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies, 58*, 277–297.
Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics, 68*, 29–51.
Balestra, P., & Varadharajan-Krishnakumar, J. (1987). Full information estimations of a system of simultaneous equations with error components structure. *Econometric Theory, 3*, 223–246.
Baltagi, B. H. (1981). Simultaneous equations with error components. *Journal of Econometrics, 17*, 189–200.
Baltagi, B. H. (2001). *Econometric analysis of panel data*. New York: Wiley.
Baltagi, B. H., & Li, Q. (1992). A note on the estimation of simultaneous equations with error components. *Econometric Theory, 8*, 113–119.
Baltagi, B. H., & Li, Q. (1994). Estimating error component models with general MA (q) disturbances. *Econometric Theory, 10*, 396–408.
Baltagi, B. H., & Li, Q. (1995). Testing AR(1) against MA(1) disturbances in an error component model. *Journal of Econometrics, 68*, 133–151.
Baltagi, B. H., & Wu, P. X. (1999). Unequally spaced panel data regressions with Ar(1) disturbances. *Economic Theory, 15*, 814–823.
Beck, T., Demirguc-Kunt, A., & Levine, R. (2000). A new database on financial development and structure. *World Bank Economic Review, 14*, 597–605.
Bhattacharya, S., & Thakor, A. V. (1993). Contemporary banking theory. *Journal of Financial Intermediation, 3*, 2–50.
Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics, 87*, 115–143.
Bun, M. J. G., & Windmeijer, F. (2010). The weak instrument problem of the system GMM estimator in dynamic panel data models. *Econometrics Journal, 13*, 95–126.
Coase, R. H. (1960). The problem of social cost. *Journal of Law and Economics, 3*, 1–44.
Diamond, D. W. (1984). Financial intermediation and delegated monitoring. *Review of Economic Studies, 51*, 393–414.
Djankov, S., La Porta, R., Lopez-de-Silvanes, F., & Shleifer, A. (2008). The law and economics of self-dealing. *Journal of Financial Economics, 88*, 430–465.
Ergungor, O. E. (2004). Market- vs. bank-based financial systems: Do rights and regulations really matter? *Journal of Banking and Finance, 28*, 2869–2887.
Fligstein, N. (2001). *The architecture of markets: An economic sociology of twenty-first-century capitalist societies*. Princeton: Princeton University Press.
Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica, 50*, 1029–1054.
Hart, O. (1995). *Firms, contracts, and financial structure*. New York: Oxford University Press.
Hart, O. (2001). Financial contracting. *Journal of Economic Literature, 39*, 1079–1100.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica, 6*, 1251–1271.
Hofstede, G. (2001). *Culture's consequences*. London: Sage.
Ibbotson, R. G., Diermeier, J. J., & Siegel, L. B. (2006). The demand for capital market returns: A new equilibrium theory. In W. N. Goetzmann & R. G. Ibbotson (Eds.), *The equity risk premium*. Oxford: Oxford University Press.
Kaufmann, D., Kraay, A., & Mastruzzi, M. (2008). Governance Matters Vii: Aggregate and Individual Governance Indicators, 1996-2007. *The World Bank*.
Kwok, C. C., & Tadesse, S. (2006). National culture and financial systems. *Journal of International Business Studies, 37*, 227–247.
Modigliani, F., & Perotti, E. (2000). Security markets versus bank finance: Legal enforcement and investor protection. *International Review of Finance, 1*, 81–96.
Moerman, G. A., & van Dijk, M. A. (2010). Inflation risk and international asset returns. *Journal of Banking and Finance, 34*, 840–855.
Morck, R., Yeung, B., & Yu, W. (2000). The information content of stock markets: Why do emerging markets have synchronous stock price movements? *Journal of Financial Economics, 58*, 215–260.
Myers, S. (1977). The determinants of corporate borrowing. *Journal of Financial Economics, 5*, 147–175.
Nerlove, M. (1971). A note on error components model. *Econometrica, 39*, 383–396.
North, D. C. (1990). *Institutions, institutional change, and economic performance*. Cambridge, UK: Cambridge University Press.
Penman, S. H. (2005). Discussion of "on Accounting-Based Valuation Formulae" and "Expected Eps and Eps Growth as Determinants of Value". *Review of Accounting Studies, 10*(2–3): 367–378.
Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies, 22*, 435–480.
Rajan, R. G. (1992). Insiders and outsiders: The choice between informed and arm's length debt. *Journal of Finance, 47*, 1367–1400.
Rajan, R. G., & Zingales, L. (1998). Which capitalism? Lessons from the East Asian crisis. *Journal of Applied Corporate Finance, 11*, 40–48.
Rajan, R. G., & Zingales, L. (2003). *Saving Capitalism from the Capitalists*. New York: Crown Business.
Shirai, S. (2004). Assessing the impact of financial and capital market reforms on firms' corporate financing patterns in India. *South Asia Economic Journal, 5*, 189–208.
Turner, P. (2003). Bond market development: What are the policy issues? In R. E. Litan, M. Pomerleano, & V. Sundararajan (Eds.), *The future of domestic capital markets in developing countries*. Washington: Brookings Institute Press.
Williamson, O. E. (1988). Corporate finance and corporate governance. *Journal of Finance, 43*, 567–592.
Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics, 126*, 25–51.
Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.

Does Banking Capital Reduce Risk? An Application of Stochastic Frontier Analysis and GMM Approach

13

Wan-Jiun Paul Chiou and Robert L. Porter

Contents

Abstract

In this chapter, we thoroughly analyze the relationship between capital and bank risk-taking. We collect cross section of bank holding company data from 1993 to 2008. To deal with the endogeneity between risk and capital, we employ stochastic frontier analysis to create a new type of instrumental variable. The unrestricted frontier model determines the highest possible profitability based solely on the book value of assets employed. We develop a second frontier based

W.-J.P. Chiou (✉)
Department of Finance and Law College of Business Administration, Central Michigan University, Mount Pleasant, MI, USA
e-mail: Chiou1P@cmich.edu

R.L. Porter
Department of Finance School of Business, Quinnipiac University, Hamden, CT, USA
e-mail: RLPorter@quinnipiac.edu; robert.porter1@quinnipiac.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_13,

on the level of bank holding company capital as well as the amount of assets. The implication of using the unrestricted model is that we are measuring the unconditional inefficiency of the banking organization.

We further apply generalized method of moments (GMM) regression to avoid the problem caused by departure from normality. To control for the impact of size on a bank's risk-taking behavior, the book value of assets is considered in the model. The relationship between the variables specifying bank behavior and the use of equity is analyzed by GMM regression. Our results support the theory that banks respond to higher capital ratios by increasing the risk in their earning asset portfolios and off-balance-sheet activity. This perverse result suggests that bank regulation should be thoroughly reexamined and alternative tools developed to ensure a stable financial system.

Keywords
Bank capital • Generalized method of moments • Stochastic frontier analysis • Bank risks • Bank holding companies • Endogeneity of variables

13.1 Introduction

Bank capital management has become an important issue to both commercial bankers and central bankers after the recent financial crisis. As the subprime mortgage debacle spread, the balkanized regulatory system designed over half a century ago appeared totally inadequate for today's complex financial system. In this study, we evaluate the role of capital in regulatory risk management by examining a wide-range bank holding company data. Historically, both theoretical and empirical papers on the relationship between capital and risk have produced mixed results.[1] Yet a new look at the role of bank capital in risk management is now critical if we are to protect the financial system of the twenty-first century.

This chapter fits into a long history of literature dealing in general with bank risk management and more specifically with the question of what constitutes an adequate level of bank capital. Our contribution consists of the analysis of a large cross section of bank holding companies over the years that the Basle Accords have been implemented. In addition, we introduce a unique, to our knowledge, method to exogenously model an instrumental variable for capital in a regression with risk.

One of the primary goals of bank regulators is to minimize the risk held on and off their balance sheets by financial institutions. In this way, the negative externalities of bank failures and the risk to taxpayers from losses from the federal bank safety net are avoided or reduced. A mandatory bank capital requirement is one of the most important tools historically used by regulators to stabilize the financial industry. The recent financial crisis, however, challenges the effectiveness of these

[1]For detailed discussion, please see Berger et al. (1995); Gatev et al. (2009); Hovakimian and Kane (2000); Shrieves and Dahl (1992); and VanHoose (2007).

mandatory capital requirements. The inherent characteristics of today's banking industry such as rapid financial innovation, high financial leverage, information asymmetry, liquidity creation, and the federal bank safety net all distort incentives and reward risk-taking. If maintaining a certain level of capital is viewed by bank managers only as a necessary evil, then critical questions emerge: How is capital related to specific measures of risk including credit risk, liquidity risk, interest rate risk, off-balance-sheet risk, market risk, and overall bank risk? Do higher levels of capital improve or lower the efficiency of banking? Answers to these questions will help to establish the role of capital in regulatory risk management.

Empirical studies of bank capital and bank risk, however, face an inherent problem. In order to measure the effect of the level of capital on bank risk-taking, it would be useful to regress risk, as the dependent variable, on capital, as the independent variable. However, there is an obvious endogeneity problem. The amount of risk a bank can undertake is dependent on its amount of capital, and the amount of capital needed is dependent on the amount of risk that a bank wants to undertake. In other words, they are jointly determined, much like price and quantity in a basic microeconomic analysis. The solution to this problem is normally either to use a simultaneous equation model or to use instrumental variables. However, a simultaneous equation model must be properly identified, and no one has yet been able to accomplish this in regard to risk and bank capital. Likewise, no one, to our knowledge, has yet found a true instrument for capital that is independent of risk.

We present a methodology for the development of an exogenous instrument for capital in a regression with risk by using stochastic frontier analysis. First, we determine the maximum possible income that can be achieved from a given level of assets. This is referred to as fitting an upper envelope. Such a frontier is obviously exogenous to any specific bank because it is determined by the data from all banks in the sample. The distance from the frontier to any specific bank's actual income can be considered a measure of bank inefficiency. Next, to develop the instrument for capital, we create a second frontier conditioned on bank capital as well as the amount of assets employed. The incremental inefficiency from the second frontier is a function of the bank's capital but independent of the bank's risk, and it is this incremental inefficiency that we propose to use as an instrument for capital.

Our analysis adds to the existing literature in several ways. First, we employ a large panel data set to consider the capital-risk relationship for a wider range of bank holding companies than typically reviewed. Previous empirical studies have commonly used market measures of risk. However, this approach necessarily limits the sample to publically owned banks or bank holding companies. In this study, we acknowledge the importance of small banks and bank holding companies, as well as the largest bank holding companies. This concern is significant since public policy related to the banking industry must consider a broad sample of banks and not only the largest organizations. As a result, we turn to the typical accounting measures of a bank's risk and utilize a large panel data set. In a second contribution, stochastic frontier analysis is applied to exogenously generate the effect of the use of capital in banking.

Finally, the results provide evidence of bank holding companies reacting to higher mandatory capital requirements by increasing the amount of risk the bank holding company accepts. We use the generalized method of moments and look at seven different measures of risk: credit risk, liquidity risk, interest rate risk, off-balance-sheet risk, market risk, overall risk, and leverage risk. In general, many results support the proposition that increased capital requirements reduce risk in BHCs. There are, however, some results that suggest the opposite – that BHCs increase risk as their capital ratios increase. This is obviously an important finding with major public policy implications. If the primary tool used by regulators to ensure a stable financial system is creating perverse results, then alternative tools must be developed.

The rest of the chapter is organized as follows. Section 13.2 summarizes the literature that deals with bank capital regulation. Section 13.3 presents our methodology, and Sect. 13.4 reports the data along with its univariate analysis. In Sect. 13.5, we present our empirical results. Section 13.6 concludes.

13.2 Literature Review

It has been argued that excessively high capital requirements can produce social costs through lower levels of intermediation. In addition, there can be unintended consequences of high capital requirements such as risk arbitrage (increasing risk to offset the increase in capital and thereby maintain the same return on capital), increased securitization, and increased off-balance-sheet activity, all of which could mitigate the benefits of increased capital standards. See Berger et al. (1995) and Santos (2001). The extent to which these unintended consequences played a role in our recent crisis is yet to be determined.

Moral hazard is high on the list of problems receiving attention in this post-financial crisis environment. The presence of a federal safety net creates moral hazard because bank management does not have to worry about monitoring by depositors (see Merton 1977; Buser et al. 1981; Laeven and Levine 2009). Absent depositor monitoring, banks are free to increase risk. If, however, deposit insurance and other elements of a federal safety net are reasons for increases in bank risk, why do they continue to exist? The answer lies in the contemporary theory of financial intermediation. It has been well established in the literature that there is need for both demand deposit contracts and the possibility of bank runs (Diamond and Dybvig 1983; Calomiris and Kahn 1991; Diamond and Rajan 2000; and Santos 2001). If the possibility of bank runs is needed, and bank runs are harmful, then government deposit insurance is an optimal solution. There is a related issue. Banks have a unique ability to resolve information asymmetries associated with risky loans. As a result, bank failures can produce a serious contraction in credit availability, especially among borrowers without access to public capital markets. The federal safety net is needed to avoid this credit contraction. Likewise, if a bank is considered "too big to fail," then the government will always bail the bank out, and there is no reason for bank management to limit risk.

It needs to be noted that not everyone is in agreement that the use of capital requirements is the best way to reduce risk in banking. Marcus (1984) and Keeley (1990) argue that a bank's charter value mitigates against increased risk. Banks operate in a regulated environment, and therefore, a charter to operate contains market power. Excessive risk increases the cost of financial distress, and this can cause a loss of charter value. Kim and Santomero (1988) argue that a simple capital ratio cannot be effective, and any ratio would need to have exactly correct risk weights in a risk-based system. Gorton and Pennacchi (1992) discuss "narrow banking" and propose splitting the deposit services of banks from the credit services. In other words, the financial system would include money market accounts and finance companies. The money market accounts would only invest in short-term high-quality assets and leave the lending to the finance companies that would not take in any deposits.

In Prescott (1997), he reviews the precommitment approach to risk management. Briefly, banks commit to a level of capital, and if that level proves to be insufficient, the bank is fined. This is used currently in the area of capital in support of a trading portfolio but cannot be used for overall capital ratios since a fine against a failed bank is not effective. Esty (1998) studies the impact of contingent liability of stockholders on risk. In the late nineteenth and early twentieth century, bank stockholders were subject to a call or an assessment for more money if needed to meet the claims on a bank. There was a negative relation between increases in risk and the possible call on bank stockholders. Calomiris (1999) makes a strong case for requiring the use of subordinated debt in bank capital structures. The need to issue unguaranteed debt and the associated market discipline would act as an effective limit to the amount of risk a bank would be able to assume. John et al. (2000) argue that a regulatory emphasis on capital ratios may not be effective in controlling risk. Since all banks will have a different investment opportunity set, an efficient allocation of funds must incorporate different risk-taking for different investment schedules. These authors go on to argue that senior bank management compensation contracts may be a more promising avenue to control risk using incentive-compatible contracts to achieve the optimal level of risk.

Marcus and Shaked (1984) show how Merton's (1977) put option pricing formula can be made operational and then used the results to estimate appropriate deposit insurance premium rates. The results of their empirical analysis indicated that the then current FDIC premiums were higher than was warranted by the ex ante default risk of the sample banks. This implies that banks are not transferring excessive risk to the deposit insurance safety net and capital regulation is effectively working.

Duan et al. (1992) address the question of the impact of fixed-rate versus risk-based deposit insurance premiums directly. The authors tested for specific risk-shifting behavior by banks. If banks were able to increase the risk-adjusted value of the deposit insurance premiums, then they had appropriated wealth from the FDIC. This is because the FDIC, at the time, could not increase the insurance premium even though risk had increased. Their empirical findings were that only 20 % of their sample banks were successful in risk-shifting behavior and therefore the problem was not widespread. This also implies that capital management has been effective.

Keeley (1992) empirically studied the impact of the establishment of objective capital-to-assets ratio requirements in the early 1980s. His evidence documents an increase in the book value capital-to-assets ratio of previously undercapitalized banks, and this, of course, was the goal of the new capital regulations. His study, however, is unable to confirm the same result when looking at the market value capital ratios. While the market value capital-to-assets ratios also increased, there was no significant difference between the undercapitalized banks and the adequately capitalized banks. Nevertheless, this was more evidence that capital regulation was working.

Hovakimian and Kane (2000) use the same empirical design as Duan et al. (1992) but for a more recent time period, and they obtain opposite results. They also start with the argument of Merton (1977) that the value of deposit insurance increases in asset return variance and leverage. They regress the change in leverage on the change in risk and find a positive rather than a negative coefficient. The coefficient must be negative if capital regulation forces banks to decrease leverage with increases in risk. In a second test, they regress the change in the value of the deposit insurance premium on the change in the asset return variance. Here again the coefficient must be negative (or zero) if there is any restraint. In this equation, the coefficient measures how much the bank can benefit from increasing the volatility of its asset returns. The option-model evidence presented shows that capital regulation has not prevented risk-shifting by banks and that it was possible for banks to extract a deposit insurance subsidy.

In Hughes et al. (2001), the authors study the joint impact of two functions of bank capital. First is the capital's influence on market value conditioned on risk, and second is its impact on production decisions incorporating endogenous risk. Efficient BHCs are determined according to frontier analysis, and then these BHCs are assumed to be value-maximizing firms. The conclusion is that these value-maximizing firms do achieve economies of scale, but the analysis of production must include capital structure and risk-taking.

Berger et al. (2008) note that US banks hold significantly more equity capital than the minimum amount required by regulators. Their evidence documents the active management of capital levels by BHCs including setting target levels of capital above regulatory minimums and moving quickly to achieve their targets. Over the 15-year period of their study, BHCs regularly used new issues of shares and share repurchase programs to actively manage their capital levels. Several reasons for differing capital ratios among BHCs are given by the authors. Banks with high earnings volatility would likely hold more capital. Banks whose customers are more sensitive to default risk via counterparty exposure may be forced to hold more capital. Firms with high charter values will want to minimize their costs of financial distress by maintaining high capital ratios. On the other hand, larger banks by asset size tend to be more diversified, enjoy scale economies in risk management, have ready access to capital markets, and are possibly viewed as "too big to fail" with attendant implicit government guarantees.

Flannery and Rangan (2008) also document a large increase in bank capital during the 1990s. The authors note the timing correlation with deregulation of the banking industry and the related increase in risk exposure. They suggest that increased diversification may have been offset by the increased risk of the newly permissible activities. As a result, it was counterparty risk that was the driving force for higher capital ratios.

13.3 Methodology

13.3.1 Instrumental Variable for Capital

We differ from previous studies that deal with the endogeneity between risk and capital using traditional methods such as a simultaneous equation approach or two- or three-stage regression analysis.[2] In this study, we follow the method and concept of Hughes et al. (2001, 2003); and others and use stochastic frontier analysis to estimate the inefficiency of our sample of bank holding companies. See Jondrow et al. (1982) for a discussion of fitting production frontier models. We then create a unique instrumental variable for bank capital to be used in regressions of capital and risk. The question we ask is: "How efficient is a bank holding company in converting the resources with which it has to work into profit?" The frontier developed is exogenous to any specific bank since it is based on the results of all banks in the sample. From this frontier, we measure the inefficiency of each bank as the distance between the frontier and that specific bank's pretax income. This measure, however, must be adjusted for those elements that are beyond the control of the bank.

Our unrestricted frontier model determines the highest possible profitability based solely on the book value of assets employed. The unrestricted model is specified as

$$\begin{aligned} &PTI(BVA, \sigma_{\text{BANK}}) = a + b_1 BVA + b_2 (BVA)^2 + e \\ &e = \xi - \varsigma \\ &\xi \sim iid\, N\left(0, \sigma_\xi^2\right), \varsigma(\geq 0) \sim iid\, N\left(0, \sigma_\varsigma^2\right) \end{aligned} \tag{13.1}$$

where *PTI* is pretax income, *BVA* is of book value of assets, ξ is statistical noise, ς is systematic shortfall (under management control), and $\varsigma \geq 0$. A quadratic specification is used to allow for a nonlinear relation between the pretax income and the book value of assets.

Our next step is to develop a second frontier based on the level of bank holding company capital as well as the amount of assets. The implication of using the unrestricted model is that we are measuring the unconditional inefficiency of the banking organization. By also conditioning the model on capital, we can develop a measure of the incremental efficiency or inefficiency of an organization due to its

[2] In Appendix 1, we explain the execution of stochastic frontier in this chapter.

capital level. It is this incremental inefficiency due to a bank's capital level that we propose to use as an instrument for capital in a regression of risk on capital. Specifically, our restricted model, again in a quadratic form, is as follows:

$$\begin{aligned} &PTI(BVA, BVC, \sigma_{\mathrm{BANK}}) = \alpha + \beta_1 BVA + \beta_2 (BVA)^2 + \beta_3 BVC + \varepsilon \\ &\varepsilon = v - u \\ &v \sim iid\, N\left(0, \sigma_v^2\right)\, u(\geq 0) \sim iid\, N\left(0, \sigma_u^2\right) \end{aligned} \tag{13.2}$$

where BVC is the book value of capital, v is statistical noise, and u denotes the inefficiency of a bank considering its use of both assets and capital.

The two assessments of inefficiency allow us to measure the difference in profitability due to the use of capital by calculating the difference in the inefficiency between the restricted and unrestricted model. Specifically,

$$\delta = u - \varsigma \tag{13.3}$$

This becomes our instrumental variable for capital. While any measure of profitability endogenously includes risk, our instrument, the difference between two measures of profitability conditioned only on capital, is related to capital but not to risk which is included in both models.

13.3.2 Generalized Method of Moments

We first apply a generalized method of moments (GMM) regression in this study.[3] There are several reasons why we need to consider the infeasibility of the OLS regression. First, the departure from normality of the variable δ due to the combined error terms should be taken into account in the analysis. There is no theory to support a Gaussian distribution of these variables. Furthermore, in practice, the ranges of the independent and dependent variables are bounded within certain intervals. Unlike other estimators, GMM is robust and does not require information on the exact distribution of the disturbances. We follow Hamilton (1994) to construct our GMM estimation. To control for the impact of size on a bank's risk-taking behavior, the book value of assets is considered in the model (Gatev et al. 2009). The relationship between the variables specifying bank behavior and the use of equity is analyzed by GMM regression. Specifically,

$$y_{k,t} = c_t + b_{k,t}\delta_{i,t} + \gamma_{k,t}\ln(BVA) + \eta_{k,t}, \tag{13.4}$$

where $y_{k,t}$ is one of the measures of risk or behavior (e.g., total equity/total asset) for bank i in year t; c is a constant; $b_{k,t}$ is the coefficient of instrumental

[3] In Appendix 2, we provide the derivation of the GMM model.

variable of capital, $\delta_{i,k}$, for k's regression in year t; $\gamma_{k,t}$ is the coefficient of natural logarithm of bank's book value; and $\eta_{k,t}$ is the error term.

13.4 Data

We obtain our data on bank holding companies from Federal Reserve reports FR Y-9C for the years 1993–2008. Data on risk-weighted assets, tier 1 capital, and tier 2 capital were not included with the FR Y-9C reports from 1993 to 1996. We were graciously provided this missing information by the authors of Berger et al. (2008).

Table 13.1 displays the descriptive statistics of the sample BHCs in our analysis. The total of 24,973 bank-year observations ranges from 2,256 in 2005 to 678 in 2008. From 2005 to 2006, there is an especially large drop in the number of BHCs included in our data. This is primarily due to a change in the reporting criteria for the FR Y-9C report. Starting in 2006, the threshold for required reporting by a BHC was increased from BHCs with $150 million in total assets to BHCs with $500 million in total assets. Note that in spite of the 57 % drop in the number of BHCs reporting in 2006 compared with 2005, the total assets represented in the sample for these 2 years decreased by only 14 %.

Our data start in 1993 because 1992 was the final year in which capital ratios were still adjusting in order to conform to the Basle I Capital Accord. As a result, 1993 represents the first year that does not include any mandated changes in the capital ratios. The entire period of 1993–2008 contains a number of significant events affecting the banking industry. For instance, the Riegle-Neal Interstate Banking and Branching Act was passed in 1994 eliminating geographic restrictions on bank expansion. In 1999 the Gramm-Leach-Bliley Financial Services Modernization Act was passed effectively repealing the Glass-Steagall Act. Together these two acts overturned 65 years of legislation and regulation intended to keep banks financially sound.

From an economic point of view, the early portion of our time period represented a time of recovery from recession. The economy then moved from recovery to growth, and the decade ended in a tech-stock boom followed by a bursting of the tech-stock price bubble and an attendant recession. The new decade brought traditional financial policies intended to stimulate the economy which, in hindsight, probably helped to lay the foundation for the housing price bubble which precipitated the 2007–2009 financial crisis. The time period from 1993 to 2008 seems to be a very appropriate period in which to analyze bank capital ratios.

Previous empirical studies have used market measures of risk and various risk measures derived from a market model based on return data. However, this approach necessarily limits the sample to publically owned banks or bank holding companies. In this study, we wish to determine the impact of capital on various measures of risk and acknowledge the importance of small banks and bank holding companies, as well as the largest bank holding companies. This concern is significant since public policy related to the banking industry must consider the broadest sample and not only the largest organizations. As a result, we utilize a large panel data set and turn to the typical accounting measures of a bank's risk.

Table 13.1 Statistical summary

Year	N	BVA (US$ million) Mean	Max	Min	SD	BE	PTI	ROE	OHE	OBS	RA	Cap	E/A	NPA	Gap
1993	1,525	2,630	216,574	22	12,260	207	43	0.135	4.10	0.120	0.438	0.161	0.085	0.069	0.160
1994	1,306	3,442	250,489	8	15,334	265	57	0.125	4.45	0.139	0.470	0.164	0.087	0.060	0.166
1995	1,355	3,526	256,853	27	16,410	283	62	0.124	4.18	0.240	0.493	0.172	0.093	0.052	0.160
1996	1,405	3,516	336,099	28	18,238	283	63	0.127	4.27	0.220	0.585	0.148	0.093	0.047	0.121
1997	1,493	3,614	365,521	30	20,027	281	65	0.127	4.04	0.222	0.607	0.146	0.094	0.043	0.086
1998	1,563	3,742	668,641	32	28,629	292	62	0.124	4.04	0.218	0.621	0.157	0.094	0.039	0.073
1999	1,658	3,845	716,937	38	29,651	292	75	0.130	4.18	0.215	0.652	0.151	0.089	0.041	0.032
2000	1,726	3,674	715,348	38	29,664	293	66	0.122	4.11	0.190	0.672	0.145	0.091	0.045	0.042
2001	1,818	3,869	693,575	38	31,651	319	56	0.117	3.77	0.224	0.682	0.145	0.091	0.055	0.026
2002	1,968	3,911	758,800	40	32,953	330	66	0.124	3.72	0.215	0.681	0.149	0.093	0.056	0.066
2003	2,129	3,960	820,103	41	34,712	334	75	0.127	3.71	0.224	0.688	0.152	0.093	0.056	0.079
2004	2,240	4,558	1,157,248	40	45,012	421	77	0.123	3.99	0.225	0.709	0.150	0.092	0.044	0.115
2005	2,252	5,407	1,494,037	40	55,416	466	91	0.130	4.15	0.242	0.725	0.147	0.090	0.041	0.115
2006	969	10,842	1,463,685	43	76,903	942	179	0.124	3.49	0.338	0.764	0.137	0.091	0.046	0.086
2007	888	10,801	1,720,688	72	88,290	1,010	152	0.110	3.82	0.322	0.778	0.133	0.092	0.085	0.076
2008	678	13,884	2,175,052	79	123,635	1,194	70	0.084	3.21	0.315	0.760	0.145	0.092	0.085	0.041

The numbers of bank holding companies (*BHCs*) and statistics of their book value of asset (*BVA*) over sample years are reported. The means of other descriptive statistics are listed: BE, book value of equity; PTI, pretax income; ROE, return of equity; OHE (*overhead efficiency*), noninterest expenses/noninterest income; OBS (*off-balance-sheet activities*), all OBS activities/total assets; RA (*risk-based asset ratio*), total risk-based assets/total assets; Cap (*total risk-based capital ratio*), (tier 1 capital + tier 2 capital)/total risk-based asset; E/A, total equity/total asset; NPA (*nonperforming assets ratio*), nonperforming assets/total equity capital; Gap, (*interest-sensitive gap*) (IS assets – IS liabilities)/total asset. The BVA, BE, and PTI are in million US dollar

13.4.1 Overall Observations of BHC Data

We see the significant events and the economic activity listed above in the statistics in Table 13.1. First, the size of BHCs measured by either their asset values or equity has increased, while the number of banks has decreased. This trend is still evident after adjusting for changes in the reporting criteria for the FR Y-9C report. The government deregulation noted above has resulted in increased concentration in the banking industry. We note also the significant cross-sectional variation in scale of BHCs that suggests the utilization and operation of their resources vary considerably.

When we look in Table 13.1 at the basic leverage ratio of equity to assets (E/A), we see a generally rising ratio. In 1993, the ratio was 8.5 %, while in 2008 it was 9.2 %. These ratios appear to be in line with mandatory capital requirements. We also see variation in this trend consistent with prevailing economic activity. For example, the decline from 9.4 % in 1998 to 8.9 % in 1999 reflects the tech-stock problems of that time period. In Table 13.1, we also see a rising trend in RA, the ratio of risk-based assets to total assets. Here, however, the trend is far more pronounced, rising from 43.80 % in 1993 to 76.00 % in 2008. Confirmation of these two trends comes from the trend in CAP, the ratio of tier 1 plus tier 2 capital to risk-based assets. This ratio declines from 16.10 % in 1993 to 14.50 % in 2008. While these ratios are substantially above the Basle Capital Accord standards, the trend is clearly down.

Another dramatic trend over this time period is the increase in off-balance-sheet activity. In Table 13.1, the off-balance-sheet activities-to-total assets ratio (OBS) has increased from 12.00 % in 1993 to 31.50 % in 2008. While this trend is not a surprise, we need to ask if there is capital to support this expansion and consider the makeup of the components of off-balance-sheet activities. It is unclear whether BHCs use off-balance-sheet activities to decrease or increase risk.

The time-varying overall performance measures of our sample of BHCs such as pretax income (PTI), return on equity (ROE), nonperforming assets ratio (NPA), and the interest-sensitive gap (Gap) are shaped by major economic occurrences and policies. Return on equity has varied in a relatively narrow band over this time period. With the exception of 2007 and 2008, the return on equity ranged from 12.20 % to 13.50 %. In line with the financial crisis that started in 2007, ROE declined to 11.00 % in 2007 and to 8.40 % in 2008. It is also noteworthy that the highest return on equity was in the first year of our sample period, 1993. Nonperforming assets appear to move in concert with economic activity. The recovery and expansion period of 1993–1998 is marked by a steady decrease in the ratio of nonperforming assets to equity. This is followed by an increase in this ratio during the tech-stock bubble and recession after which we see another decline until the crisis of 2007 and 2008.

Since the industrial structure of financial services changes intertemporally, we analyze the risk and use of capital by BHCs year by year. The analysis suggests banks progressively depend more on aggressive funding sources and new product lines over our sample period. Given that financial leverage (e.g., equity/asset ratio) must remain approximately stable due to regulatory requirements, bankers may try to improve their ROE by (1) enhancing overhead efficiency (OHE), (2) engaging in more off-balance-sheet activities (OBS), and (3) using interest-sensitive gap management in an attempt to decrease their total risk-based capital ratio (Cap) while maintaining an attractive ROE. The above developments in the banking industry generate potential improvement in performance but also intensify uncertainties and complexities of bank management. Therefore, a study to investigate the impact of the use of capital on the riskiness of banks is an indispensible element in bank management.

13.4.2 Instrumental Variable

The statistical summary of our instrumental variable, δ, for each year is shown in Table 13.2. Consistent with the findings documented by Hughes et al. (2001), John et al. (2000), Keeley (1990), and Kim and Santomero (1988), the use of equity capital by banks, on average, triggers a loss in efficiency. The dispersion of δ is substantial both cross-sectionally and intertemporally. For our sample, the distribution of δ in the same year tends to be skewed to the left-hand side and leptokurtic (i.e., has positive excess kurtosis). Therefore, we look at nonparametric statistics and use a normality-free regression model in our analysis to avoid the possible errors in estimation.

Table 13.2 Distribution of instrumental variables

	Mean	SD	Skewness	Kurtosis	Max	Min
1993	−0.038	0.106	−0.167	1.936	0.470	−0.494
1994	−0.040	0.112	−0.386	3.394	0.511	−0.805
1995	−0.036	0.114	−0.196	3.791	0.551	−0.753
1996	−0.041	0.111	−0.037	1.852	0.506	−0.557
1997	−0.223	0.162	−2.848	16.357	0.302	−1.253
1998	−0.032	0.112	−0.018	3.332	0.544	−0.744
1999	−0.053	0.144	−0.266	1.274	0.502	−0.714
2000	−0.054	0.144	−0.101	1.358	0.541	−0.761
2001	−0.049	0.144	−0.146	1.286	0.558	−0.624
2002	−0.042	0.153	−0.160	1.707	0.726	−0.739
2003	−0.021	0.151	0.078	0.573	0.540	−0.559
2004	−0.021	0.121	0.070	1.425	0.531	−0.634
2005	−0.024	0.122	0.044	2.229	0.555	−0.841
2006	−0.016	0.137	−0.506	4.024	0.519	−1.006
2007	−0.015	0.190	−0.438	1.536	0.542	−0.616
2008	−0.012	0.049	−2.758	27.015	0.129	−0.581

Descriptive statistics of the instrumental variable for capital over sample years are presented. The instrumental variable $\delta = u - \varsigma$ is a measure of incremental bank inefficiency due to capital level, where the stochastic frontiers are $PTI = a + b_1BVA + b_2(BVA)^2 + e$, $e = \xi - \varsigma$, and $PTI = \alpha + \beta_1BVA + \beta_2(BVA)^2 + \beta_3BVC + \varepsilon$, $\varepsilon = v - u$

13.4.3 Measures of Bank Risk

We investigate the risks faced by banks from various aspects. Table 13.3 displays the measures of risk used in this study: credit risk, liquidity risk, interest rate risk, off-balance-sheet (OBS) risk, market risk, and finally leverage risk. Credit risk is concerned with the quality of a bank's assets. Historically this has focused on a bank's loan portfolio, but recent events have shown the importance of looking at all bank assets in light of potential default risk. Liquidity risk measures the ability of a bank to meet all cash needs at a reasonable cost whenever they arise. Interest rate risk is the extent to which banks have protected themselves from market-driven changes in the level of interest rates. Banks have the opportunity to use asset/liability management tools to mitigate the impact of changes in interest rates on both bank earnings and bank equity. We also collect data on off-balance-sheet activities and investigate their relationship with bank capital. Market risk is the risk of changes in asset prices that are beyond the control of bank management. Finally, leverage risk is the risk arising from the capital structure decisions of the BHC. The first five measures of risk relate to the various elements of business risk confronting bank management. Leverage risk, on the other hand, relates directly to the financial decisions taken in terms of the amount of capital employed. From another perspective, it can be said that minimum capital requirements (i.e., maximum leverage standards) are mandated by regulators to mitigate the various elements of business risk that the BHC accepts.

Table 13.4 displays the Spearman correlation coefficients between bank size and our instrumental variable over the sample period. We look at this nonparametric test due to the non-normal distribution of the instrument and variables. We believe that the generally insignificant correlation between our instrument and the book value of assets in combination with the generally significant correlation of our instrument and the book value of equity justifies the use of delta as an instrument for capital. In addition, Table 13.4 shows that, measured by book value of equity and pretax income, large BHCs tend to suffer a greater loss in efficiency than their smaller counterparts at a statistically significant level. On the other hand, the value of assets does not necessarily demonstrate a negative relation with bank efficiency. These findings suggest that the inefficiency of BHCs comes from the use of equity capital but is not directly led by the expansion of business scale and/or scope. Therefore, a careful investigation of the impact of capital on banking risks is appropriate.

13.5 Empirical Results

We look at seven different measures of risk: credit risk, liquidity risk, interest rate risk, off-balance-sheet risk, market risk, composite risk, and leverage risk. While many of the results support the proposition that increased capital requirements reduce risk in BHCs, there are some very significant results that suggest the opposite – that BHCs increase risk as their capital ratios increase.

Table 13.3 Variables

Symbol	Definition
Overall risk	
Eq/A	Total equity/total asset
RA/A	Total risk-based assets/total assets
RC/RA	Capital requirement ratio (total risk-based capital/total risk-based assets)
Tier 1/RA	Tier 1 capital/total risk-based assets
Tier 2/tier 1	Tier 2 capital/tier 1 capital
Credit risk	
NPL/LL	Nonperforming assets/total loans and leases
NPL/E	Nonperforming assets/total equity capital
Charge-offs/L	Net loan charge-offs/total loans and leases
Provision/L	Annual provision for loan losses/total loans and leases
Provision/E	Annual provision for loan losses/total equity capital
Allowance/L	Allowance for loan losses/total loans and leases
Allowance/E	Allowance for loan losses/total equity capital
Liquidity risk	
STPF/A	Short-term purchased funds (Eurodollars, federal funds, security RPs, large CDs, and commercial paper)/total assets
Cash/A	Cash and due from other banks/total assets
HLA/A	Cash assets and government securities/total assets
FFS/A	(Federal funds sold + reverse RPs – sum of federal funds purchased – RPs)/total assets
FFP/A	(Federal funds purchased + RPs)/total assets
Cash/STPF	Cash and due from other banks/short-term purchased funds (Eurodollars, federal funds, security RPs, large CDs, and commercial papers)
Interest rate risk	
Gap	Interest-sensitive gap (IS assets – IS liabilities)/total assets
Off-balance-sheet risk	
OBS/A	Off-balance-sheet assets/total assets
Der/A	Credit equivalent amount of off-balance-sheet derivative contracts/total assets
Der/RA	Credit equivalent amount of off-balance-sheet derivative contracts/total risk-based assets
IR Der	Notional amount of interest rate derivatives held for trading/notional amount of interest rate derivatives held for other purposes
FX Der	Notional amount of foreign exchange derivatives held for trading/notional amount of foreign exchange derivatives held for other purposes
Eq Der	Notional amount of equity derivatives held for trading/notional amount of equity derivatives held for other purposes
Cmd Der	Notional amount of commodity derivatives held for trading/notional amount of commodity derivatives held for other purposes
OBS/E	Total OBS LC, commitments, credit card lines of credit, and loan/total equity
Der/A	Total derivatives/total assets

(*continued*)

Table 13.3 (continued)

Symbol	Definition
Market risk	
Trading assets	Trading account assets/total assets
Trading A/L	Trading account assets/trading account liabilities
Investment M/B	Market value of investment portfolio/book value of investment portfolio
Performance	
PTI/A	Pretax income/asset
ROE	Return on equity
ROA	Return on asset
ATR	Average tax rate (taxes/pretax income)
Spread	Earning spread (interest income/(loan + investment)−(interest expenses/deposits))
OHE	Overhead efficiency (noninterest expenses/noninterest income)

Our results are displayed in Tables 13.5, 13.6, 13.7, 13.8, 13.9, and 13.10 and provide a number of interesting insights. To enhance robustness, we present both Spearman's rank correlation coefficient between the tested variable and the instrumental variable for capital, δ, and the coefficient of δ in GMM regressions. To control for the size of the bank holding companies, each coefficient of δ is generated by GMM regression with a constant and the natural logarithm of the book value of assets. While the coefficient of the control variable and the constant term are omitted from the tables, they are available upon request. In Table 13.5, we find a positive relationship between ratio of total equity to total assets and our instrument for capital (see Eq/A). The coefficient on our instrument is strictly positive and statistically significant. This is clearly what we would expect. As leverage decreases, so does risk; therefore, higher capital should be associated with higher levels of this risk measure. In other words, it should be a positive relationship, and it is. However, for the ratio of risky assets to total assets, risk increases as the ratio increases. Therefore, higher capital should be associated with lower levels of this risk measure (a negative relationship), and again that is what we find. When we look at just tier 1 capital to total risky assets, we find the expected positive relationship, and when we look at the ratio of tier 2 to tier 1 capital, we find the expected negative relationship.

In Table 13.6, we look at some traditional measures of credit risk. As the ratio of nonperforming loans to total loans increases, so does risk. Therefore, the coefficient on capital should be negative, and they are with several exceptions over the years. Our second measure of credit risk is the ratio of nonperforming loans to total equity. Here again higher levels of the ratio imply higher risk, so we expect to find a negative relationship and we do, and this time without exception and at high levels of significance. When we look at the ratio of loan charge-offs to loans outstanding, we have more exceptions, but in general we find an expected negative relationship.

Table 13.4 Explanatory variables and instrument for capital of BHCs: Spearman's rank correlation coefficient

	2008	2007	2006	2005	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1994	1993
ρ(BVA, δ)	0.27	0.02	−0.01	−0.01	0.01	−0.07	−0.01	−0.03	−0.03	0.01	0.03	0.25	0.07	0.07	0.00	−0.01
p-value	0.001	0.409	0.422	0.429	0.412	0.101	0.413	0.323	0.327	0.450	0.337	0.000	0.145	0.136	0.495	0.467
ρ(BE, δ)	0.51	0.27	0.23	0.32	0.34	0.26	0.32	0.31	0.30	0.33	0.34	0.51	0.34	0.33	0.26	0.20
p-value	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ρ(PTI, δ)	0.38	0.09	0.04	0.13	0.13	0.04	0.13	0.16	0.20	0.20	0.17	0.35	0.23	0.23	0.16	0.17
p-value	0.000	0.123	0.304	0.005	0.006	0.247	0.007	0.002	0.000	0.000	0.003	0.000	0.000	0.000	0.008	0.002

Spearman's rank correlation coefficients between book value of assets (*BVA*), book value of equity (*BE*), and pretax income (*PTI*) with the instrumental variable for capital (δ) in each year are reported

Table 13.5 Leverage risk

Panel A: Spearman's rank correlation coefficient between variable and δ																
	2008	2007	2006	2005	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1994	1993
Eq/A	0.71	0.82	0.83	0.90	0.89	0.83	0.89	0.89	0.93	0.91	0.92	0.72	0.90	0.90	0.91	0.92
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
RA/A	−0.10	−0.18	−0.15	−0.21	−0.22	−0.15	−0.20	−0.20	−0.17	−0.12	−0.16	−0.10	−0.11	−0.16	−0.15	−0.20
p-value	0.144	0.012	0.022	0.000	0.000	0.001	0.000	0.000	0.002	0.020	0.004	0.047	0.047	0.008	0.012	0.000
RC/RA	0.41	0.42	0.46	0.43	0.44	0.44	0.44	0.44	0.39	0.42	0.42	0.41	0.42	0.42	0.36	0.40
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Tier 1/RA	0.38	0.45	0.50	0.44	0.42	0.46	0.44	0.44	0.45	0.44	0.44	0.44	0.44	0.44	0.44	0.44
p-value	0.000	0.000	0.000	0.000	0.009	0.003	0.004	0.005	0.004	0.004	0.005	0.004	0.004	0.004	0.004	0.004
Tier 2/tier 1	−0.23	−0.37	−0.38	−0.33	−0.36	−0.35	−0.35	−0.35	−0.35	−0.35	−0.35	−0.35	−0.35	−0.35	−0.35	−0.35
p-value	0.006	0.000	0.000	0.002	0.001	0.001	0.006	0.002	0.003	0.004	0.003	0.003	0.003	0.003	0.003	0.003
Panel B: GMM regressions																
	2008	2007	2006	2005	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1994	1993
Eq/A	0.312	0.102	0.144	0.206	0.202	0.161	0.161	0.165	0.174	0.165	0.222	0.098	0.221	0.215	0.202	0.204
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
RA/A	−0.035	−0.105	−0.045	−0.199	−0.242	−0.152	−0.155	−0.152	−0.154	−0.114	−0.165	−0.049	−0.066	0.021	0.070	−0.118
p-value	0.375	0.000	0.150	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.106	0.116	0.372	0.172	0.225

(*continued*)

Table 13.5 (continued)

Panel B: GMM regressions	2008	2007	2006	2005	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1994	1993
RC/RA	0.192	0.068	0.102	0.249	0.246	0.200	0.216	0.240	0.256	0.263	0.381	0.165	0.399	0.259	0.253	0.284
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Tier 1/RA	0.212	0.080	0.114	0.267	0.266	0.212	0.229	0.250	0.261	0.266	0.380	0.167	0.406	0.421	0.411	0.427
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Tier 2/tier 1	−0.441	−0.216	−0.229	−0.379	−0.393	−0.276	−0.292	−0.280	−0.222	−0.196	−0.231	−0.055	−0.306	−0.196	−0.255	−0.338
$p-$value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.081	0.000	0.000	0.000	0.000

Spearman's rank correlation coefficients between each risk measure variable with δ, the instrument for capital, are reported for each year in Panel A. In Panel B, the GMM regression coefficients of $b_{k,t}$ for each dependent variable k in each year t are reported. Specifically, $y_{k,t} = c_t + b_{k,t}\delta_{i,t} + \gamma_{k,t}\ln(BVA) + \eta_{k,t}$, where $y_{k,t}$ is one of the measures of risk or performance (e.g., total equity/total asset) for bank i in year t; c is a constant; $b_{k,t}$ is the coefficient of instrumental variable of capital, $\delta_{i,k}$, for k's regression in year t; $\gamma_{k,t}$ is the coefficient of natural logarithm of bank's book value; an $\eta_{k,t}$ is the error term

Table 13.6 Credit risk

Panel A: Spearman's rank correlation coefficient between variable and δ

	2008	2007	2006	2005	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1994	1993
NPL/LL	−0.02	−0.03	0.02	−0.02	−0.03	−0.02	0.01	−0.03	−0.05	−0.05	−0.02	0.01	−0.04	−0.08	−0.05	−0.07
p-value	0.412	0.340	0.418	0.346	0.285	0.290	0.407	0.291	0.209	0.189	0.381	0.453	0.246	0.118	0.215	0.123
NPL/E	−0.23	−0.22	−0.17	−0.18	−0.14	−0.16	−0.21	−0.25	−0.26	−0.26	−0.22	−0.16	−0.23	−0.25	−0.22	−0.27
p-value	0.005	0.003	0.015	0.000	0.002	0.001	0.000	0.000	0.000	0.000	0.000	0.005	0.000	0.000	0.000	0.000
Charge-offs/L	0.01	0.02	0.05	0.03	0.02	0.02	0.06	0.03	0.02	0.00	0.00	0.10	−0.01	−0.04	−0.07	−0.08
p-value	0.452	0.415	0.270	0.328	0.341	0.360	0.142	0.290	0.396	0.489	0.498	0.048	0.410	0.282	0.148	0.106
Provision/L	0.01	−0.13	−0.18	−0.11	−0.14	−0.14	−0.08	−0.14	−0.13	−0.15	−0.15	−0.06	−0.15	−0.17	−0.13	−0.07
p-value	0.442	0.052	0.010	0.083	0.052	0.052	0.065	0.005	0.010	0.005	0.006	0.160	0.008	0.004	0.022	0.142
Provision/E	−0.24	−0.33	−0.39	−0.41	−0.38	−0.35	−0.37	−0.41	−0.40	−0.41	−0.38	−0.24	−0.34	−0.33	−0.27	−0.26
p-value	0.004	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Allowance/L	0.18	0.08	0.08	0.13	0.12	0.13	0.17	0.13	0.15	0.14	0.14	0.17	0.11	0.06	0.05	0.02
p-value	0.021	0.144	0.135	0.045	0.069	0.068	0.001	0.012	0.005	0.009	0.011	0.003	0.038	0.168	0.204	0.350
Allowance/E	−0.40	−0.54	−0.57	−0.60	−0.57	−0.50	−0.56	−0.57	−0.55	−0.52	−0.50	−0.33	−0.43	−0.44	−0.43	−0.49
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Panel B: GMM regressions

	2008	2007	2006	2005	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1994	1993
NPL/LL	−0.074	−0.006	−0.002	0.002	0.004	0.007	−0.001	−0.001	−0.001	−0.001	0.000	−0.004	−0.004	−0.009	−0.001	−0.003
p-value	0.022	0.017	0.219	0.134	0.007	0.000	0.258	0.402	0.373	0.246	0.415	0.012	0.028	0.003	0.437	0.204
NPL/E	−9.510	−0.159	−0.103	−0.088	−0.087	−0.004	−0.122	−0.118	−0.087	−0.088	−0.090	−0.064	−0.116	−0.173	−0.191	−0.195
p-value	0.096	0.000	0.000	0.000	0.000	0.463	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Charge-offs/L	−0.007	−0.001	0.000	0.001	0.002	0.005	0.000	0.003	0.000	0.000	−0.001	−0.003	−0.002	−0.003	−0.002	−0.004
p-value	0.074	0.138	0.431	0.134	0.037	0.000	0.358	0.142	0.435	0.313	0.199	0.010	0.045	0.044	0.113	0.021

(*continued*)

Table 13.6 (continued)

Panel B: GMM regressions																
	2008	2007	2006	2005	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1994	1993
Provision/L	−0.009	−0.003	−0.003	−0.002	−0.001	0.001	−0.002	0.001	−0.001	−0.001	−0.003	−0.004	−0.006	−0.016	−0.013	−0.006
p-value	0.074	0.000	0.000	0.008	0.047	0.044	0.024	0.419	0.182	0.089	0.000	0.003	0.008	0.105	0.090	0.029
Provision/E	−15.83	−0.064	−0.055	−0.074	−0.070	−0.037	−0.089	−0.065	−0.061	−0.063	−0.087	−0.051	−0.076	−0.064	−0.062	−0.075
p-value	0.096	0.000	0.000	0.000	0.000	0.000	0.000	0.005	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Allowance/L	0.009	0.002	0.002	0.006	0.006	0.005	0.004	0.006	0.004	0.005	0.006	0.001	0.004	0.001	0.008	0.026
p-value	0.013	0.049	0.015	0.000	0.000	0.000	0.000	0.005	0.003	0.000	0.000	0.332	0.043	0.306	0.086	0.145
Allowance/E	−6.659	−0.118	−0.153	−0.201	−0.201	−0.133	−0.183	−0.163	−0.163	−0.165	−0.202	−0.095	−0.199	−0.237	−0.265	−0.300
p-value	0.096	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Please see the description of Table 13.5

Table 13.7 Liquidity and interest rate risk

Panel A: Spearman's rank correlation coefficient between variable and δ																
	2008	2007	2006	2005	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1994	1993
STPF/A	−0.05	−0.10	−0.11	−0.30	−0.34	−0.38	−0.38	−0.38	−0.36	−0.34	−0.41	−0.16	−0.31	−0.40	−0.39	−0.45
p-value	0.287	0.104	0.080	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.005	0.000	0.000	0.000	0.000
Cash/A	−0.01	0.01	0.04	0.03	0.04	0.06	0.00	0.04	0.00	−0.01	−0.02	−0.06	−0.09	−0.13	−0.01	−0.07
p-value	0.440	0.474	0.290	0.252	0.209	0.133	0.487	0.260	0.490	0.446	0.379	0.176	0.073	0.025	0.430	0.142
HLA/A	−0.01	0.01	0.02	0.05	0.07	0.09	0.03	0.07	0.08	0.10	0.09	0.04	0.07	0.04	0.12	−0.07
p-value	0.476	0.443	0.382	0.168	0.089	0.041	0.274	0.096	0.088	0.042	0.059	0.250	0.130	0.291	0.032	0.142
FFS/A	−0.16	−0.04	0.01	0.05	0.05	0.17	0.06	0.08	0.08	0.08	0.03	−0.10	−0.01	−0.04	0.09	0.00
p-value	0.037	0.294	0.471	0.177	0.148	0.000	0.127	0.081	0.070	0.077	0.287	0.048	0.461	0.279	0.091	0.485
FFP/A	0.04	−0.04	−0.15	−0.32	−0.33	−0.43	−0.39	−0.35	−0.32	−0.29	−0.34	−0.12	−0.65	−1.00	−0.99	−1.00
p-value	0.346	0.289	0.024	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.023	0.000	0.000	0.000	0.000
Cash/STPF	−0.04	−0.03	−0.04	−0.10	−0.10	−0.10	−0.13	−0.11	−0.10	−0.09	−0.13	−0.08	−0.20	−0.31	−0.24	−0.31
p-value	0.431	0.434	0.374	0.266	0.241	0.196	0.315	0.240	0.275	0.261	0.288	0.250	0.277	0.266	0.259	0.295
Gap	0.07	−0.07	−0.08	−0.05	−0.08	−0.03	−0.11	−0.05	−0.06	−0.02	−0.07	−0.03	−0.05	−0.06	−0.09	−0.10
p-value	0.207	0.019	0.015	0.014	0.016	0.255	0.021	0.163	0.141	0.352	0.117	0.326	0.203	0.185	0.092	0.054
Panel B: GMM regressions																
	2008	2007	2006	2005	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1994	1993
STPF/A	−0.091	−0.015	−0.038	−0.045	−0.044	−0.024	−0.028	−0.031	−0.040	−0.046	−0.029	−0.019	−0.026	−0.011	−0.021	−0.011
p-value	0.004	0.016	0.018	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.002	0.003	0.060	0.016	0.073
Cash/A	0.001	−0.001	−0.012	−0.004	0.005	0.007	0.005	0.006	0.002	−0.004	0.005	−0.018	−0.023	−0.031	−0.004	0.002
p-value	0.487	0.397	0.046	0.409	0.224	0.131	0.272	0.121	0.405	0.218	0.252	0.015	0.001	0.000	0.300	0.441

(*continued*)

Table 13.7 (continued)

Panel B: GMM regressions	2008	2007	2006	2005	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1994	1993
HLA/A	0.011	0.000	−0.020	0.014	0.029	0.024	0.020	0.026	0.035	0.039	0.060	0.027	0.069	0.043	0.127	0.002
p-value	0.346	0.484	0.013	0.234	0.002	0.003	0.021	0.001	0.006	0.000	0.000	0.024	0.002	0.031	0.000	0.441
FFS/A	0.029	0.001	0.011	0.018	0.028	0.054	0.020	0.026	0.036	0.045	0.040	0.004	0.021	0.001	0.053	1.324
p-value	0.323	0.461	0.204	0.027	0.002	0.000	0.020	0.009	0.001	0.000	0.006	0.333	0.047	0.464	0.000	0.154
FFP/A	−0.105	−0.005	−0.055	−0.016	−0.021	−0.020	−0.017	−0.027	−0.029	−0.024	−0.034	−0.013	0.000	0.000	0.000	0.000
p-value	0.060	0.353	0.019	0.039	0.011	0.004	0.004	0.007	0.001	0.002	0.003	0.014	0.241	0.430	0.092	0.261
Cash/STPF	172.87	−21.99	−904.47	276.21	193.34	72.96	58.75	88.84	−4.764	39.15	195.89	−0.744	−5434.9	356.92	2.226	−1990.9
p-value	0.112	0.347	0.180	0.007	0.137	0.143	0.035	0.248	0.470	0.197	0.235	0.489	0.157	0.063	0.493	0.088
Gap	−0.082	−0.072	−0.123	−0.050	−0.110	−0.012	−0.133	−0.071	−0.077	−0.020	−0.080	−0.068	−0.112	−0.116	−0.103	−0.126
p-value	0.258	0.026	0.004	0.084	0.001	0.340	0.000	0.015	0.037	0.261	0.026	0.012	0.007	0.002	0.009	0.001

Please see the description of Table 13.5

Table 13.8 Off-balance-sheet risk

Panel A: Spearman's rank correlation coefficient between variable and δ

	2008	2007	2006	2005	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1994	1993
OBS/A	−0.10	−0.16	−0.19	−0.29	−0.28	−0.26	−0.29	−0.26	−0.28	−0.29	−0.28	−0.18	−0.29	−0.28	−0.30	−0.30
p-value	0.328	0.247	0.128	0.070	0.064	0.113	0.084	0.101	0.104	0.153	0.101	0.150	0.120	0.113	0.117	0.116
Der/A	−0.93	−0.99	−0.96	−0.96	−0.94	−0.92	−1.00	−0.92	−0.92	−0.90	−0.89	−0.96	−1.02	−1.00	−1.00	−1.00
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Der/RA	−0.21	−0.20	−0.29	−0.28	−0.26	−0.29	−0.26	−0.28	−0.22	−0.27	−0.27	−0.27	−0.26	−0.26	−0.26	−0.24
p-value	0.319	0.227	0.195	0.140	0.129	0.225	0.168	0.201	0.208	0.306	0.202	0.288	0.240	0.225	0.176	0.175
IR Der	−0.14	−0.16	−0.20	−0.25	−0.27	−0.27	−0.28	−0.27	−0.26	−0.26	−0.27	−0.25	−0.25	−0.27	na	na
p-value	0.256	0.276	0.179	0.119	0.087	0.103	0.115	0.111	0.121	0.149	0.152	0.149	0.158	0.143	na	na
FX Der	−0.16	−0.48	−0.56	−0.59	−0.60	−0.60	−0.61	−0.62	−0.60	−0.59	−0.58	−0.59	−0.62	−0.63	na	na
p-value	0.085	0.103	0.093	0.065	0.045	0.039	0.043	0.045	0.046	0.053	0.059	0.060	0.061	0.060	na	na
Eq Der	−0.71	−0.35	−0.38	−0.42	−0.43	−0.44	−0.44	−0.44	−0.43	−0.42	−0.43	−0.43	na	na	na	na
p-value	0.000	0.122	0.123	0.103	0.080	0.086	0.094	0.092	0.099	0.119	0.123	0.122	na	na	na	na
Cmd Der	−0.89	−0.45	−0.31	−0.31	−0.34	−0.35	−0.36	−0.36	−0.35	−0.35	−0.35	−0.35	na	na	na	na
p-value	0.000	0.109	0.135	0.110	0.083	0.073	0.079	0.082	0.084	0.095	0.106	0.109	na	na	na	na
OBS/E	−0.26	−0.43	−0.49	−0.49	−0.48	−0.44	−0.47	−0.48	−0.47	−0.40	−0.42	−0.21	−0.33	−0.30	−0.31	−0.33
p-value	0.019	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Der/A	−0.39	−0.54	−0.44	−0.44	−0.45	−0.45	−0.45	−0.46	−0.46	−0.44	−0.42	−0.38	−0.26	−0.12	−0.07	−0.06
p-value	0.132	0.167	0.145	0.107	0.085	0.105	0.100	0.106	0.111	0.144	0.129	0.146	0.123	0.126	0.130	0.257

(*continued*)

Table 13.8 (continued)

Panel B: GMM regressions

	2008	2007	2006	2005	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1994	1993
OBS/A	−0.002	−0.001	0.000	−0.001	−0.002	−0.001	−0.001	−0.001	−0.003	−0.003	−0.005	−0.002	−0.005	−0.004	0.000	0.000
p-value	0.377	0.092	0.124	0.005	0.014	0.065	0.081	0.144	0.030	0.074	0.088	0.062	0.052	0.065	0.500	0.500
Der/A	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
p-value	0.410	0.080	0.277	0.038	0.050	0.093	0.394	0.427	0.058	0.083	0.069	0.118	0.064	0.500	0.500	0.500
Der/RA	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	−0.001	0.000	0.000	0.000
p-value	0.474	0.083	0.285	0.054	0.063	0.223	0.338	0.462	0.061	0.097	0.071	0.118	0.075	0.500	0.500	0.500
IR Der	−0.002	−0.304	−0.173	−0.104	−0.291	−0.157	−0.121	−0.019	−4.707	−1.127	−3.398	−1.459	1.383	−2.868	na	na
p-value	0.397	0.146	0.070	0.006	0.071	0.065	0.084	0.440	0.097	0.228	0.080	0.210	0.334	0.010	na	na
FX Der	−0.106	−5.001	−8.842	−102.09	−85.289	−1.508	−0.716	−0.478	−0.964	0.078	−1.691	1.829	−4.028	−9.633	na	na
p-value	0.236	0.184	0.149	0.147	0.144	0.006	0.190	0.280	0.292	0.483	0.210	0.321	0.024	0.036	na	na
Eq Der	−0.003	−0.304	−0.415	1.004	−0.876	−0.219	0.804	0.692	2.029	−0.903	−3.828	−2.313	na	na	na	na
p-value	0.412	0.200	0.178	0.271	0.201	0.391	0.215	0.367	0.468	0.259	0.112	0.043	na	na	na	na
Cmd Der	−0.005	−0.020	−0.035	−0.265	−0.228	−0.117	−0.074	−0.094	−0.171	−0.122	−0.331	0.004	0.895	0.367	na	na
p-value	0.390	0.298	0.403	0.442	0.171	0.081	0.483	0.452	0.408	0.381	0.156	0.476	0.220	0.322	na	na
OBS/E	−1.910	−2.453	−4.405	−4.672	−4.639	−2.778	−2.912	−3.953	−3.203	−2.939	−4.320	−1.913	−4.142	−3.237	−3.536	−3.104
p-value	0.096	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
Der/A	−0.406	−0.396	−0.128	−0.382	−0.560	−0.413	−0.378	−0.242	−0.478	−0.493	−0.693	−0.165	−0.724	−0.450	0.000	0.000
p-value	0.312	0.095	0.189	0.004	0.060	0.071	0.103	0.163	0.097	0.102	0.089	0.050	0.048	0.057	0.500	0.500

Please see the description of Table 13.5

Table 13.9 Market risk

Panel A: Spearman's rank correlation coefficient between variable and δ																
	2008	2007	2006	2005	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1994	1993
Trading assets	−0.75	−0.92	−0.96	−1.00	−0.99	−0.97	−0.98	−0.98	−0.98	−0.98	−0.98	−0.98	−0.98	−0.98	−0.98	−0.98
p-value	0.000	0.000	0.000	0.000	0.000	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Trading A/L	−0.57	−0.13	−0.12	−0.12	−0.12	−0.12	−0.12	−0.12	−0.12	−0.12	−0.12	−0.12	−0.11	−0.10	−0.09	n.a
p-value	0.044	0.153	0.207	0.204	0.176	0.163	0.164	0.165	0.170	0.190	0.206	0.214	0.214	0.206	0.206	n.a
Investment M/B	0.20	0.21	0.16	0.15	0.16	0.14	0.14	0.17	0.13	0.16	0.14	0.13	0.16	0.13	0.13	n.a
p-value	0.059	0.070	0.118	0.070	0.059	0.052	0.050	0.050	0.051	0.041	0.057	0.063	0.041	0.038	0.059	n.a
Panel B: GMM regressions																
	2008	2007	2006	2005	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1994	1993
Trading assets	−0.028	0.001	−0.023	−0.007	−0.007	−0.002	−0.005	−0.003	−0.009	−0.006	−0.009	−0.004	−0.014	−0.010	−0.009	−0.003
p-value	0.062	0.399	0.049	0.003	0.020	0.173	0.056	0.104	0.040	0.075	0.067	0.047	0.077	0.077	0.074	0.252
Trading A/L	−16.484	0.227	0.970	−1.365	−1.467	3.105	−0.992	0.925	0.752	−1.869	−0.169	−2.241	−1.526	−1.025	−0.366	na
p-value	0.000	0.497	0.388	0.183	0.188	0.148	0.230	0.146	0.285	0.167	0.324	0.147	0.128	0.057	0.391	na
Investment M/B	0.002	0.001	−0.006	0.002	0.009	0.007	0.010	0.012	0.026	0.019	0.009	0.001	0.001	0.006	0.010	na
p-value	0.066	0.030	0.106	0.046	0.014	0.000	0.001	0.000	0.035	0.001	0.083	0.042	0.098	0.055	0.009	na

Please see the description of Table 13.5

Table 13.10 Performance

Panel A: Spearman's rank correlation coefficient between variable and δ																
	2008	2007	2006	2005	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1994	1993
PTI/A	0.12	0.06	0.06	0.22	0.16	0.32	0.26	0.34	0.42	0.39	0.32	0.19	0.39	0.35	0.36	0.44
p-value	0.100	0.215	0.235	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000
ROE	0.20	−0.37	−0.43	−0.40	−0.43	−0.28	−0.32	−0.24	−0.18	−0.28	−0.31	−0.35	−0.26	−0.26	−0.26	−0.19
p-value	0.138	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.001	0.000	0.000	0.000	0.000	0.000	0.000	0.001
ROA	0.10	0.07	0.06	0.20	0.17	0.31	0.29	0.36	0.43	0.40	0.36	0.24	0.45	0.42	0.41	0.48
p-value	0.136	0.175	0.214	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ATR	0.19	0.03	0.00	0.02	−0.03	−0.06	−0.08	−0.06	−0.06	−0.04	−0.10	−0.11	−0.14	−0.13	−0.06	−0.09
p-value	0.018	0.353	0.488	0.312	0.245	0.120	0.079	0.148	0.163	0.236	0.056	0.043	0.012	0.024	0.164	0.064
Spread	0.05	0.19	0.20	0.16	0.17	0.14	0.23	0.21	0.22	0.10	0.17	0.05	0.10	0.10	0.08	0.17
p-value	0.287	0.009	0.005	0.057	0.019	0.058	0.000	0.000	0.000	0.041	0.002	0.226	0.050	0.060	0.124	0.003
OHE	−0.15	−0.04	−0.05	−0.08	−0.05	−0.05	0.00	−0.01	−0.02	0.23	0.04	−0.10	−0.01	0.00	−0.01	0.00
p-value	0.049	0.295	0.244	0.053	0.167	0.154	0.490	0.439	0.366	0.113	0.264	0.054	0.454	0.477	0.453	0.489
Panel B: GMM regressions																
	2008	2007	2006	2005	2004	2003	2002	2001	2000	1999	1998	1997	1996	1995	1994	1993
PTI/A	0.027	0.006	0.009	0.013	0.011	0.010	0.015	0.017	0.017	0.019	0.017	0.015	0.023	0.020	0.020	0.024
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ROE	0.024	−0.061	−0.124	−0.146	−0.136	−0.108	−0.058	−0.037	−0.055	−0.065	−0.098	−0.010	−0.058	−0.063	−0.015	−0.083
p-value	0.317	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.003	0.000	0.000	0.001	0.000
ROA	0.022	0.004	0.005	0.009	0.008	0.008	0.011	0.013	0.012	0.013	0.013	0.011	0.017	0.016	0.016	0.018
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
ATR	3.887	0.223	0.355	0.194	0.180	0.078	0.442	0.154	0.040	0.001	0.061	0.480	0.049	−0.021	0.163	0.103
p-value	0.145	0.007	0.005	0.000	0.001	0.005	0.023	0.117	0.367	0.489	0.197	0.006	0.162	0.382	0.020	0.080
Spread	0.042	0.009	0.030	0.019	0.017	0.021	0.019	0.021	0.019	0.011	0.020	0.006	0.032	0.016	0.117	0.042
p-value	0.059	0.000	0.010	0.000	0.000	0.049	0.000	0.000	0.000	0.000	0.000	0.033	0.019	0.000	0.179	0.001
OHE	2.238	−0.628	−2.920	−1.924	0.209	−0.627	0.461	0.490	0.809	1.220	1.617	0.275	10.730	3.912	1.622	0.478
p-value	0.321	0.281	0.157	0.009	0.341	0.068	0.120	0.238	0.188	0.246	0.036	0.321	0.139	0.085	0.291	0.290

Please see the description of Table 13.5

The ratio of the provision for loan loss to total loans is an ambiguous measure. A high provision could indicate bank management is expecting high loan losses. On the other hand, a high ratio could indicate conservative bank management that is taking no chances on an underfunded allowance for loan loss. When we look at the provision as a percentage of total equity, we again have the same ambiguous results possible. In general, we find that both of these risk measures produce a negative coefficient on our measure of capital. The ambiguity seems to be resolved in that as the provision for loan losses increases, so does risk. The alternative explanation is that risk should decrease with this ratio, but higher capital levels produce counterintuitive results. The ratio of the allowance for loan loss to total loans moves inversely with capital while the ratio of allowance for loan loss to total equity moves in the opposite direction.

We turn now to the allowance for loan losses. Like the provision for loan loss ratios, we have the same ambiguous expectations, but now we find conflicting results. There is no clear expectation for the impact of this ratio on risk. In other words, both the ratio of the allowance for loan losses to total loans and the ratio of the allowance for loan losses to total equity can be reflecting either high risk or low risk. What we find is that the allowance for loan loss as a percentage of loans produces a positive sign for the coefficient on capital, while the allowance for loan loss as a percentage of equity produces a negative sign on the coefficient.

We find even more consistent counterintuitive results when we look at the relationship between capital and liquidity risk in Table 13.7. Lower capital ratios are generally related to higher levels of short-term purchased liabilities (see STPF/A and FFP/A). Since short-term purchased money is more volatile than core deposits, for example, we would expect high levels of purchased money to be associated with high levels of capital, yet this is not what we find. On the other hand, we do find that higher capital ratios are related to more liquid assets (HLA/A) and better coverage of short-term liabilities (FFS/A). Since both of these ratios imply higher levels of liquidity, we expected them to be related to lower levels of capital. Apparently liquidity risk is not reflected in a BHC's capital level.

When we look at a BHC's exposure to interest rate risk, we again find counterintuitive results. As noted above, Table 13.7 shows that low capital ratios are related to high levels of short-term purchased funds. This can result in a fundamental liquidity problem if some markets for short-term borrowing completely dry up as we have observed in the recent financial crisis. Our more direct measure of interest rate risk is the interest-sensitive gap (gap) which we define as interest-sensitive assets minus interest-sensitive liabilities divided by total assets. Here we find ambiguous results. It is obvious that wider gaps expose banks to more risk if interest rates move against the bank. However, wide gaps can be held in both a positive and negative direction. A high positive gap indicates a BHC has a large amount of interest-sensitive assets in relation to interest-sensitive liabilities and will be hurt by falling interest rates. A high negative gap indicates a BHC has a large amount of interest-sensitive liabilities in relation to interest-sensitive assets and will be hurt if interest rates rise. Our results in Table 13.7 indicate that wider gaps are associated with lower levels of capital, but this is only a true measure if BHCs typically held a positive gap.

In Table 13.8, we turn our attention to BHCs' off-balance-sheet activity. Rather surprisingly, off-balance-sheet exposures seem to be inversely related to capital levels in spite of the Basle Capital Accords. Recall that the Basle Accords require the maintenance of capital in support of off-balance-sheet activity. Yet all of our measures of off-balance-sheet risk are associated with low capital levels with one exception. The notional amount of commodity derivatives held for trading compared with the notional amount of commodity derivatives held for other purposes is associated with a higher level of capital. Our interpretation of the ratios that measure the amount of derivatives held for trading compared with the derivatives held for other purposes is that the derivatives not held for trading are held to hedge an existing position on the books of the BHC. As a result, a high ratio implies more trading activity in relation to hedging activity, and therefore, more capital should be required. However, we again find high OBS ratios associated with low levels of capital.

In Table 13.9, we look at two measures of market risk: the size of the BHCs' trading account and the amount of unrealized gains or losses on the BHCs' investment portfolio. We again find what we believe are counterintuitive results. First, larger trading portfolios inherently contain a larger amount of market risk. On the other hand, a large amount of unrealized gains in the investment portfolio mitigates market risk, at least to some extent. We find, however, low levels of capital associated with higher trading portfolios, while high levels of capital are associated with higher unrealized gains in the investment portfolio.

Our results concerning performance measures are shown in Table 13.10. Here we find evidence of higher capital ratios being associated with higher return on assets ratios and with higher net interest spreads. Since higher returns on earning assets are logically associated with higher risk, it is appropriate that higher capital is used in support of the additional risk. However, while the direction of the causality is not clear, this could be interpreted as more evidence that BHCs increase risk to maintain a target return on equity in the face of higher capital requirements.

13.6 Conclusions

In this study, we thoroughly analyze a large cross section of bank holding company data from 1993 to 2008 to determine the relationship between capital and bank risk-taking. Our sample includes a minimum of almost 700 BHCs in 2008 and a maximum of about 1,500 BHCs in 1993. This produces nearly 25,000 company-year observations of BHCs starting with the year that risk-based capital requirements were first in place. Our data cover a period containing significant changes in the banking industry and varying levels of economic activity. The Riegle-Neal and Gramm-Leach-Bliley acts were passed during this time period, and the tech-stock and housing bubbles both burst with attendant recessions. By including a larger size range of BHCs in our analysis over a long sample period, our results are applicable to relatively small BHCs as well as to the largest 200 or so BHCs traditionally included in empirical studies.

We employ stochastic frontier analysis to create a new type of instrumental variable for capital to be used in regressions of risk and capital, thereby mitigating the obvious endogeneity problem. The instruments are validated to confirm their high correlation with capital and limited correlation with risk. We conclude that they are suitable for use in our models and employ a GMM estimator to acknowledge the non-normal distribution of the instruments.

Our results are consistent with the theory that BHCs respond to higher capital ratios by increasing the risk in the earning asset portfolios. We find an inverse relationship between the proportion of risky assets held by a bank holding company and the amount of capital they hold. We also find lower levels of capital associated with measures of credit risk that indicate a riskier loan portfolio. For example, the amount of nonperforming assets held by the bank holding company is inversely related to the bank holding company's capital.

Our findings also demonstrate a counterintuitive relationship between bank capital and liquidity risk. Less liquid banks tend to have low capital ratios, while more liquid banks tend to have high capital ratios. These same results suggest that a higher level of interest rate risk is also related to lower levels of capital. Our direct measure of the mismatch between interest-sensitive assets and liabilities provides additional evidence, although somewhat ambiguously, of higher interest rate risk being associated with lower capital. High levels of off-balance-sheet activity and of market risk exposure are likewise surprisingly related to low capital levels. Finally, we note the association of high levels of capital with high return on asset ratios. This association at least suggests that bank holding companies do increase the risk of their earning assets in order to provide an adequate return on their capital.

Our analysis adds to the existing literature with three contributions. First, we employ a large panel data set to consider the capital-risk relationship for a wider range of bank holding companies than previously reviewed. Second, stochastic frontier analysis is applied to exogenously generate the effect of the use of capital in banking. Finally, our results provide what we believe are important findings with potentially major public policy implications. If the primary tool used by bank regulators to ensure a stable financial system is, instead, creating perverse results, then alternative tools must be developed. Further exploring the relationship between the efficiency of capital and the risk strategy adopted by a bank would be a contribution to this literature.

Appendix 1: Stochastic Frontier Analysis

Stochastic frontier analysis (SFA) is an economic modeling method that is introduced by Jondrow et al. (1982). The frontier without random component can be written as the following general form:

$$PTI_i = TE_i \cdot f(\mathbf{x}_i; b), \tag{13.5}$$

where PTI_i is the pretax income of the bank i, $i = 1,..N$, TE_i denotes the technical efficiency defined as the ratio of observed output to maximum feasible output, $\mathbf{x}_i$ is a vector of J inputs used by the bank i, $f(\mathbf{x}_i, b)$ is the frontier, and b is a vector of technology parameters to be estimated. Since the frontier provides an estimate of the maximum feasible output, we then can measure the shortfall of the observed output from the maximum feasible output. Considering a stochastic component that describes random shocks affecting the production process in this model, the stochastic frontier becomes

$$PTI_i = e^{v_i} \cdot TE_i \cdot f(\mathbf{x}_i; b). \tag{13.6}$$

The shock, e^{v_i}, is not directly attributable to the bank or the technology but may come from random white noises in the economy, which is considered as a two-sided Gaussian distributed variable.

We further describe TE_i as a stochastic variable with a specific distribution function. Specifically,

$$TE_i = e^{-u_i}, \tag{13.7}$$

where u_i is the nonnegative technical inefficiency component, since it is required that $TE_i \leq 1$. Thus, we obtain the following equation:

$$PTI_i = e^{v_i - u_i} \cdot f(\mathbf{x}_i; b). \tag{13.8}$$

We then can describe the frontier according to a specific production model. In our case, we assume that bank's profitability can be specified as the log-linear Cobb-Douglas function:

$$PTI_i = a + \sum_{h=1}^{H} b_h \ln x_{i,h} + v_i - u_i. \tag{13.9}$$

Because both v_i and u_i constitute a compound error term with a specific distribution to be determined, hence the SFA is often referred as composed error model.

In our study, we use the above stochastic frontier with different inputs to generate the net effect of bank capital without mixing the impact of risk. The unrestricted model (without including bank equity) is

$$\begin{array}{l} PTI(BVA, \sigma_{\mathrm{BANK}}) = a + b_1 BVA + b_2 (BVA)^2 + e \\ e = \xi - \varsigma \\ \xi \sim iidN\left(0, \sigma_\xi^2\right), \varsigma \sim iidN\left(0, \sigma_\varsigma^2\right) \end{array}, \tag{13.10}$$

where BVA is the natural logarithm of book value of assets, ξ is statistical noise, ς is systematic shortfall (under management control), and $\varsigma \geq 0$. Our restricted model is as follows:

$$\begin{aligned} &PTI(BVA, BVC, \sigma_{\text{BANK}}) = \alpha + \beta_1 BVA + \beta_2 (BVA)^2 + \beta_3 BVC + \varepsilon. \\ &\varepsilon = v - u \\ &v \sim iidN(0, \sigma_v^2) u \sim iidN(0, \sigma_u^2)\text{'} \end{aligned} \tag{13.11}$$

where BVC is the natural logarithm of book value of capital, v is statistical noise, and u denotes the inefficiency of a bank considering its use of both assets and capital. The difference in the inefficiency between the restricted and unrestricted model,

$$\delta = u - \varsigma, \tag{13.12}$$

is our instrumental variable. The instrumental variable for capital can be used in regressions of various measures of risk, as the dependent variable, on our instrument for capital, as the independent variable, while controlling for BHC size.

Appendix 2: Generalized Method of Moments

Hansen (1982) develops generalized method of moments (GMM) to estimate parameters that its full shape of the distribution function is not known. The method requires that a certain number of moment conditions were specified for the model. These moment conditions are functions of the model parameters and the data, such that their expectation is zero at the true values of the parameters. The GMM method then minimizes a certain norm of the sample averages of the moment conditions.

Suppose the error term $\boldsymbol{\varepsilon}_t = \boldsymbol{\varepsilon}(\mathbf{x}_t, \theta)$ is a $(T \times 1)$ vector that contains T observations of the error term $\boldsymbol{\varepsilon}_t$, where $\mathbf{x}_t$ includes the data relevant for the model and θ is a vector of N_β coefficients. Assume there are N_H instrumental variables in an $(N_H \times 1)$ column vector, $\mathbf{h}_t$ and T observations of this vector form a $(T \times N_H)$ matrix $\mathbf{H}$. We define

$$\mathbf{f}_t(\theta) \equiv \mathbf{h}_t \otimes \boldsymbol{\varepsilon}(\mathbf{x}_t, \theta). \tag{13.13}$$

The notation $\otimes$ denotes the Kronecker product of the two vectors. Therefore, $\mathbf{f}_t\ (\theta)$ is a vector containing the cross product of each instrument in $\mathbf{h}$ with each element of $\boldsymbol{\varepsilon}$. The expected value of this cross product is a vector with $N_\varepsilon N_{\mathrm{H}}$ elements of zeros at the parameter vector:

$$\mathrm{E}[\mathbf{f}_t(\theta_0)] = 0. \tag{13.14}$$

Since we do not observe the true expected values of $\mathbf{f}$, thus we must work instead with the sample mean of $\mathbf{f}$,

$$\mathbf{g}_t(\theta) \equiv T^{-1} \sum_{t=1}^{T} \mathbf{f}_t(\theta) = T^{-1} \sum_{t=1}^{T} \mathbf{h}_t \boldsymbol{\varepsilon}_t(\theta) = T^{-1}\ \mathbf{H}' \boldsymbol{\varepsilon}_t(\theta). \tag{13.15}$$

We can minimize the quadratic form

$$\mathbf{Q}_T(\theta) \equiv \mathbf{g}_T(\theta)' \mathbf{W}_T \mathbf{g}_T(\theta), \tag{13.16}$$

where $\mathbf{W}_T$ is an $(N_H \times N_H)$ symmetric, positive definite weighting matrix. We then find the first-order condition is

$$\mathbf{D}_T\left(\hat{\theta}_T\right)' \mathbf{W}_T \mathbf{g}_T\left(\hat{\theta}_T\right) = 0, \tag{13.17}$$

where $\mathbf{D}_T\ (\theta_T)$ is a matrix of partial derivatives defined by

$$\mathbf{D}_T(\theta_T) = \partial \mathbf{g}_T(\theta_T)/\partial\theta'.$$

Note the above problem is nonlinear; thus, the optimization must be solved numerically.

Applying the asymptotic distribution theory, the coefficient estimate $\hat{\theta}_T$ is

$$\sqrt{T}\left(\hat{\theta}_T - \theta_0\right) \xrightarrow{d} N(0, \Omega), \tag{13.18}$$

where $\Omega = (\mathbf{D}_0{}'\mathbf{W}\mathbf{D}_0)^{-1}\ \mathbf{D}_0{}'\mathbf{W}\mathbf{S}\mathbf{W}\mathbf{D}_0\ (\mathbf{D}_0{}'\mathbf{W}\mathbf{D}_0)^{-1}$. $\mathbf{D}_0$ is a generalization of M_{HX} in those equations and is defined by $\mathbf{D}_0 \equiv \mathrm{E}[\partial \mathbf{f}(\mathbf{x}_t, \theta_0)/\partial\theta_0]$. $\mathbf{S}$ is defined as

$$\mathbf{S} \equiv \lim_{T\to\infty} \mathrm{Var}\left[T^{1/2}\sum_{t=1}^{T} \mathbf{f}_t(\theta_0)\right] = \lim_{T\to\infty} \mathrm{Var}\left[T^{1/2}\mathbf{g}_T(\theta_0)\right]. \tag{13.19}$$

The GMM estimators are known to be consistent, asymptotically normal, and efficient in the class of all estimators that do not use any extra information aside from that contained in the moment conditions. For more discussion of the execution of the GMM, please refer to Campbell et al. (1997) and Hamilton (1994).

References

Berger, A. N., Herring, R. J., & Szego, G. P. (1995). The role of capital in financial institutions. *Journal of Banking and Finance, 24*, 1383–1398.

Berger, A. N., DeYoung, R., Flannery, M. J., Lee, D., & Oztekin, O. (2008). How do large banking organizations manage their capital ratios? *Journal of Financial Services Research, 343*, 123–149.

Buser, S. A., Chen, A. H., & Kane, E. J. (1981). Federal deposit insurance, regulatory policy, and optimal bank capital. *Journal of Finance, 36*, 51–60.

Calomiris, C. W. (1999). Building an incentive-compatible safety net. *Journal of Banking and Finance, 23*, 1499–1519.

Calomiris, C. W., & Kahn, C. M. (1991). The role of demandable debt in structuring optimal banking arrangements. *American Economic Review, 81*, 497–513.

Campbell, J., Lo, A., & MacKinlay, A. C. (1997). *The econometrics of financial markets*. Princeton: Princeton University Press.

Diamond, D. W., & Dybvig, P. H. (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy, 91*, 401–419.

Diamond, D. W., & Rajan, R. (2000). A theory of bank capital. *Journal of Finance, 55*, 2431–2465.

Duan, J., Moreau, A. F., & Sealey, C. W. (1992). Fixed rate deposit insurance and risk-shifting behavior at commercial banks. *Journal of Banking and Finance, 16*, 715–742.

Esty, B. C. (1998). The impact of contingent liability on commercial bank risk taking. *Journal of Financial Economics, 47*, 189–218.

Flannery, M. J., & Rangan, K. P. (2008). What caused the bank capital build-up of the 1990s? *Review of Finance, 12*, 391–430.

Gatev, E., Schuermann, T., & Strahan, P. E. (2009). Managing bank liquidity risk: How deposit-loan synergies vary with market conditions. *Review of Financial Studies, 22*, 995–1020.

Gorton, G., & Pennacchi, G. (1992). Financial innovation and the provision of liquidity services. In J. R. Barth & R. Dan Brumbaugh (Eds.), *Reform of federal deposit insurance*. New York: Harper Collins.

Hamilton, J. D. (1994). *Time series analysis*. Princeton: Princeton University Press.

Hansen, L. P. (1982). Large sample properties of Generalized Method of Moments estimators, *Econometrica, 50*, 1029–1054.

Hovakimian, A., & Kane, E. J. (2000). Effectiveness of capital regulation at U.S. commercial banks: 1985 to 1994. *Journal of Finance, 55*, 451–468.

Hughes, J. P., Mester, L. J., & Moon, C. (2001). Are scale economies in banking elusive or illusive? Incorporating capital structure and risk into models of bank production. *Journal of Banking and Finance, 25*, 2169–2208.

Hughes, J. P., Lang, W. W., Mester, L. J., Moon, C., & Pagano, M. (2003). Do bankers sacrifice value to build empires? Managerial incentives, industry consolidation, and financial performance. *Journal of Banking and Finance, 27*, 417–447.

John, K., Saunders, A., & Senbet, L. (2000). A theory of bank regulation and management compensation. *Review of Financial Studies, 13*, 95–125.

Jondrow, J., Lovell, C. A., Materov, I. S., & Schmidt, P. (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics, 19*, 233–238.

Keeley, M. C. (1990). Deposit insurance, risk, and market power in banking. *American Economic Review, 80*, 1183–1200.

Keeley, M. C. (1992). Bank capital regulation in the 1980s: Effective or ineffective? In S. Anthony, G. F. Udell, & L. J. White (Eds.), *Bank management and regulation: A book of readings*. Mountain View: Mayfield Publishing Company.

Kim, D., & Santomero, A. M. (1988). Risk in banking and capital regulation. *Journal of Finance, 42*, 1219–1233.

Laeven, L. A., & Levine, R. (2009). Bank governance, regulation and risk taking. *Journal of Financial Economics, 93*, 259–275.

Marcus, A. J. (1984). Deregulation and bank financial policy. *Journal of Banking and Finance, 8*, 557–565.

Marcus, A. J., & Shaked, I. (1984). The valuation of FDIC deposit insurance using option-pricing estimates. *Journal of Money, Credit and Banking, 16*, 446–460.
Merton, R. C. (1977). An analytic derivation of the cost of deposit insurance and loan guarantees. *Journal of Banking and Finance, 1*, 3–11.
Prescott, E. S. (1997). The pre-commitment approach in a model of regulatory banking capital. *FRB of Richmond Economic Quarterly, 83*, 23–50.
Santos, J. A. (2001). Bank capital regulation in contemporary banking theory: A review of the literature. *Financial Markets, Institutions and Instruments, 10*, 41–84.
Shrieves, E., & Dahl, D. (1992). The relationship between risk and capital in commercial banks. *Journal of Banking and Finance, 16*, 439–457.
VanHoose, D. (2007). Theories of bank behavior under capital regulation. *Journal of Banking and Finance, 31*, 3680–3697.

Evaluating Long-Horizon Event Study Methodology

14

James S. Ang and Shaojun Zhang

Contents

Abstract

We describe the fundamental issues that long-horizon event studies face in choosing the proper research methodology and summarize findings from existing simulation studies about the performance of commonly used methods. We document in details how to implement a simulation study and report our own findings on large-size samples. The findings have important implications for future research.

We examine the performance of more than 20 different testing procedures that fall into two categories. First, the buy-and-hold benchmark approach uses

J.S. Ang (✉)
Department of Finance, College of Business, Florida State University, Tallahassee, FL, USA
e-mail: jang@cob.fsu.edu; jang@garnet.acns.fsu.edu

S. Zhang
School of Accounting and Finance, Faculty of Business, Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong
e-mail: afszhang@inet.polyu.edu.hk

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_14,

a benchmark to measure the abnormal buy-and-hold return for every event firm and tests the null hypothesis that the average abnormal return is zero. Second, the calendar-time portfolio approach forms a portfolio in each calendar month consisting of firms that have had an event within a certain time period prior to the month and tests the null hypothesis that the intercept is zero in the regression of monthly portfolio returns against the factors in an asset-pricing model. We find that using the sign test and the single most correlated firm being the benchmark provides the best overall performance for various sample sizes and long horizons. In addition, the Fama-French three-factor model performs better in our simulation study than the four-factor model, as the latter leads to serious over-rejection of the null hypothesis.

We evaluate the performance of bootstrapped Johnson's skewness-adjusted *t*-test. This computation-intensive procedure is considered because the distribution of long-horizon abnormal returns tends to be highly skewed to the right. The bootstrapping method uses repeated random sampling to measure the significance of relevant test statistics. Due to the nature of random sampling, the resultant measurement of significance varies each time such a procedure is used. We also evaluate simple nonparametric tests, such as the Wilcoxon signed-rank test or the Fisher's sign test, which are free from random sampling variation.

Keywords

Long-horizon event study • Johnson's skewness-adjusted *t*-test • Weighted least squares regression • Bootstrap test • Calendar-time portfolio approach • Fama-French three-factor model • Johnson's skewness-adjusted *t*-statistic • Large-scale simulations

14.1 Introduction

A large number of papers in finance literature have documented evidence that firms earn abnormal returns over a long time period (ranging from 1 to 5 years) after certain corporate events. Kothari and Warner (2007) report that a total of 565 papers reporting event study results were published between 1974 and 2000 in five leading journals: the *Journal of Business* (JB), *Journal of Finance* (JF), *Journal of Financial Economics* (JFE), *Journal of Financial and Quantitative Analysis* (JFQA), and the *Review of Financial Studies* (RFS). Approximately 200 of the 565 event studies use a maximum window length of 12 months or more.

The evidence of long-horizon abnormal returns contradicts the efficient market hypothesis that stock prices adjust to information fully within a narrow time window (a few days). To reconcile the contradiction, Fama (1998) argues that "Most important, consistent with the market efficiency prediction that apparent anomalies can be due to methodology, most long-term return anomalies tend to disappear with reasonable changes in technique." Several simulation studies such as Kothari and Warner (1997) and Barber and Lyon (1997) document evidence that statistical inference in long-horizon event studies is sensitive to the choice of methodology. Therefore, it is crucial to gain an understanding of the properties and limitations of the available approaches before choosing a methodology for a long-horizon event study.

At the core of a long-horizon event study lie two tasks: the first is to measure the event-related long-horizon abnormal returns, and the second is to test the null hypothesis that the distribution of these long-horizon abnormal returns concentrates around zero. A proper testing procedure for long-horizon event studies has to do both tasks well. Otherwise, two types of error could arise and lead to incorrect inference. The first error occurs when the null hypothesis is rejected, not because the event has generated true abnormal returns, but because a biased benchmark has been used to measure abnormal returns. A biased benchmark shifts the concentration of abnormal returns away from zero and leads to too many false rejections of the null hypothesis. The second error occurs when the null hypothesis is accepted, not because the event has no impact, but because the test itself does not have enough power to statistically discriminate the mean abnormal return from zero. A test with low power is undesirable, as it will lead researchers to reach the incorrect inference that long-term effect is statistically insignificant. Thus, the researchers would want a procedure that minimizes both sources of error or at least choose a balance between them.

Two approaches have been followed in recent finance literature to measure and test long-term abnormal returns. The first approach uses a benchmark to measure the abnormal buy-and-hold return for every event firm in a sample and tests whether the abnormal returns have a zero mean. The second approach forms a portfolio in each calendar month consisting of firms that have had an event within a certain time period prior to the month and tests the null hypothesis that the intercept is zero in the regression of monthly calendar-time portfolio returns against the factors in an asset-pricing model. To follow either approach, researchers need to make a few choices as illustrated in Fig. 14.1. For the calendar-time portfolio approach, researchers choose an asset-pricing model and an estimation technique to fit the model. Among the most popular asset-pricing models are Fama and French's (1993) three-factor model and its four-factor extension proposed by Carhart (1997) that includes an additional momentum-related factor. Two techniques are commonly used to fit the pricing model: the ordinary least squares (OLS) technique and the weighted least squares (WLS) technique. On the other hand, if adopting the buy-and-hold benchmark approach, researchers choose either a reference portfolio or a single control firm as the benchmark for measuring abnormal returns and select either parametric or nonparametric statistic for testing the null hypothesis of zero abnormal return.

Permutations of these choices under both approaches generate a large number of possible testing procedures that can be used in a long-horizon event study. It is neither practical nor sensible to implement all the testing procedures in an empirical study of a financial event. Therefore, it would be very useful to provide guidance on the strength and weakness of the procedures based on simulation results. Simulation study generates large number of repetitions under various circumstances for each testing procedure, which allows the tabulations of these two types of error for comparison.

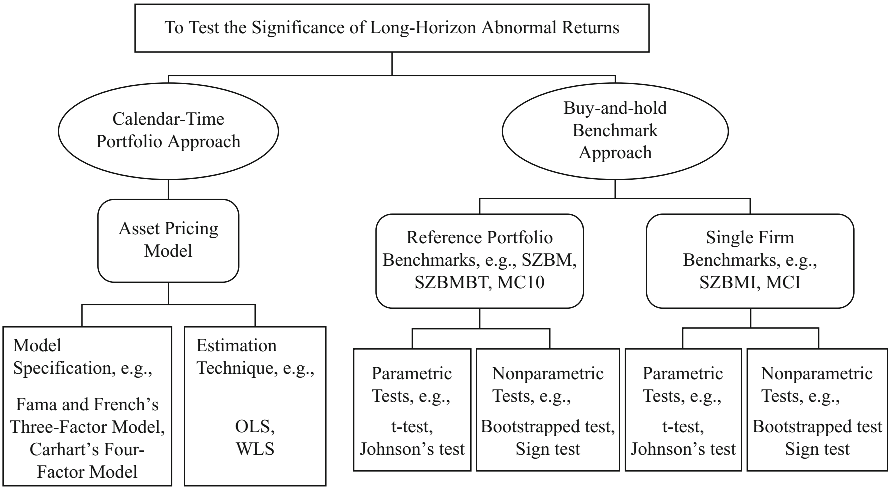


Fig 14.1 Overview of the two approaches to choose a methodology for long-horizon event study

We organize this chapter as follows. Section 14.2 discusses the fundamental issues in long-horizon event studies that have been documented in the literature. Section 14.3 reviews existing simulation studies. Section 14.4 reports results from a simulation study of large-size samples. Section 14.5 contains some suggestions for future research.

14.2 Fundamental Issues in Long-Horizon Event Studies

14.2.1 The Buy-and-Hold Benchmark Approach

The long-term buy-and-hold abnormal return of firm i, denoted as AR_i, is calculated as

$$AR_i = R_i - BR_i, \tag{14.1}$$

where R_i is the long-term buy-and-hold return of firm i and BR_i is the long-term return on a particular benchmark of firm i. The buy-and-hold return of firm i over τ months is obtained by compounding monthly returns, that is,

$$R_i = \prod_{t=1}^{\tau}(1 + r_{it}) - 1, \tag{14.2}$$

where r_{it} is firm i's return in month t. Calculation of the benchmark return BR_i is given below. The benchmark return, BR_i, estimates the return that an event firm would have had if the event had not happened.

Several articles clearly show that long-term abnormal returns are very sensitive to choice of benchmarks; see, e.g., Ikenberry et al. (1995), Kothari and Warner (1997), Barber and Lyon (1997), and Lyon et al. (1999). If wrong benchmarks were used in measuring long-term abnormal returns, inference on the significance of a certain event would be erroneous. Most existing studies use either a single matched firm or a matched reference portfolio as the benchmark. Barber and Lyon (1997) point out that the control firm approach eliminates the new listing bias, the rebalancing bias, and the skewness problem. It also yields well-specified test statistics in virtually all the situations they consider. Further, Lyon et al. (1999) advocate a reference portfolio of firms that match on size and BE/ME. The issue on choice of the benchmark is practically unresolved. Ang and Zhang (2004) additionally argue that the control firm method overcomes another important problem that is associated with the event firm not being representative in important aspects of the respective matched portfolio in the reference portfolio approach. This leads to the matched portfolio return generating a biased estimate of expected firm return. This problem is particularly severe with small firms.

A common practice in computing an event firm's long-term abnormal return is to utilize a benchmark that matches the event firm on size and BE/ME. The practice is often justified by quoting the findings in Fama and French (1992) that size and BE/ME combine to capture the cross-sectional variation in average monthly stock returns and that market beta has no additional power in explaining cross-sectional return differences. However, in a separate paper, Fama and French (1993) demonstrate that expected monthly stock returns are related to three factors: a market factor, a size-related factor, and a book-to-market equity ratio (BE/ME)-related factor. To resolve this issue, Ang and Zhang (2004) show that matching based on beta in addition to size and BE/ME does not improve the performance of the approach.

A recent trend is to use computation-intensive bootstrapping-based tests, such as the bootstrapped Johnson's skewness-adjusted *t*-statistic (e.g., Sutton 1993 and Lyon et al. 1999) and the simulated empirical p-values (e.g., Brock et al. 1992 and Ikenberry et al. 1995). These procedures rely on repeated random sampling to measure the significance of relevant test statistics. Due to the nature of random sampling, the resultant measurement of significance varies every time such a procedure is used. As a consequence, different researchers could reach contradictory conclusions using the same procedure on the same sample of event firms. In contrast, simple nonparametric tests, such as the Wilcoxon signed-rank test or the Fisher's sign test, are free from random sampling variation. Barber and Lyon (1997) examined the performance of the Wilcoxon signed-rank test in a large-scale simulation study. They show that the performance depends on choice of the benchmark. The signed-rank test is well specified when the benchmark is a single size and BE/ME matched firm and misspecified when the benchmark is a size and BE/ME matched reference portfolio. However, Barber and Lyon (1997) present only simulation results for 1-year horizon. *No simulation study in the finance literature has examined the performance of these simple nonparametric tests for 3- or 5-year horizons, which are the common holding periods in long-horizon event studies.*[1]

[1]The sign test has an advantage over the signed-rank test in that it does not require a symmetric underlying distribution, while the signed-rank test does.

Power is an important consideration in statistical hypothesis testing. Lyon et al. (1999) report that bootstrapping-based tests are more powerful than Student's *t*-test in testing 1-year abnormal returns in a large-scale simulation study. However, they do not report evidence on the power of these tests for the longer 3- or 5-year horizon. In statistics literature, bootstrapping is primarily for challenging situations when the sampling distribution of the test statistic is either indeterminate or difficult to obtain and that bootstrapping is less powerful in hypothesis testing than other parametric or simple nonparametric methods when both bootstrapping and other methods are applicable (see, e.g., Efron and Tibshirani 1993, Chap. 16 and Davison and Hinkley 1997, Chap. 4). In a recent study on 5-year buy-and-hold abnormal returns to holders of the seasoned equity offerings, Eckbo et al. (2000) note that bootstrapping gives lower significance level relative to the Student's *t*-test.

Ang and Zhang (2004) find that most testing procedures have very low power for samples of medium size over long event horizons (3 or 5 years). This raises concern about how to interpret long-horizon event studies that fail to reject the null hypothesis. Failure to reject is often interpreted as evidence that supports the null hypothesis. However, when power of the test is low, such interpretation may no longer be warranted. This problem gets even worse when event firms are primarily small firms. They observe that all tests, except the sign test, have much lower power for samples of small firms.

More recently, Schultz (2003) argue via simulation that the long-run IPO underperformance could be related to the endogeneity of the number of new issues. Firms choose to go IPO at the time when they expect to obtain high valuation in the stock market. Therefore, IPOs cluster after periods of high abnormal returns on new issues. In such a case, even if the ex ante returns on IPO are normal, the ex post measures of abnormal returns may be negative on average. Schultz suggests using calendar-time returns to overcome the bias. However, Dahlquist and de Jong (2008) find that it is unlikely that the endogeneity of the number of new issues explains the long-run underperformance of IPOs. Viswanathan and Wei (2008) present a theoretical analysis on event abnormal returns when returns predict events. They show that, when the sample size is fixed, the expected abnormal return is negative and becomes more negative as the holding period increases. This implies that there is a small-sample bias in the use of long-run event returns. Asymptotically, abnormal returns converge to zero provided that the process of the number of events is stationary. Nonstationarity in the process of the number of events is needed to generate a large negative bias.

The issues discussed above are associated with the buy-and-hold approach to testing long-term abnormal returns.[2] In addition, this approach suffers from the cross-correlation problem and the bad model problem (Fama 1998; Brav 1999; Mitchell and Stafford 2000). The cross-correlation problem arises because matching on firm-specific characteristics fails to completely remove the correlation between

[2]Variations of this approach have been used extensively; see, e.g., Ritter (1991); Ikenberry et al. (1995); Ikenberry et al. (1996); and Desai and Jain (1997), among many others.

event firms' returns. The bad model problem arises because no benchmark gives perfect estimate of the *counterfactual* (i.e., what if there was no event) return of an event firm and benchmark errors are multiplied in computing long-term buy-and-hold returns. Therefore, Fama (1998) advocates a calendar-time portfolio approach.[3]

14.2.2 The Calendar-Time Portfolio Approach

In the calendar-time portfolio approach, for each calendar month, an event portfolio is formed, consisting of all firms that have experienced the same event within the τ months prior to the given month. Monthly return of the event portfolio is computed as the equally weighted average of monthly returns of all firms in the portfolio. Excess returns of the event portfolio are regressed on the Fama-French three factors as in the following model:

$$R_{pt} - R_{ft} = \alpha + \beta\left(R_{mt} - R_{ft}\right) + sSMB_t + hHML_t + \varepsilon_t, \tag{14.3}$$

where R_{pt} is the event portfolio's return in month t; R_{ft} is the 1-month Treasury bill rate, observed at the beginning of the month; R_{mt} is the monthly market return; SMB_t is the monthly return on the zero investment portfolio for the common size factor in stock returns; and HML_t is the monthly return on the zero investment portfolio for the common book-to-market equity factor in stock returns.[4] Under the assumption that the Fama-French three-factor model provides a complete description of expected stock returns, the intercept, α, measures the average monthly abnormal return on the portfolio of event firms and should be equal to zero under the null hypothesis of no abnormal performance.

A later modification that has gained popularity is the four-factor model that added a momentum-related factor to the Fama-French three factors:

$$R_{pt} - R_{ft} = \alpha + b\left(R_{mt} - R_{ft}\right) + sSMB_t + hHML_t + pPR12_t + \varepsilon_t, \tag{14.4}$$

where $PR12_t$ is the momentum-related factor advocated by Carhart (1997). Typically, we compute $PR12_t$ by first ranking all firms by their previous 11-month stock return lagged 1 month and then taking the average return of the top one third (i.e., high past return) stocks minus the average return of the bottom one third (i.e., low past return) stocks.

Under the assumption that the asset-pricing model adequately explains variation in expected stock returns, the intercept, α, measures the average monthly abnormal return of the calendar-time portfolio of event firms and should be equal to zero under the null hypothesis of no abnormal performance. If the test concludes that the

[3]Loughran and Ritter (1995), Brav and Gompers (1996), and Brav et al. (2000), among others, have used the calendar-time portfolio approach.

[4]See Fama and French (1993) for details on construction of the mimicking portfolios for the common size and book-to-market equity factors. We thank Eugene Fama for providing us with returns on R_{ft}, R_{mt}, SMB_t, and HML_t.

time series conforms to the asset-pricing model, the event is said to have had no significant long-term effect; otherwise, the event has produced significant long-term abnormal returns. Lyon et al. (1999) report that the calendar-time portfolio approach together with the Fama-French three-factor model, which shall be referred to as the Fama-French calendar-time approach later, is well specified for random samples in their simulation study.

However, we do not know how much power the Fama-French calendar-time approach has. Loughran and Ritter (1999) criticize the approach as having very low power. They argue that reduction in power is caused by using returns on contaminated portfolios as factors in the regression, by weighting each month equally and by using value-weighted returns of the calendar-time portfolios. However, their empirical evidence is based only on one carefully constructed sample of firms and is hardly conclusive. No large-scale simulation study has been done to examine power of the Fama-French calendar-time approach, which we will remedy in this paper.

The Fama-French calendar-time approach, estimated with the ordinary least squares (OLS) technique, could suffer from a potential heteroskedasticity problem due to unequal and changing number of firms in the calendar-time portfolios. The weighted least squares (WLS) technique, which is helpful in addressing the heteroskedasticity problem, has been suggested as a way to deal with the changing size of calendar-time portfolios. When applying WLS, we use the monthly number of firms in the event portfolio as weights.

14.3 A Review of Simulation Studies on Long-Horizon Event Study Methodology

Several papers have documented performance of testing procedures in large-scale simulations. Table 14.1 surveys these papers with reference to testing procedures under their investigation and their simulation settings. The simulation technique was pioneered by Brown and Warner (1980, 1985) to evaluate size and power of testing procedures. In this section, we review these simulation studies.

As shown in Fig. 14.1, there are two approaches for a long-term event study: the calendar-time portfolio approach versus the buy-and-hold benchmark approach. There has been a debate on which approach prescribes the best procedure for long-term event studies. Both approaches have been under criticisms. The buy-and-hold benchmark approach is susceptible to biases associated with cross-sectional correlation, insufficient matching criteria, new equity issues, periodic balancing, and skewed distribution of long-term abnormal returns, while the calendar-time portfolio approach may suffer from an improper asset-pricing model and heteroskedasticity in portfolio returns. See Kothari and Warner (1997), Barber and Lyon (1997), Fama (1998), Loughran and Ritter (1999), Lyon et al. (1999); and others for more detailed discussions. Kothari and Warner (1997) argue that the combined effect of these issues is difficult to specify a priori and, thus, "a simulation study with actual security return data is a direct way to study the joint impact, and is helpful in identifying the potential problems that are empirically most relevant."

Table 14.1 Summary of existing simulation studies

	Procedures under investigation				Validation settings		
	Calendar-time portfolio approach		Buy-and-hold benchmark approach				
Authors (year)	Pricing model	Estimation method	Matching criteria	Test statistics	Simulation design	Evidence on specification at horizon of	Evidence on power at horizon of
Kothari and Warner (1997)	This paper examines procedures that are based on cumulating monthly abnormal returns. Such procedures are severely misspecified in most cases and are not recommended				250 simulated samples of 200 firms each	1 month, 1, 2, and 3 years	1 year
Barber and Lyon (1997)			Size, BE/ME	*t*-test, Wilcoxon test	1,000 simulated samples of 200 firms each	1, 3, and 5 years	1 year
Lyon et al. (1999)	Three-factor model	OLS	Size, BE/ME	*t*-test, Johnson's test, bootstrapped test	1,000 simulated samples of 200 firms each	1, 3, and 5 years	1 year, only for the buy-and-hold approach
Mitchell and Stafford (2000)	Three-factor model	OLS			1,000 simulated samples of 2,000 firms each	3 years	3 years
Cowan and Sergeant (2001)			Size, BE/ME	*t*-test, two-group test with winsorized data	1,000 simulated samples with sample size of 50, 200, and 1,000	1, 3, and 5 years	3 years
Ang and Zhang (2004)	Three-factor model, four-factor model	OLS, WLS	Size, BE/ME, beta, and correlation coefficient	*t*-test, Johnson's test, bootstrapped test, sign test	250 simulated samples with sample size of 200 and 1,000	1, 3, and 5 years	1, 3, and 5 years
Jegadeesh and Karceski (2009)			Size, BE/ME	*t*-test, *t*-test adjusted for heteroskedasticity and serial correlation	1,000 simulated samples of 200 firms each	1, 3, and 5 years	1, 3, and 5 years

In their simulation study, Kothari and Warner (1997) measure the long-term (up to 3 years) impact of an event by cumulative monthly abnormal returns, where monthly abnormal returns are computed against four common models: the market-adjusted model, the market model, the capital asset-pricing model, and the Fama-French three-factor model. They find that tests for cumulative abnormal returns are severely misspecified. They identify sample selection, survival bias, and bias in variance estimation as potential sources of the misspecification and suggest that nonparametric and bootstrap tests are likely to reduce misspecification.

Barber and Lyon (1997) address two main issues in their simulation study. First, they argue that buy-and-hold return is a better measure of investors' actual experience over a long horizon and should be used in long-term event study (up to 5 years). They show simulation evidence that approaches using cumulative abnormal returns cause severe misspecification, which is consistent with the observation in Kothari and Warner (1997). Second, they use simulations to measure both size and power of testing procedures that follow the buy-and-hold benchmark approach. An important finding is that using a single control firm as benchmark yields well-specified tests, whereas using reference portfolio causes substantial over-rejection.

In a later paper, Lyon et al. (1999) report another simulation study (for up to the 5-year horizon) that investigates the performance of both buy-and-hold benchmark approach and calendar-time portfolio approach. They find that using the Fama-French three-factor model yields a well-specified test. However, they advocate a test that uses carefully constructed reference portfolio as benchmark and the bootstrapped Johnson's statistic for testing abnormal returns. They present evidence that this test is well specified and has high power at the 1-year horizon.

Two questions remain unanswered in Lyon et al. (1999). First, how much power does the bootstrap test have for event horizons longer than 1 year (e.g., 3 or 5 years that is common in long-horizon studies)? It is known in statistics literature that a bootstrap test is not as powerful as simple nonparametric tests in many occasions (see Efron and Tibshirani 1993, Chap. 16 and Davison and Hinckley 1997, Chap. 4). It is necessary to know the actual power of such test for event horizons beyond 1 year. Second, is the calendar-time portfolio approach as powerful as the buy-and-hold benchmark approach? Loughran and Ritter (2000) argue that the calendar-time portfolio approach has low power, using simulations and empirical evidence from a sample of new equity issuers. However, they do not measure how much power the approach actually has, which makes it impossible to compare the two approaches directly in more general settings.

Mitchell and Stafford (2000) is the only study that empirically measures power of the calendar-time portfolio approach using simulations. Their main focus is to assess performance of several testing procedures in three large samples of major managerial decisions, i.e., mergers, seasoned equity offerings, and share repurchases (up to 3 years). They find that different procedures lead to contradicting conclusions and argue that the calendar-time portfolio approach is preferred. To resolve Loughran and Ritter's (2000) critique that the calendar-time portfolio approach has low power, they conduct simulations to measure the empirical power and find that the power is actually very high with an empirical rejection rate of 99 % for induced abnormal returns of ±15 % over a 3-year horizon. Since they have a large sample size, this finding is actually consistent with what we document in Table 14.5. However, their simulations focus on only samples of 2,000 firms. Many event studies have much smaller sample sizes, especially after researchers slice and dice a whole sample into subsamples. More evidence is needed in order to have great confidence in applying the calendar-time portfolio approach in such studies.

Cowan and Sergeant (2001) focus on the buy-and-hold benchmark approach in their simulations. They find that using the reference portfolio approach cannot overcome the skewness bias discussed in Barber and Lyon (1997) and that the larger the sample size, the smaller the magnitude of the skewness bias. They also argue that cross-sectional dependence among event firms' abnormal returns increases in event horizon due to partially contemporaneous holding periods, which may cause the overlapping horizon bias. They propose a two-group test using abnormal returns winsorized at three standard deviations to deal with these two biases and report evidence that this test yields correct specifications and considerable power in many situations.

All previous simulation studies use only size and BE/ME to construct benchmarks, which is often justified by the findings in Fama and French (1992) that size and BE/ME together adequately capture the cross-sectional variations in average monthly stock returns. Ang and Zhang (2004) use two other matching criteria to explore whether better benchmarks could be used for future studies. The two criteria are market beta and pre-event correlation coefficient. Using market beta is motivated by the fact that Fama and French's (1993) three-factor model has a market factor, a size-related factor, and a BE/ME-related factor. Matching on the basis of size and BE/ME does not account for the influence of the market factor. The rationale for using pre-event correlation coefficient is that matching on size and BE/ME may fail to control for other factors that could influence stock returns, such as industry factor, seasonal factor, momentum factor, and other factors shared by only firms of the same characteristics, such as geographical location, ownership, and governance structures. Matching on the basis of pre-event correlation coefficient helps remove the effect of these factors on the event firm's long-term return.

The main findings in Ang and Zhang (2004) include the following. First, the four-factor model is inferior to the well-specified three-factor model in the calendar-time portfolio approach in that the former causes too many rejections of the null hypothesis relative to the specified significance level. Second, WLS improves the performance of the calendar-time portfolio approach over OLS, especially for long event horizons. Third, the Fama-French three-factor model has relatively high power in detecting abnormal returns, although power decreases sharply as event horizon increases. Fourth, the simple sign test is well specified when it is applied with a single firm benchmark, but misspecified when used with reference portfolio benchmarks. More importantly, the combination of the sign test and the benchmark with the single most correlated firm consistently has much higher power than any other test in our simulations and is the only testing procedure that performs well in samples of small firms.

Jegadeesh and Karceski (2009) propose a new test of long-run performance that allows for heteroskedasticity and autocorrelation. Previous tests used in Lyon et al. (1999) implicitly assume that the observations are cross-sectionally uncorrelated.

This assumption is frequently violated in nonrandom samples such as samples with industry clustering or with overlapping returns. To overcome the cross-correlation bias in event firms' returns, they recommend a t-statistic that is computed using a generalized version of the Hansen and Hodrick (1980) standard error. Their simulation studies show that the new tests they propose are reasonably well specified in random samples, in samples that are concentrated in particular industries, and also in samples where event firms enter the sample on multiple occasions within the holding period.

In summary, these simulation studies show that testing procedures differ dramatically in performance. Some procedures reject the null hypothesis at an excessively high rate, while others have very low power. These findings confirm the Fama (1998) statement that evidence for long-term return anomalies is dependent upon methodology and suggest that caution must be exercised in choosing the proper methodology for a long-term event study.

14.4 A Simulation Study of Large-Size Samples

A simulation study of large-size samples serves two purposes. First, it is well documented that the distribution of buy-and-hold abnormal returns tends to be skewed to the right. Kothari and Warner (2007) mention that the extent of skewness bias is likely to decline with sample size. It is of interest to provide evidence on how much is the level of right-skewness in the average abnormal returns of large-size samples. Second, although it is expected that testing power increases with sample size, it is of practical interest to know more precisely how much power a test can have in a sample of 1,000 observations. Large sample simulation defines the limits of a procedure.

14.4.1 Research Design

In this simulation study, we construct 250 samples each consisting of 1,000 event firms. To produce one sample, we randomly select, with replacement, 1,000 event months between January 1980 and December 1992, inclusively.[5,6] This allows us to calculate 5-year abnormal returns until December 1997. For each selected event month, we randomly select, without replacement, one firm from a list of qualified firms. The qualified firms satisfy the following requirements: (i) They are publicly traded firms, incorporated in the USA, and have ordinary common shares with Center for Research in Security Prices (CRSP) share codes 10 and 11; (ii) they have return data found in the CRSP monthly returns database for the 24-month period

[5]We use a pseudorandom number generator developed by Matsumoto and Nishimura (1998) to ensure high quality of random sampling.

[6]Kothari and Warner (1997) use 250 samples, each of 200 event months between January 1980 and December 1989 inclusively. Barber and Lyon (1997) use 1,000 samples, each of 200 event months in a much longer period from July 1963 to December 1994. The period under our study, between January 1980 and December 1992, is of similar length to Kothari and Warner's.

prior to the event month; (iii) they have nonnegative book values on COMPUSTAT prior to the event month so that we can calculate their book-to-market equity ratios.

The 250 samples, each of 1,000 randomly selected firms, comprise the simulation setting for comparing the performance of different testing procedures.[7] We apply all testing procedures under our study to the same samples. Such controlled comparison is more informative because it eliminates difference in performance due to variation in the samples.

For the buy-and-hold approach, we compute the long-term buy-and-hold abnormal return of firm i as the difference between the long-term buy-and-hold return of firm i and the long-term return of a benchmark. The buy-and-hold return of firm i over τ months is obtained by compounding monthly returns. In case that firm i does not have return data for all τ months, we replace missing returns by the same-month returns of a size and BE/ME matched reference portfolio.[8] We evaluate a total of five benchmarks and four test statistics in this study. We briefly describe them in the following and give the details in the Appendix.

Three of the benchmarks are reference portfolios. The *first* reference portfolio consists of firms that are similar to the event firm in both size and BE/ME. We follow the same procedure as in Lyon et al. (1999) to construct the two-factor reference portfolio. We use the label "SZBM" for this benchmark. The *second* reference portfolio consists of firms that are similar to the event firm not only in size and BE/ME but also in market beta. We use the label "SZBMBT" for this benchmark. The *third* reference portfolio consists of ten firms that are most correlated with the event firm prior to the event. We use the label "MC10" for this benchmark.

The other two of the five benchmarks consist of a single firm. The *first* single firm benchmark is the firm that matched the event firm in both size and BE/ME. To find the two-factor single firm benchmark, we first identify all firms whose market value is within 70–130 % of the event firm's market value and then choose the firm that has the BE/ME ratio closest to that of the event firm. We use the label "SZBM1" for this benchmark. The *second* single firm benchmark is the firm that has the highest correlation coefficient with the event firm prior to the event. We use the label "MC1" for this benchmark.

We apply four test statistics to test the null hypothesis that the mean long-term abnormal return is zero. They include Student's t-test, Fisher's sign test, Johnson's skewness-adjusted t-test, and the bootstrapped Johnson's t-test. Fisher's sign test is a nonparametric test and is described in details in Hollander and Wolfe (1999, Chap. 3). Johnson's skewness-adjusted t-statistic was developed by Johnson (1978) to deal with the skewness-related misspecification error in Student's t-test. Sutton (1992) proposes to apply Johnson's t-test with a computationally intensive bootstrap resampling

[7]Ang and Zhang (2004) examine two other simulation settings. Under one setting, they have another 250 samples of 200 event firms, a smaller sample size than the setting in this chapter. Under the other setting, they have the sample size of 200 with the requirement that event firms belong to the smallest quintile sorted by NYSE firm size. The second setting is used to examine the effect of small firms.

[8]Filling in missing returns is a common practice in calculating long-term buy-and-hold returns; e.g., see Barber and Lyon (1997), Lyon et al. (1999), and Mitchell and Stafford (2000).

technique when the population skewness is severe and the sample size is small. Lyon et al. (1999) advocate use of the bootstrapped Johnson's *t*-test because long-term buy-and-hold abnormal returns are highly skewed when buy-and-hold reference portfolios are used as benchmarks. We follow Lyon et al. (1999) and set the resampling size in the bootstrapped Johnson's *t*-test to be one quarter of the sample size.

For the Fama-French calendar-time approach, we use both the Fama-French three-factor model and the four-factor model. We apply both ordinary least squares (OLS) and weighted least squares (WLS) techniques to estimate parameters in the pricing model. The WLS is used to correct the heteroskedasticity problem due to the monthly variation in the number of firms in the calendar-time portfolio. When applying WLS, we use the number of event firms in the portfolio as weights.

14.4.2 Simulation Results for the Buy-and-Hold Benchmark Approach

In this section, we examine the performance of testing procedures that follow the buy-and-hold benchmark approach. Implementation of the buy-and-hold benchmark approach involves choosing both benchmark and test statistic. For this reason, rather than focusing on what is the best among all benchmarks, or focusing on what is the best among all test statistics, we address the more practical question of finding the best combination of benchmark and test statistic. Combination of the five benchmarks and the four test statistics yields 20 testing procedures, out of which we look for the best combination.

For each sample of 1,000 abnormal returns, we compute mean, median, standard deviation, interquartile range, skewness coefficient, and kurtosis coefficient. Table 14.2 reports the average of these statistics over 250 samples.

Since these event firms, being randomly selected, may not experience any event or may experience events that have offsetting effects on averaged stock returns, we expect their abnormal returns to concentrate around zero. In Table 14.2, means are close to zero for all five benchmarks at all three holding periods, but medians differ systematically according to the type of benchmark used. Medians are clearly negative under the three reference portfolio benchmarks (i.e., SZBM, SZBMBT, and MC10), but close to zero under the two single firm benchmarks (i.e., SZBM1 and MC1). The evidence suggests that reference portfolio benchmarks overestimate holding period returns of many event firms, resulting in far too many event firms having negative abnormal returns under the portfolio-based benchmarks. The extent of the overestimation bias by portfolio-based benchmarks is quite severe and gets worse as the time horizon lengthens. The bias, as measured by the magnitude of median, ranges from around 4 % at a 1-year horizon to 12 % at a 3-year horizon and to more than 20 % at a 5-year horizon. Bias of this magnitude could cause too many events to be falsely identified as having significant long-term impact.

Volatility of abnormal returns increases with the length of holding period under all five benchmarks. For the same holding period, volatility is higher under the two single firm benchmarks than under the three reference portfolio benchmarks. This is expected

Table 14.2 Descriptive statistics of abnormal returns in samples of *1,000 firms*

	Descriptive statistics					
Benchmark	Mean	Median	Standard deviation	Interquartile range	Skewness coefficient	Kurtosis coefficient
Panel A: 1-year holding period						
SZBM	0.009	−0.032	0.574	0.453	4.332	60.763
SZBMBT	−0.001	−0.043	0.586	0.462	4.074	58.462
MC10	0.000	−0.040	0.591	0.463	3.853	56.733
SZBM1	0.005	0.005	0.814	0.638	−0.203	53.034
MC1	0.002	−0.003	0.780	0.584	0.229	53.202
Panel B: 3-year holding period						
SZBM	0.034	−0.112	1.240	0.963	4.561	57.644
SZBMBT	−0.001	−0.139	1.264	0.982	4.258	54.616
MC10	0.000	−0.126	1.286	0.982	3.996	53.153
SZBM1	0.023	0.022	1.746	1.305	−0.137	51.176
MC1	0.016	−0.006	1.658	1.200	0.736	43.430
Panel C: 5-year holding period						
SZBM	0.068	−0.209	2.034	1.490	5.287	67.521
SZBMBT	0.002	−0.248	2.073	1.514	4.982	64.364
MC10	0.007	−0.223	2.106	1.516	4.652	61.091
SZBM1	0.054	0.039	2.802	1.979	0.269	41.428
MC1	0.036	0.000	2.745	1.834	0.500	50.365

This table reports descriptive statistics that characterize the probability distribution of long-term abnormal returns, in samples of *1,000 firms*. Abnormal return is calculated as the difference in holding period return between the event firm and its benchmark. We use five benchmarks: a reference portfolio matched by size and BE/ME (SZBM); a reference portfolio matched by size, BE/ME, and beta (SZBMBT); a reference portfolio consisting of ten firms, within the event firm's size and BE/ME matched portfolio, whose returns are most correlated with the event firm's MC10; a single firm matched by size and BE/ME (SZBM1); and a single firm, from the event firm's size and BE/ME matched portfolio, whose returns have the highest correlation with the event firm's MC1. We compute mean, median, standard deviation, interquartile range, skewness coefficient, and kurtosis coefficient for abnormal returns in every sample. Since there are 250 samples in the simulation, entries in the table are the average of these statistics over the 250 samples

because reference portfolios have lower volatility due to averaging. As for kurtosis, all five benchmarks produce highly leptokurtic abnormal returns, with kurtosis coefficients ranging from 41.4 to 67.5, which are far greater than three, the kurtosis coefficient of any normal distribution. At last, skewness coefficients for the two single firm benchmarks are close to zero regardless of event horizons, while skewness coefficients for the three portfolio benchmarks are excessively positive.

To sum up, probability distributions of long-term abnormal returns exhibit different properties, depending on whether the benchmark is a reference portfolio or a single firm. Under a reference portfolio benchmark, the distribution is highly leptokurtic and positively skewed, with a close-to-zero mean but a highly negative median. Under a single firm benchmark, the distribution is highly leptokurtic but symmetric, with both mean and median close to zero. Statistical properties of long-term abnormal returns have important bearings on performance of test statistics. Overall, it seems single firm benchmarks have more desirable properties. Between the two single firm benchmarks, MC1 shows better performance than SZBM1, because the abnormal returns based on MC1 have both mean and median being closer to zero and smaller standard deviation.

A superior test should control for the probability of committing two errors. First, it is important to control for the probability of misidentifying an insignificant event as having statistical significance; in other words, the empirical size of the test, which is computed from simulations, is close to the prespecified significance level at which the test is conducted. When this happens, the test is well specified. Second, power of the test should be large, that is, the probability of finding a statistically significant event if one did exist.

Table 14.3 reports empirical size of all 20 tests for three holding periods. Empirical size is calculated as the proportion of 250 samples that rejects the null hypothesis at the 5 % nominal significance level. With only a few exceptions, Student's *t*-test is well specified against the two-sided alternative hypothesis. Despite excessively high skewness in abnormal returns from reference portfolio benchmarks, Student's *t*-test is well specified against two-sided alternative hypothesis because the effect of skewness at both tails cancels out (see, e.g., Pearson and Please 1975). When testing against the two-sided alternative hypothesis, Johnson's skewness-adjusted *t*-test is in general misspecified, but its bootstrapped version is well specified in most situations. The sign test is misspecified when applied to abnormal returns from reference portfolio benchmarks, and the extent of misspecification is quite serious and increases in the length of holding period. This is not surprising because abnormal returns from reference portfolio benchmarks have highly negative medians.

Table 14.4 reports empirical power of testing the null hypothesis of zero abnormal return against the two-sided alternative hypothesis. We follow Brown and Warner (1980, 1985) to measure empirical power by intentionally forcing the mean abnormal return away from zero with induced abnormal returns. We induce nine levels of abnormal returns ranging from −20 % to 20 % at an increment of 5 %. To induce an abnormal return of −20 %, for example, we add −20 % to the observed holding period return of an event firm. Empirical power is calculated as the proportion of 250 samples that rejects the null hypothesis at 5 % significance level.

With a large sample size of 1,000, the power of these tests remains reasonably high at the longer holding period. Ang and Zhang (2004) report that, with the sample size of 200, the power of all tests deteriorates sharply as holding period lengthens from 1 to 3 and to 5 years and is alarmingly low at the 5-year horizon. For example, when the induced abnormal return is −20 % over a 5-year horizon, the highest power of the bootstrapped Johnson's *t*-test is 13.6 % for a sample of 200 firms, whereas the highest power is 62.8 % for a sample of 1,000 firms.

We compare the power of the three test statistics: Student's *t*-test, the bootstrapped Johnson's skewness-adjusted *t*-test, and the sign test. All three test statistics are applied together with the most correlated single firm benchmark. The evidence shows that all three tests are well specified. However, the sign test clearly has much higher power than the other two tests.

Table 14.3 Specification of tests in samples of *1,000 firms*

Benchmark	Two-tailed test				Lower-tailed test				Upper-tailed test			
	t	Jt	BJt	sign	t	Jt	BJt	sign	t	Jt	BJt	sign
Panel A: 1-year holding period												
SZBM	4.0	7.2	6.4	75.6*	2.8	2.0*	1.6*	85.6*	8.4*	15.2*	13.6*	0.0*
SZBMBT	5.2	6.4	4.0	92.0*	9.6*	9.2*	7.6	96.0*	2.0*	4.8	4.0	0.0*
MC10	5.6	6.4	6.4	85.6*	10.0*	8.0*	6.8	92.8*	2.4*	5.6	4.8	0.0*
SZBM1	4.4	5.6	3.6	4.0	2.8	4.4	2.8	1.6*	6.0	8.0*	6.4	10.8*
MC1	3.6	5.2	3.2	9.6*	6.0	8.0*	4.8	12.8*	6.8	7.6	6.4	2.4
Panel B: 3-year holding period												
SZBM	11.2*	14.4*	12.8*	99.6*	1.2*	1.2*	0.8*	100.0*	17.6*	22.8*	21.6*	0.0*
SZBMBT	5.2	5.2	5.6	100.0*	7.6	7.6	5.6	100.0*	2.8	3.2	3.6	0.0*
MC10	4.8	6.8	5.6	100.0*	6.8	5.6	4.4	100.0*	3.2	6.0	5.6	0.0*
SZBM1	6.0	7.6	5.2	9.2*	2.0*	2.8	1.6*	0.4*	11.2*	13.6*	10.0*	15.6*
MC1	6.8	8.4*	6.4	6.4	2.8	2.8	2.8	7.6	7.6	8.0*	6.8	1.2*

(*continued*)

Table 14.3 (continued)

	Two-tailed test				Lower-tailed test				Upper-tailed test			
Benchmark	t	Jt	BJt	sign	t	Jt	BJt	sign	t	Jt	BJt	sign
Panel C: 5-year holding period												
SZBM	17.6*	20.4*	19.6*	100.0*	0.4*	0.4*	0.4*	100.0*	22.8*	28.8*	28.0*	0.0*
SZBMBT	3.6	4.4	2.8	100.0*	5.6	4.0	3.2	100.0*	2.8	5.6	3.6	0.0*
MC10	2.0*	4.4	2.8	100.0*	3.6	3.2	2.4	100.0*	3.6	6.4	5.6	0.0*
SZBM1	8.0*	10.4*	6.4	12.0*	1.2*	1.2*	1.2*	0.0*	13.6*	15.2*	11.2*	19.6*
MC1	6.0	8.4*	5.2	2.4	1.6*	2.0*	2.0*	4.0	10.8*	12.4*	9.2*	3.2

This table reports empirical size of testing the null hypothesis of zero abnormal return against two-tailed, lower-tailed, and upper-tailed alternative hypothesis, in samples of *1,000 firms*. Empirical size is calculated as the proportion of 250 samples that reject the null hypothesis at 5 % significance level. Abnormal return is calculated as the difference in holding period return between the event firm and its benchmark. We use *five benchmarks* (a reference portfolio matched by size and BE/ME (SZBM); a reference portfolio matched by size, BE/ME, and beta (SZBMBT); a reference portfolio consisting of ten firms, within the event firm's size and BE/ME matched portfolio, whose returns are most correlated with the event firm's MC10; a single firm matched by size and BE/ME (SZBM1); and a single firm, from the event firm's size and BE/ME matched portfolio, whose returns have the highest correlation with the event firm's MC1) and *four test statistics* (the conventional t-test (t), Fisher's sign test (sign), Johnson's skewness-adjusted t-test (Jt), and the bootstrapped Johnson's skewness-adjusted t-test (BJt)). It is indicated by * that the empirical size is significantly different from the 5 % significance level. The significance is judged against the critical values $0.05 \pm 1.96\sqrt{0.05(1-0.05)/250}$, where 0.05 is the theoretical size, 1.96 is the 97.5th percentile of the standard normal distribution, and 250 is the sample size

Table 14.4 Power of tests in samples of *1,000 firms*

		Induced abnormal return over the holding period (%)								
Test	Benchmark	−20	−15	−10	−5	0	5	10	15	20
Panel A: 1-year holding period										
t	SZBM	100.0	99.6	98.4	62.0	4.0	92.8	100.0	100.0	100.0
	SZBMBT	100.0	99.6	98.8	76.8	5.2	79.2	100.0	100.0	100.0
	MC10	100.0	99.6	98.4	73.6	5.6	77.6	100.0	100.0	100.0
	SZBM1	100.0	99.6	93.6	46.8	4.4	58.4	97.6	99.2	100.0
	MC1	100.0	99.6	97.2	50.8	3.6	58.0	96.8	99.6	100.0
Jt	SZBM	89.2	94.4	93.2	55.2	7.2	94.4	100.0	100.0	100.0
	SZBMBT	89.6	94.4	95.2	69.6	6.4	83.2	100.0	100.0	100.0
	MC10	91.2	95.6	95.2	66.4	6.4	80.0	100.0	100.0	100.0
	SZBM1	98.4	97.6	92.0	47.6	5.6	58.8	94.8	98.0	98.0
	MC1	98.0	98.4	95.6	50.0	5.2	59.2	95.6	98.0	98.0
BJt	SZBM	80.8	86.0	85.2	47.6	6.4	93.2	100.0	100.0	100.0
	SZBMBT	79.2	85.2	86.0	57.2	4.0	81.2	100.0	100.0	100.0
	MC10	81.6	86.4	87.2	56.4	6.4	78.8	100.0	100.0	100.0
	SZBM1	96.0	96.0	87.2	40.4	3.6	51.6	90.0	95.2	94.0
	MC1	95.6	95.6	88.8	44.4	3.2	51.6	91.6	95.6	95.6
Sign	SZBM	100.0	100.0	100.0	100.0	75.6	28.4	100.0	100.0	100.0
	SZBMBT	100.0	100.0	100.0	100.0	92.0	10.4	99.2	100.0	100.0
	MC10	100.0	100.0	100.0	100.0	85.6	17.2	100.0	100.0	100.0
	SZBM1	100.0	100.0	100.0	72.0	4.0	92.0	100.0	100.0	100.0
	MC1	100.0	100.0	100.0	93.6	9.6	90.4	100.0	100.0	100.0
Panel B: 3-year holding period										
t	SZBM	96.0	80.8	43.2	9.6	11.2	58.0	96.0	100.0	100.0
	SZBMBT	98.4	93.2	70.8	30.8	5.2	19.2	73.2	98.4	100.0
	MC10	98.4	92.4	70.0	26.8	4.8	19.2	72.4	98.8	100.0
	SZBM1	88.4	63.6	30.8	10.0	6.0	27.6	64.4	85.6	96.4
	MC1	92.4	74.0	36.4	10.4	6.8	22.4	64.0	91.2	97.6
Jt	SZBM	91.2	74.8	38.4	9.6	14.4	66.4	96.4	100.0	100.0
	SZBMBT	94.8	88.0	65.6	26.0	5.2	24.4	78.4	98.8	100.0
	MC10	94.0	87.6	62.8	24.4	6.8	23.2	76.8	99.2	100.0
	SZBM1	86.4	63.2	32.4	12.4	7.6	29.2	64.0	84.8	94.4
	MC1	90.4	72.4	36.0	12.0	8.4	24.4	64.8	90.4	97.6
BJt	SZBM	84.8	66.0	32.4	7.6	12.8	62.4	96.0	100.0	100.0
	SZBMBT	88.8	82.0	58.4	21.6	5.6	21.6	74.8	98.4	100.0
	MC10	90.4	79.6	54.8	19.6	5.6	21.6	73.6	98.0	100.0
	SZBM1	81.6	56.4	27.6	8.0	5.2	24.4	56.4	79.6	88.8
	MC1	86.0	65.6	29.6	8.8	6.4	20.4	55.2	86.8	94.4
Sign	SZBM	100.0	100.0	100.0	100.0	99.6	63.6	6.0	27.2	88.8
	SZBMBT	100.0	100.0	100.0	100.0	100.0	89.6	27.6	6.8	64.4
	MC10	100.0	100.0	100.0	100.0	100.0	78.4	15.6	12.4	78.8
	SZBM1	100.0	94.8	56.4	14.0	9.2	54.8	94.0	100.0	100.0
	MC1	100.0	100.0	95.2	50.0	6.4	37.2	86.4	100.0	100.0

(*continued*)

Table 14.4 (continued)

		Induced abnormal return over the holding period (%)								
Test	Benchmark	−20	−15	−10	−5	0	5	10	15	20
Panel C: 5-year holding period										
t	SZBM	58.0	28.4	9.6	1.6	17.6	40.8	79.2	97.6	99.2
	SZBMBT	84.8	63.2	37.6	14.8	3.6	8.4	32.8	66.0	92.0
	MC10	80.4	61.2	32.8	11.6	2.0	10.4	32.4	69.6	92.0
	SZBM1	38.0	18.4	7.2	4.0	8.0	21.6	41.6	64.8	82.4
	MC1	50.4	23.6	10.4	4.0	6.0	17.2	38.0	61.2	81.2
Jt	SZBM	44.4	23.6	7.2	5.6	20.4	52.0	85.2	98.8	99.6
	SZBMBT	72.0	51.6	27.6	11.2	4.4	14.8	40.0	73.2	94.8
	MC10	71.6	51.2	27.6	7.6	4.4	15.6	38.0	73.6	96.4
	SZBM1	38.0	20.0	8.8	6.0	10.4	23.6	42.4	65.2	82.0
	MC1	48.8	24.4	11.2	6.0	8.4	18.8	38.4	60.4	79.6
BJt	SZBM	35.2	19.6	5.2	2.8	19.6	48.0	82.8	98.0	99.6
	SZBMBT	62.8	43.2	21.6	8.4	2.8	12.8	36.8	70.0	94.0
	MC10	60.0	42.4	20.0	6.4	2.8	14.4	36.0	72.0	94.8
	SZBM1	30.0	16.0	5.6	4.0	6.4	16.4	33.6	54.4	70.4
	MC1	38.4	17.6	9.6	3.2	5.2	15.2	30.8	50.4	70.8
Sign	SZBM	100.0	100.0	100.0	100.0	100.0	95.6	72.0	22.8	3.2
	SZBMBT	100.0	100.0	100.0	100.0	100.0	99.6	93.2	61.6	19.2
	MC10	100.0	100.0	100.0	100.0	100.0	98.4	81.6	40.8	7.6
	SZBM1	91.2	63.2	20.8	4.0	12.0	48.8	86.0	97.2	100.0
	MC1	99.2	92.8	59.2	22.8	2.4	20.4	67.2	93.6	99.2

This table reports empirical power of testing the null hypothesis of zero abnormal return against the two-sided alternative hypothesis, in samples of *1,000 firms*. Empirical power is calculated as the proportion of 250 samples that reject the null hypothesis at 5 % significance level. Abnormal return is calculated as the difference in holding period return between the event firm and its benchmark. We use *five benchmarks* (a reference portfolio matched by size and BE/ME (SZBM); a reference portfolio matched by size, BE/ME, and beta (SZBMBT); a reference portfolio consisting of ten firms, within the event firm's size and BE/ME matched portfolio, whose returns are most correlated with the event firm's MC10; a single firm matched by size and BE/ME (SZBM1); and a single firm, from the event firm's size and BE/ME matched portfolio, whose returns have the highest correlation with the event firm's MC1) and *four test statistics* (the conventional *t*-test (t), Johnson's skewness-adjusted *t*-test (Jt), the bootstrapped Johnson's skewness-adjusted *t*-test (BJt), and Fisher's sign test (sign)). We study power at nine levels of induced abnormal return, ranging from −20 % to 20 % at an increment of 5 %

14.4.3 Simulation Results for the Calendar-Time Portfolio Approach

Table 14.5 reports the rejection frequency of the calendar-time portfolio approach in testing the null hypothesis that the intercept is zero in the regression of monthly calendar-time portfolio returns, against the two-sided alternative hypothesis. Rejection frequency is measured as the proportion of the total 250 samples that reject the null hypothesis. We compute rejection frequencies at nine *nominal* levels of induced abnormal returns, ranging from −20 % to 20 % at an increment of 5 %. Since monthly returns of the calendar-time portfolio are used in fitting the model, to examine the power of testing the intercept, we need to induce abnormal returns by

Table 14.5 Rejection frequency of calendar-time portfolio approach in samples of *1,000 firms*

Panel A: 1-year holding period										
		Average effective induced holding period return (%)								
		−20.4	−15.7	−10.7	−5.5	0	5.7	11.7	17.9	24.4
Three factors	OLS	100.0	100.0	99.2	53.2	2.4	78.8	100.0	100.0	100.0
	WLS	100.0	100.0	99.6	74.4	2.0*	82.8	100.0	100.0	100.0
Four factors	OLS	100.0	99.2	90.8	18.0	28.0*	97.6	100.0	100.0	100.0
	WLS	100.0	99.6	93.2	20.8	25.2*	98.8	100.0	100.0	100.0
Panel B: 3-year holding period										
		Average effective induced holding period return (%)								
		−25.2	−19.3	−13.2	−6.8	0	7.1	14.5	22.3	30.4
Three factors	OLS	98.0	86.8	38.0	3.6	2.4	32.0	84.8	99.6	99.6
	WLS	100.0	97.2	65.2	10.0	1.2*	36.0	91.6	100.0	100.0
Four factors	OLS	69.2	22.0	1.6	6.4	55.2*	94.0	99.6	100.0	100.0
	WLS	92.0	38.0	4.0	10.4	75.6*	99.6	100.0	100.0	100.0
Panel C: 5-year holding period										
		Average effective induced holding period return (%)								
		−31.1	−23.9	−16.3	−8.3	0	8.7	17.9	27.4	37.5
Three factors	OLS	64.8	31.2	10.0	0.8	4.0	27.6	62.4	90.8	99.6
	WLS	94.4	58.4	14.8	0.4	4.0	36.0	81.2	99.2	100.0
Four factors	OLS	12.4	1.6	5.2	32.8	70.8*	89.2	98.8	100.0	100.0
	WLS	14.0	1.2	14.8	62.4	94.0*	100.0	100.0	100.0	100.0

This table reports rejection frequency in testing the null hypothesis that the intercept in the regression of monthly calendar-time portfolio returns is zero, in samples of *1,000 firms*. Both the Fama-French three-factor model and the four-factor model are used in the regression. Model parameters are estimated with both OLS and WLS estimation techniques. Rejection frequency is equal to the proportion of 250 samples that reject the null hypothesis at 5 % significance level. We measure rejection frequency at nine levels of induced abnormal returns. We induce abnormal returns by adding an extra amount to monthly returns of every event firm before forming the calendar-time portfolios. The *effective* induced holding period return of an event firm is equal to the difference in the firm's holding period return between before and after adding the monthly extra amount. The *average* effective induced holding period return is computed over all event firms in the 250 samples

adding an extra amount to actual monthly returns of every event firm before forming the calendar-time portfolios. For example, in order to induce the −20 % nominal level of abnormal holding period return, we add the extra amount of −1.67 % (=−20 %/12) to an event firm's 12 monthly returns for a 1-year horizon, or add the abnormal amount of −0.56 % (= −20 %/36) to the firm's 24 monthly returns for a 3-year horizon, or the abnormal amount of −0.33 % (=−20 %/60) to the firm's 60 monthly returns for a 5-year horizon.

Note that the *nominal* induced holding period return is different from the *effective* induced abnormal holding period return, because adding the abnormal amount each month does not guarantee that an event firm's holding period return will be increased or decreased by the exact nominal level. We measure the *effective* induced holding period return of an event firm as the difference in the firm's holding period return between before and after adding the monthly abnormal amount. The *average* effective induced holding period return is computed over all event firms in the 250 samples. The average induced holding period return allows us to compare power of the buy-and-hold benchmark approach with that of the calendar-time portfolio approach at the scale of holding period return.

We first examine empirical size of the calendar-time portfolio approach, which is equal to the rejection frequency when no abnormal return is induced. In Table 14.5, the empirical size is in the column with zero induced return. It is very surprising that when the four-factor model is used, the test has excessively high rejection frequency at 3-year and 5-year horizons. The rejection frequency, for example, is 94.0 % at the 5-year horizon with the WLS estimation! In contrast, when the Fama-French three-factor model is used, the empirical sizes are not significantly different from the 5 % significance level. The evidence strongly suggests that the three-factor model is preferred for the calendar-time portfolio approach, whereas the four-factor model suffers from overfitting and should not be used.

Table 14.5 shows that, for a sample of 1,000 firms, the power of this approach remains high as event horizon increases. WLS estimation does improve the power of the procedure over the OLS, and the extent of improvement becomes greater as holding period gets longer. By comparing Tables 14.4 and 14.5, we find that the power of the Fama-French calendar-time approach implemented with WLS technique (i.e., FF, WLS) has almost the same power as the buy-and-hold benchmark approach implemented with the most correlated single firm and the sign test (i.e., MC1, sign), at the 1-year horizon, but slightly less at the 3- and 5-year horizons.

14.5 Conclusion

Comparing the simulation results in Sect. 14.4 with those in Ang and Zhang (2004), we find that sample size has a significant impact on the performance of tests in long-horizon event studies. With a sample size of 1,000, a few tests perform reasonably well, including the Fama-French calendar-time approach implemented with WLS technique and the buy-and-hold benchmark approach implemented with the most correlated single firm (MC1) and the sign test. In particular, they have reasonably high power even for the long 5-year holding period. On the contrary, with a sample size of 200, Ang and Zhang (2004) find that the power of most well-specified tests is very low for the 5-year horizon, only in the range of 10–20 % against a high level of induced abnormal returns, while the combination of the most correlated single firm and the sign test stands out with a power of 41.2 %. Thus, the most correlated single firm benchmark dominates for most practical sample sizes, and in addition, the simplicity of the sign test is appealing.

The findings have important implications for future research. For long-horizon event studies with a large sample, it is likely to be more fruitful to spend efforts on understanding the characteristics of the sample firms, than on implementing various sophisticated testing procedures. The simulation results here show that the commonly used tests following both the Fama-French calendar-time approach and the buy-and-hold benchmark approach perform reasonably well. In a recent paper, Butler and Wan (2010) reexamine the long-run underperformance of bond-issuing firms and find that straight debt and convertible debt issuers appear to have systematically better liquidity than benchmark firms, and controlling for liquidity by having an additional matching criterion eliminates the underperformance. This resonates well with Barber and Lyon's (1997) suggestion that "as future research in financial economics discovers additional variables that explain the cross-sectional variation in common stock returns, it will also be important to consider these additional variables when matching sample firms to control firms" (pp. 370–71). One reason why the benchmark with a single most correlated firm performs well in our simulations may be that returns of highly correlated firms are likely to move in tandem in response to changes in risk factors that are well known, such as the market, size, and book-to-market ratio, but also changes in other factors, such as industry, liquidity, momentum, and seasonality.

On the other hand, for long-horizon event studies with a small sample, it may be necessary to use a wide range of tests and interpret their outcome with care. This prompts researchers to continue searching for better test statistics. For example, Kolari and Pynnonen (2010) find that even relatively low cross-correlation among abnormal returns in a short event window causes serious over-rejection of the null hypothesis. They propose both cross-correlation and volatility-adjusted as well as cross-correlation-adjusted scaled test statistics and demonstrate that these statistics perform well in samples of 50 firms. It is an open and interesting question whether these statistics have high power in long-horizon event studies with a small sample.

Appendix

This appendix includes the details on the benchmarks and the test statistics that are used in our simulation studies. We use five benchmarks. The *first benchmark* is a reference portfolio constructed on the basis of firm size and BE/ME. We follow Lyon et al. (1999) to form 70 reference portfolios at the end of June in each year from 1979 to 1997. At the end of June of year t, we calculate the size of every qualified firm as price per share multiplied by shares outstanding. We sort all NYSE firms by firm size into ten portfolios, each having the same number of firms, and then place all AMEX/NASDAQ firms into the ten portfolios based on firm size. Since a majority of NASDAQ firms are small, approximately 50 % of all firms fall in the smallest size decile. To obtain portfolios with the same number of firms, we further partition the smallest size decile into five subportfolios by firm size without regard to listing exchange. We now have 14 size portfolios. Next, we calculate each qualified firm's BE/ME as the ratio of the book equity value (COMPUSTAT data item 60) of the firm's fiscal year ending in year $t-1$ to its market equity value at the end of December of year $t-1$. We then divide each of the 14 portfolios into five subportfolios by BE/ME and conclude the procedure with 70 reference portfolios on the basis of size and BE/ME.

The size and BE/ME matched reference portfolio of an event firm is taken to be the one of the 70 reference portfolios constructed at the month of June prior to the event month that matches the event firm in size and BE/ME. The return on a size and BE/ME matched reference portfolio over τ months is calculated as

$$BR_i^{SZBM} = \prod_{t=0}^{\tau-1}\left[1 + \frac{\sum_{j=1}^{n_t} r_{jt}}{n_t}\right] - 1, \tag{14.5}$$

where month $t = 0$ is the event month, n_t is the number of firms in month t, and r_{jt} is the monthly return of firm j in month t. We use the label “SZBM” for the benchmark that is based on firm size and BE/ME.

The *second benchmark* is a reference portfolio constructed on the basis of firm size, BE/ME, and market beta. The Fama-French three-factor model suggests that expected stock returns are related to three factors: a market factor, a size-related factor, and a BE/ME-related factor. Reference portfolios constructed on the basis of size and BE/ME account for the systematic portion of expected stock returns due to the size and BE/ME factors, but not the portion due to the market factor. Our second benchmark is based on firm size, BE/ME, and market beta to take all three factors into account.

To build a three-factor reference portfolio for a given event firm, we first construct the 70 size and BE/ME reference portfolios as above and identify the one that matches the event firm. Next, we pick firms within the matched portfolio that have returns in CRSP monthly returns database for all 24 months prior to the event month and compute their market beta by regressing the 24 monthly returns on the value-weighted CRSP return index. Lastly, we divide these firms that have market beta into three portfolios by their rankings in beta and pick the one that matches the event firm in beta as the three-factor reference portfolio. The return on a three-factor portfolio over τ months is calculated as

$$BR_i^{SZBMBT} = \prod_{t=0}^{\tau-1}\left[1 + \frac{\sum_{j=1}^{l_t} r_{jt}}{n_t}\right] - 1, \tag{14.6}$$

where month $t = 0$ is the event month, n_t is the number of firms in month t, and r_{jt} is the monthly return of firm j in month t. We use the label “SZBMBT” to indicate that the benchmark is based on firm size, BE/ME, and market beta.

The *third benchmark* is a reference portfolio constructed on the basis of firm size, BE/ME, and pre-event correlation coefficient. The rational for using pre-event correlation coefficient as an additional dimension is that returns of highly correlated firms are likely to move in tandem in response to not only changes in “global” risk factors, such as the market factor, the size factor, and the BE/ME factor in the Fama-French model, but also changes in other “local” factors, such as the industry factor, the seasonal factor, liquidity factor, and the momentum factor. Over a long time period following an event, both global and local factors experience changes that affect stock returns. It is reasonable to expect more correlated stocks would be affected by these factors similarly and should have resulting stock return patterns that are closer to each other. Therefore, returns of a reference portfolio on the basis of pre-event size, BE/ME, and pre-event correlation coefficient are likely to be better estimate of the *status quo* (i.e., what if there was no event) return of an event firm.

To build a reference portfolio on the basis of size, BE/ME, and pre-event correlation coefficient, we first construct the same 70 size and BE/ME reference portfolios as above and identify the combination that matches the event firm. Next, we pick firms within the matched size and BE/ME reference portfolio that have returns in CRSP monthly returns database for all 24 months prior to the event month and compute their correlation coefficients with the event firm over the pre-event 24 months. Lastly, we choose the ten firms that have the highest pre-event correlation coefficient with the event firm to form the reference portfolio. Return of the portfolio over τ months is calculated as

$$BR_i^{MC10} = \sum_{j=1}^{10} \frac{\prod_{t=0}^{\tau-1}\left(1 + r_{jt}\right) - 1}{10}, \tag{14.7}$$

where month $t = 0$ is the event month and r_{jt} is the monthly return of firm j in month t. We use the label "MC10" to indicate that the benchmark consists of the most correlated ten firms. The benchmark return is the return of investing equally in the ten most correlated firms over the τ months beginning with the event month. The benchmark is to be considered as a hybrid between the reference portfolio discussed above and the matching firm approach shown below.

The *fourth benchmark* is a single firm matched to the event firm in size and BE/ME. Barber and Lyon (1997) report that using a size and BE/ME matched firm as benchmark gives measurements of long-term abnormal return that is free of the new listing bias, the rebalancing bias, and the skewness bias documented in Kothari and Warner (1997) and Barber and Lyon (1997). To select the size and BE/ME matched firm, we first identify all firms that have a market equity value between 70 % and 130 % of that of the event firm and then choose the firm with BE/ME closest to that of the event firm. The buy-and-hold return of the matched firm is computed as in Eq. 14.2. We use the label "SZBM1" to represent the single size and BE/ME matched firm.

The *fifth and last benchmark* is a single firm that has the highest pre-event correlation coefficient with the event firm. Specifically, to select the firm, we first construct the 70 size and BE/ME reference portfolios and identify the one that matches the event firm. Next, we pick firms within the matched size and BE/ME reference portfolio that have returns in CRSP monthly returns database for all 24 months prior to the event month and compute their correlation coefficients with the event firm over the pre-event 24 months. We choose the firm with the highest pre-event correlation coefficient with the event firm as the benchmark. The buy-and-hold return of the most correlated firm is computed as in Eq. 14.2. We use the label "MC1" to represent the most correlated single firm.

We apply four test statistics to test the null hypothesis of no abnormal returns: (a) Student's t-test, (b) Fisher's sign test, (c) Johnson's skewness-adjusted t-test, and (d) bootstrapped Johnson's t-test.

(a) Student's t-test

Given the long-term buy-and-hold abnormal returns for a sample of n event firms, we compute Student's t-statistic as follows:

$$t = \frac{\overline{AR}}{s(AR)/\sqrt{n}}, \tag{14.8}$$

where $\overline{AR}$ is the sample mean and $s(AR)$ the sample standard deviation of the given sample of abnormal returns. The Student's t-statistic tests the null hypothesis that the population mean of long-term buy-and-hold abnormal returns is equal to zero. The usual assumption for applying the Student's t-statistic is that abnormal returns are mutually independent and follow the same normal distribution.

(b) Fisher's sign test

To test the null hypothesis that the population median of long-term buy-and-hold abnormal returns is zero, we compute Fisher's sign test statistic as follows:

$$B = \sum_{i=1}^{n} \mathrm{I}(AR_i > 0), \tag{14.9}$$

where $\mathrm{I}(AR_i > 0)$ equals 1 if the abnormal return on the ith firm is greater than zero and 0 otherwise. At the chosen significance level of α, the null hypothesis is rejected in favor of the alternative of nonzero median if $B \geq b(\alpha/2, n, 0.5)$ or $B < [n - b(\alpha/2, n, 0.5)]$, or in favor of positive median if $B \geq b(\alpha, n, 0.5)$, or in favor of negative median if $B < [n - b(\alpha, n, 0.5)]$. The constant $b(\alpha, n, 0.5)$ is the upper α percentile point of the binomial distribution with sample size n and success probability of 0.5. The usual assumption for applying the sign test is that abnormal returns are mutually independent and follow the same continuous distribution. Note that application of the sign test does not require the population distribution to be symmetric. When the population distribution is symmetric, the population mean equals the population median, and the sign test then indicates the significance of the population mean (see Hollander and Wolfe 2000, Chap. 3).

(c) Johnson's skewness-adjusted t-test

Johnson (1978) developed the following skewness-adjusted t-test to correct the misspecification of Student's t-test caused by the skewness of the population distribution. Johnson's test statistic is computed as follows:

$$J = t + \frac{1}{3\sqrt{n}} t^2 \gamma + \frac{1}{6\sqrt{n}} \gamma, \tag{14.10}$$

where t is Student's t-statistic given in Eq. 14.8 and γ is an estimate of the coefficient of skewness given by $\gamma = \sum_{i=1}^{n} \left(AR_i - \overline{AR}\right)^3 / s(AR)^3 n$. Johnson's

t-test is applied to test the null hypothesis of zero mean under the assumption that abnormal returns are mutually independent and follow the same continuous distribution. At the chosen significance level of α, the null hypothesis is rejected in favor of the alternative of nonzero mean if $J > t(\alpha/2, \upsilon)$ or $J < -t(\alpha/2, \upsilon)$, or in favor of positive mean if $J > t(\alpha, \upsilon)$, or in favor of negative mean if $J < -t(\alpha, \upsilon)$. The constant $t(\alpha, \upsilon)$ is the upper α percentile point of the Student's t distribution with the degrees of freedom $\upsilon = n - 1$.

(d) Bootstrapped Johnson's skewness-adjusted t-test
Sutton (1992) proposes to apply Johnson's t-test with a computer-intensive bootstrap resampling technique when the population skewness is severe and the sample size is small. He demonstrates it by an extensive Monte Carlo study that the bootstrapped Johnson's t-test reduces both type I and type II errors compared to Johnson's t-test. Lyon et al. (1999) advocate the bootstrapped Johnson's t-test in that long-term buy-and-hold abnormal returns are highly skewed when buy-and-hold reference portfolios are used as benchmarks. They report that the bootstrapped Johnson's t-test is well specified and has considerable power in testing abnormal returns at the 1-year horizon. In this paper, we document its power at 3- and 5-year horizons.

We apply the bootstrapped Johnson's t-test as follows. From the given sample of n event firms, we draw m firms randomly with replacement counted as one resample until we have 250 resamples. We calculate Johnson's test statistic as in Eq. 14.10 for each resample and end up with 250 J values, labeled as $J_1, \cdots, J_{250}$. Let J_0 denotes the J value of the original sample. To test the null hypothesis of zero mean at the significance level of α, we first determine two critical values, C_1 and C_2, such that the percentage of J values less than C_1 equals $\alpha/2$ and the percentage of J values greater than c_2 equals $\alpha/2$, and then reject the null hypothesis if $J_0 < C_1$ or $J_0 > C_2$. We follow Lyon et al. (1999) to apply the bootstrapped Johnson's t-test with $m = 50$.[9]

References

Ang, J. S., & Zhang, S. (2004). An evaluation of testing procedures for long horizon event studies. *Review of Quantitative Finance and Accounting, 23*, 251–274.

Barber, B. M., & Lyon, J. D. (1997). Detecting long-run abnormal stock returns: The empirical power and specification of test statistics. *Journal of Financial Economics, 43*, 341–372.

Brav, A., & Gompers, P. A. (1997). Myth or reality? The long-run underperformance of initial public offerings: Evidence from venture and nonventure capital-backed companies. *Journal of Finance, 52*, 1791–1821.

Brav, A., Geczy, C., & Gompers, P. A. (2000). Is the abnormal return following equity issuances anomalous? *Journal of Financial Economics, 56*, 209–249.

[9]Noreen (1989, ► Chap. 4) cautions that bootstrap hypothesis tests can be unreliable and that extensive research is necessary to determine which one of many possible specifications can be trusted in a particular hypothesis testing situation. We also apply the bootstrapped Johnson's t-test with $m = 100, 200$. We find no significant difference in the test's performance.

Brown, S. J., & Warner, J. B. (1980). Measuring security price performance. *Journal of Financial Economics, 8*, 205–258.
Brown, S. J., & Warner, J. B. (1985). Using daily stock returns: The case of event studies. *Journal of Financial Economics, 14*, 3–31.
Butler, A. W., & Wan, H. (2010). Stock market liquidity and the long-run stock performance of debt issuers. *Review of Financial Studies, 23*, 3966–3995.
Carhart, M. M. (1997). On persistence in mutual fund performance. *Journal of Finance, 52*, 57–82.
Cowan, A. R., & Sergeant, A. M. A. (2001). Interacting biases, non-normal return distributions and the performance of tests for long-horizon event studies. *Journal of Banking and Finance, 25*, 741–765.
Dahlquist, M., & de Jong, F. (2008). Pseudo market timing: A reappraisal. *Journal of Financial and Quantitative Analysis, 43*, 547–580.
Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their application*. Cambridge, UK: Cambridge University Press.
Eckbo, B. E., Masulis, R. W., & Norli, O. (2000). Seasoned public offerings: Resolutions of the 'new issues puzzle'. *Journal of Financial Economics, 56*, 251–291.
Eckbo, B. E., & Norli, O. (2005). Liquidity risk, leverage, and long-run IPO returns. *Journal of Corporate Finance, 11*, 1–35.
Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. New York: Chapman & Hall.
Fama, E. F. (1998). Market efficiency, long-term returns and behavioral finance. *Journal of Financial Economics, 49*, 283–306.
Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance, 47*, 427–465.
Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*, 3–56.
Hollander, M., & Wolfe, D. A. (1999). *Nonparametric statistical methods*. New York: Wiley.
Jegadeesh, N. (2000). Long-term performance of seasoned equity offerings: Benchmark errors and biases in expectations. *Financial Management, 29*, 5–30.
Jegadeesh, N., & Karceski, J. (2009). Long-run performance evaluation: Correlation and heteroskedasticity-consistent tests. *Journal of Empirical Finance, 16*, 101–111.
Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance, 48*, 65–91.
Jegadeesh, N., & Titman, S. (2001). Profitability of momentum strategies: An evaluation of alternative explanations. *Journal of Finance, 56*, 699–720.
Johnson, N. J. (1978). Modified *t* tests and confidence intervals for asymmetrical populations. *Journal of the American Statistical Association, 73*, 536–544.
Kolari, J. W., & Pynnonen, S. (2010). Event study testing with cross-sectional correlation of abnormal returns. *Review of Financial Studies, 23*, 3996–4025.
Kothari, S. P., & Warner, J. B. (1997). Measuring long-horizon security price performance. *Journal of Financial Economics, 43*, 301–340.
Kothari, S. P., & Warner, J. B. (2007). Econometrics of event studies. In B. E. Eckbo (Ed.), *Handbooks of corporate finance: Empirical corporate finance*. Amsterdam: Elsevier/North-Holland.
Loughran, T., & Ritter, J. R. (1995). The new issues puzzle. *Journal of Finance, 50*, 23–52.
Loughran, T., & Ritter, J. R. (2000). Uniformly least powerful tests of the market efficiency. *Journal of Financial Economics, 55*, 361–389.
Lyon, J. D., Barber, B. M., & Tsai, C.-L. (1999). Improved methods for tests of long-run abnormal stock returns. *Journal of Finance, 54*, 165–201.
Matsumoto, M., & Nishimura, T. (1998). Mersenne twister: A 623-dimensionally equidistributed uniform pseudorandom number generator. *ACM Transactions on Modeling and Computer Simulation, 8*, 3–30.
Mitchell, M. L., & Stafford, E. (2000). Managerial decisions and long-term stock price performance. *Journal of Business, 73*, 287–329.

Noreen, E. W. (1989). *Computer intensive methods for testing hypotheses: An introduction*. New York: Wiley.

Pearson, E. S., & Please, N. W. (1975). Relation between the shape of population distribution and the robustness of four simple test statistics. *Biometrika, 62*, 223–241.

Schultz, P. (2003). Pseudo market timing and the long-run underperformance of IPOs. *Journal of Finance, 58*, 483–517.

Sutton, C. D. (1993). Computer-intensive methods for tests about the mean of an asymmetrical distribution. *Journal of the American Statistical Association, 88*, 802–808.

Viswanathan, S., & Wei, B. (2008). Endogenous events and long-run returns. *Review of Financial Studies, 21*, 855–888.

The Effect of Unexpected Volatility Shocks on Intertemporal Risk-Return Relation

15

Kiseok Nam, Joshua Krausz, and Augustine C. Arize

Contents

Abstract

We suggest that an unexpected volatility shock is an important risk factor to induce the intertemporal relation, and the conflicting findings on the relation could be attributable to an omitting variable bias resulting from ignoring the effect of an unexpected volatility shock on the relation. With the effect of an unexpected volatility shock incorporated in estimation, we find a strong positive

K. Nam (✉) • J. Krausz
Yeshiva University, New York, NY, USA
e-mail: knam@yu.edu; krausz@yu.edu

A.C. Arize
Texas A & M University-Commerce, Commerce, TX, USA
e-mail: chuck.arize@tamuc.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_15,

intertemporal relation for the US monthly excess returns for 1926:12–2008:12. We reexamine the relation for the sample period studied by Glosten, Jagannathan, and Runkle (Journal of Finance 48, 1779–1801, 1993) and find that their sample period is indeed characterized by a positive (negative) relation under a positive (negative) volatility shock with the effect of a volatility shock incorporated in estimation. We also find a significant link between the asymmetric mean reversion and the intertemporal relation in that the quicker reversion of negative returns is attributed to the negative intertemporal relation under a prior negative return shock.

For estimations we employ the ANST-GARCH model that is capable of capturing the asymmetric volatility effect of a positive and negative return shock. The key feature of the model is the regime-shift mechanism that allows a smooth, flexible transition of the conditional volatility between different states of volatility persistence. The regime-switching mechanism is governed by a logistic transition function that changes values depending on the level of the previous return shock. With a negative (positive) return shock, the conditional variance process is described as a high (low)-persistence-in-volatility regime. The ANST-GARCH model describes the heteroskedastic return dynamics more accurately and generates better volatility forecasts.

Keywords

Intertemporal risk-return relation • Unexpected volatility shocks • Time-varying rational expectation hypothesis • Stock market overreaction • Expected market risk premium • Volatility feedback effect • Asymmetric mean reversion • Asymmetric volatility response • Time-varying volatility • Volatility regime switching • ANST-GARCH model

15.1 Introduction

The trade-off between risk and return is a core tenet in financial economics. In particular, the intertemporal risk-return relation is a key element to explain the predictable variation of expected asset returns.[1] Despite its importance in asset pricing, there has been a long-standing debate on the empirical sign of the intertemporal relation, with findings that are mixed and inconclusive.

Criticisms of the mixed results often refer to a lack of conditional information. If the predetermined conditional information set does not contain an important variable that affects the risk-return trade-off, the econometric modeling of market expectations suffers from the model misspecification problem and leads to a wrong conclusion on the empirical nature of the intertemporal risk-return relation.

[1]Fama and French (1989) argue that systematic patterns in the predictable variations of expected returns are consistent with the intertemporal asset pricing model by Lucas (1978) and Breeden (1979) and the consumption smoothing idea by Modigliani and Brumberg (1955) and Friedman (1957). Ferson and Harvey (1991) and Evans (1994) also document the relative importance of the time-varying risk premia to the conditional betas to explain predictable variations in expected returns.

In this chapter, we suggest that an unexpected volatility shock is an important risk factor to induce the intertemporal risk-return trade-off, and the conflicting findings on the relation could be attributable to an omitting variable bias resulting from ignoring the effect of an unexpected volatility shock on the relation. Incorporating the effect of an unexpected volatility shock in estimation of the relation, we find a strong *positive* intertemporal relation between the expected market risk premium and the predictable volatility for the US monthly excess returns of market indices for the period of 1926:01–2008:12.

Conventional belief about the intertemporal relation is that a positive risk-return relation is consistent with the time-varying rational expectation hypothesis in the sense that a substantial amount of predictable variations of the expected risk premium is induced by the risk-averse investors' revision of their expectations in responding to changing volatility. For example, Pindyck (1984) empirically shows that much of the decline in stock prices during the 1970s in the US stock market is attributable to the upward shift in risk premium arising from high stock market volatility.[2] Studies that support a positive relation include French et al. (1987), Fama and French (1988), Ball and Kothari (1989), Turner et al. (1989), Harvey (1989), Cecchetti et al. (1990), Haugen et al. (1991), Campbell and Hentschel (1992), Scruggs (1998), Kim et al. (2001), Ghysel et al. (2005), Ludvigson and Ng (2007), Pastor et al. (2008), Bali (2008), Darrat et al. (2011), and Nyberga (2012).

Although a positive intertemporal relation is consistent with Merton's (1980) dynamic CAPM, there is another side to the argument that the equilibrium asset pricing does not necessarily imply a positive relation. Abel (1988) suggests that a positive risk-return relation is consistent with the general equilibrium model only when the coefficient of relative risk aversion is less than one. Barky (1989) suggests that the directional effect of an increase in riskiness on stock prices depends on the curvature of the utility function. Showing evidence of a strong negative relation for their sample period, Glosten et al. (hereinafter GJR) (1993) suggest that both positive and negative intertemporal relations are consistent with the equilibrium asset pricing theory.[3] Among others, Campbell (1987), Pagan and Hong (1989), Breen et al. (1989), Nelson (1991), Backus and Gregory (1993), Harvey (2001), and Ang et al. (2006) support a negative intertemporal relation.[4]

[2]He suggested that a substantial portion of time variation in the expected risk premium is associated with time-varying risk factors in investment opportunities.

[3]They argue that investors may not require a large premium for bearing risk, but rather may reduce the risk premium when they perceive exceptionally optimistic expectations on the future performance of stock prices.

[4]Brandt and Kang (2004) find that the conditional mean and volatility are negatively correlated contemporaneously but positively correlated unconditionally due to the positive lead-lag relation between the two moments of stock returns. Poterba and Summers (1986) suggested that, due to the low level of volatility persistence, the volatility effect on the expected risk premium dissipates so quickly that it cannot have a major effect on stock price movements. Note that there are some studies that report weak evidence of the intertemporal relation. See Baillie and DeGennaro (1990), Whitelaw (1994, 2000), Boudoukh et al. (1997), Yu and Yuan (2011), and Müller et al. (2011).

A price shock generates two sources of forecasting errors: a return forecasting error and a volatility forecasting error. Unlike a return forecasting error, the unanticipated volatility shock has not been paid much attention by the literature. We refer the volatility forecasting error to as an unexpected volatility shock. An unexpected volatility shock is indeed an important risk factor to affect investors' pricing behavior in that rational risk-averse investors revise their expectations in responding not only to an underlying volatility but also to an unexpected volatility shock. This implies that the intertemporal behavior of the expected market risk premium is driven by both the predictable volatility and the unexpected volatility change.

While the underlying return volatility has been widely examined by many studies, the impact of an unexpected volatility shock has not been given much attention by the literature. Almost all of the previous empirical studies on this topic completely ignore the effect of an unexpected volatility shock on the relation, so that their empirical results reflect only a partial intertemporal risk-return relation. Thus, considering an unexpected volatility shock as the conditional information in estimation, we reexamine the nature of the intertemporal relation.

Especially, we examine an asymmetrical effect of a positive and negative unexpected volatility shock on the relation. We conjecture that a higher volatility level than predicted would increase the expected risk premium and induce a stronger positive intertemporal relation. We define a positive (negative) unexpected volatility shock as the case where actual market volatility is higher (lower) than expected. Our estimation results show that a positive unexpected volatility shock (the case in which the actual volatility is higher than expected) causes a stronger positive intertemporal relation than does a negative unexpected volatility shock (the case in which the actual volatility is lower than expected).

The effect of an unexpected volatility shock considered in this chapter is different from the volatility feedback effect proposed by Campbell and Hentschel (1992). While the volatility feedback effect focuses on the *contemporaneous* effect of concurrent volatility shocks on the expected returns, our volatility shock effect implies the consequence of a prior unexpected volatility shock on the intertemporal relation. However, the estimation of their contemporaneous volatility feedback effect is subject to the endogeneity problem, as it is theoretically impossible to obtain the concurrent volatility forecasting error due to an unavailability of the current volatility information.[5]

[5]To avoid the endogeneity problem, French et al. (1987) examined the volatility feedback effect using ex post unexpected volatility changes. They found a strong negative relation between unexpected returns and ex post unexpected volatility changes and interpreted it as evidence supporting a positive intertemporal relation.

15.2 Theoretical Models

Merton (1973) suggests the intertemporal risk-return relation as a function of stock market volatility, which can be specified in the following general form:

$$E\left(R_{mt} - rf_t|\Omega_{t-1}\right) = f\left(\sigma^p_{mt}\right),\ p = 1,2, \tag{15.1}$$

Where $E(\cdot)$ is the expectation operator, $E(R_{mt}-rf_t|\Omega_{t-1})$ is the time-varying expected market risk premium conditional on the information set Ω_{t-1}, R_{mt} is the return on a stock market index portfolio, and rf_t is the risk-free interest rate. The market volatility is represented by either σ_{mt} or σ^2_{mt}, where $\sigma^2_{mt} = E_{t-1}[(R_{mt} - rf_t) - \mathrm{E}_{t-1}(R_{mt} - rf_t)]^2$. Although $\frac{\partial E(R_{mt}-rf_t|\Omega_{t-1})}{\partial \sigma^2_{mt}} > 0$ is consistent with the equilibrium asset pricing theory, there has been a long-standing controversy in the sign of the relation.

The intertemporal relation is driven not only by predictable volatility but also by an unexpected volatility change. We thus consider the intertemporal relation in the following form:

$$E(R_{mt} - rf_t|\Omega_{t-1}) = f_1\left(\hat{\sigma}^2_{mt}\right) + f_2\left(\varepsilon^2_{mt-1} - \hat{\sigma}^2_{mt-1}\right), \tag{15.2}$$

where ε^2_{mt-1} and $\hat{\sigma}^2_{mt-1}$ are the actual realized volatility series and the predicted volatility series, respectively, such that $\varepsilon^2_{mt-1} - \hat{\sigma}^2_{mt-1}$ represents a prior unexpected volatility shock. Many studies employ the GARCH models to conditionally estimate the predictable volatility series $\hat{\sigma}^2_{mt}$. Equation 15.2 implies that $\frac{\partial E(R_{mt}-rf_t|\Omega_{t-1})}{\partial \sigma^2_{mt}}$ consists of two components: $\frac{\partial f_1}{\partial \hat{\sigma}^2_{mt}}$ and $\frac{\partial f_2}{\partial\left(\varepsilon^2_{mt-1}-\hat{\sigma}^2_{mt-1}\right)}$. While the first term $\frac{\partial f_1}{\partial \hat{\sigma}^2_{mt}}$ measures the effect of predictable market volatility on the relation, the second term $\frac{\partial f_2}{\partial\left(\varepsilon^2_{mt-1}-\hat{\sigma}^2_{mt-1}\right)}$ captures the effects of an unexpected volatility shock on the intertemporal relation. However, most of the previous empirical studies on this topic completely ignore the second term and focus only on the first term. Consequently, their empirical results show only a partial intertemporal relation.

Thus, we examine the empirical nature of the full intertemporal relation by considering not only the predictable conditional volatility but also the effect of an unexpected volatility shock on the relation. Especially, we define an unexpected volatility shock in two separate cases of a positive and negative unexpected volatility shock to examine an asymmetrical effect of a positive and negative unexpected shock on the relation, if any. A positive (negative) unexpected volatility shock is denoted as $\varepsilon^2_{mt-1} > \hat{\sigma}^2_{mt-1}$ ($\varepsilon^2_{mt-1} < \hat{\sigma}^2_{mt-1}$), implying that the actual market volatility is higher (lower) than the predicted conditional market volatility. We then examine the sign of intertemporal relation under each case of $\varepsilon^2_{mt-1} > \hat{\sigma}^2_{mt-1}$ and $\varepsilon^2_{mt-1} < \hat{\sigma}^2_{mt-1}$, respectively.

We specify the linear form of intertemporal relation with a dummy variable to capture the asymmetric effect of a positive and negative unexpected volatility shock on the relation

$$R_{mt} - rf_t = \alpha + \lambda_1 \hat{\sigma}_{mt} + \lambda_2 \hat{\sigma}_{mt} \cdot d_t + \varepsilon_{mt}, \tag{15.3}$$

where ε_{mt} is a series of white noise innovations, and $\hat{\sigma}_{mt}$ is the conditional standard deviation of market portfolio. Note that we use $\hat{\sigma}_{mt}$ instead of $\hat{\sigma}^2_{mt}$ as the conditional forecasts of stock market volatility. The use of $\hat{\sigma}_{mt}$ is suggested as the slope of the capital market line by Merton (1980). Also, in estimation, using $\hat{\sigma}_{mt}$ is expected to yield an improvement in the statistical efficiency, mainly due to a reduction in the mean square error of the regression. d_t is the dummy variable for a positive or negative unexpected volatility shock. It takes the value 1 with a prior unexpected positive volatility shock (i.e., $\varepsilon^2_{mt-1} > \hat{\sigma}^2_{mt-1}$) and 0 otherwise. The intertemporal relation is thus measured by $\lambda_1 + \lambda_2$ when $\varepsilon^2_{mt-1} > \hat{\sigma}^2_{mt-1}$ with $d_{\mathrm{t}} = 1$ or by λ_1 otherwise with $d_t = 0$. The asymmetrical effect of a positive and negative unexpected volatility shock on the intertemporal relation, if any, is captured by λ_2. Specifically, $\lambda_2 > 0$ (or $\lambda_1 + \lambda_2 > \lambda_1$) implies that a positive volatility shock has a positive impact on the intertemporal relation.

15.3 Empirical Models

15.3.1 Asymmetric Nonlinear Smooth Transition GARCH Model

To generate the forecast of time-varying market volatility, we employ the asymmetric nonlinear smooth transition (ANST) GARCH model that is capable of capturing the asymmetric volatility effect of a positive and negative return shock. The key feature of the model is the regime-shift mechanism that allows a smooth, flexible transition of volatility between different states of volatility persistence. For monthly excess return series r_t, we specify

$$h_t = \left[a_0 + a_1 \varepsilon^2_{t-1} + a_2 h_{t-1}\right] + \left[b_0 + b_1 \varepsilon^2_{t-1} + b_2 h_{t-1}\right] F(\varepsilon_{t-1}), \tag{15.4}$$

where $F(\varepsilon_{t-1}) = \{1 + \exp[-\gamma(\varepsilon_{t-1})]\}^{-1}$ and $\varepsilon_t = r_t - E(r_t|\Omega_{t-1})$ with $\varepsilon_t|\Omega_{t-1} \sim N(0, h_t)$. Given $\varepsilon_t = v_t \cdot \sqrt{h_t}$, the normalized residuals are distributed as $v_t \overset{iid}{\sim} N(0,1)$. The logistic transition function $F(\varepsilon_{t-1})$ is a smooth and continuous function of ε_{t-1} and the speed parameter γ and takes a value between 0 and 1: $0 < F(\varepsilon_{t-1}) < 0.5$ for $\varepsilon_{t-1} < 0$, $0.5 < F(\varepsilon_{t-1}) < 1$ for $\varepsilon_{t-1} > 0$, and $F(\varepsilon_{t-1}) = 0.5$ for $\varepsilon_{t-1} = 0$. The volatility persistence is measured by $(a_1 + a_2) + (b_1 + b_2)F$, and the condition $b_1 + b_2 < 0$ captures the excess volatility of a negative return shock. For any negative return shock that causes $0 < F(\varepsilon_{t-1}) < 0.5$, the current volatility is described as a *high-persistence-in-volatility regime*. For any positive return shock causing $0.5 < F(\varepsilon_{t-1}) < 1$, the current volatility is described as a *low-persistence-in-volatility regime*. When $\varepsilon_{t-1} = 0$, $F(\varepsilon_{t-1}) = 0.5$, which implies that the current volatility h_t is halfway between the upper and lower volatility regimes. γ governs the speed of transition between volatility regimes. When γ approaches ∞, our model with $b_0 = b_2 = 0$ degenerates into the GJR model.

Note that the volatility transition mechanism in GARCH models has been applied in the several models, such as the modified GARCH model by Glosten et al. (1993), the smooth transition GARCH model by Gonzalez-Rivera (1998), the SVSARCH (sign- and volatility-switching ARCH) model by Fornari and Mele (1997), and the MSVARCH (Markov switching volatility ARCH) model by Turner et al. (1989) and Hamilton and Susmel (1994). For more details, see Harvey (1993), Lutkepol (1993), Hamilton (1994), Campbell et al. (1997), Gourieroux and Monfort (1997), and Rothman (1999).

15.3.2 Empirical Models for the Intertemporal Relation

We examine a simple linear form of the intertemporal relation in the following model for monthly excess return series r_t:

Model 1:

$$\begin{aligned} r_t &= \mu + \phi r_{t-1} + \delta\sqrt{h_t} + \varepsilon_t \\ h_t &= \left[a_0 + a_1\varepsilon_{t-1}^2 + a_2 h_{t-1}\right] + \left[b_0 + b_1\varepsilon_{t-1}^2 + b_2 h_{t-1}\right]F(\varepsilon_{t-1}), \end{aligned} \tag{15.5}$$

where $F(\varepsilon_{t-1}) = \{1 + \exp[-\gamma(\varepsilon_{t-1})]\}^{-1}$. We include the first-order autoregressive term in the mean equation to capture the serial dependence in returns. The intertemporal relation is measured by the coefficient δ.

Model 1, however, ignores the asymmetric effect of an unexpected volatility shock on the intertemporal relation; hence, it suffers from the omitting variable problem. The estimate of δ in Model 1 measures only a partial intertemporal relation. To measure the full intertemporal relation, we present Model 2 in the following specification:

Model 2:

$$\begin{aligned} r_t &= \mu + \phi r_{t-1} + [\delta + \tau M_t)]\sqrt{h_t} + \varepsilon_t \\ h_t &= \left[a_0 + a_1\varepsilon_{t-1}^2 + a_2 h_{t-1}\right] + \left[b_0 + b_1\varepsilon_{t-1}^2 + b_2 h_{t-1}\right]F(\varepsilon_{t-1}), \end{aligned} \tag{15.6}$$

where M_t is a dummy variable to capture the asymmetrical effect of a positive and negative unexpected volatility shock on the intertemporal relation. We define $\hat{\varepsilon}_{t-1}^2 - h_{t-1} > 0 \cdot \left(\hat{\varepsilon}_{t-1}^2 - h_{t-1} < 0\right)$ as a *positive* (*negative*) unexpected volatility shock, which implies that actual volatility is greater (less) than expected. It thus takes a value 1 if $\hat{\varepsilon}_{t-1}^2 > h_{t-1}$ or 0 otherwise.

The sign of the intertemporal relation is measured by the sign of $\delta + \tau$ under a positive volatility shock ($\hat{\varepsilon}_{t-1}^2 > h_{t-1}$), while it is measured by the sign of δ under a negative volatility shock ($\hat{\varepsilon}_{t-1}^2 < h_{t-1}$). Thus, τ captures an asymmetrical effect of a positive and negative volatility shock on the relation. We put an empirical focus on $\tau > 0$ (or $\delta + \tau > \delta$), which indicates a positive impact of a positive unexpected volatility shock on the intertemporal relation.[6]

We specify Model 3 to examine whether allowing the asymmetry for the constant term affects the estimation of the intertemporal relation. A possibility is that the asymmetrical effect of a positive and negative volatility shock on the relation might disappear.

Model 3:

$$\begin{aligned} r_t &= (\mu_1 + \mu_2 M_t) + \phi r_{t-1} + (\delta + \tau M_t)\sqrt{h_t} + \varepsilon_t \\ h_t &= \left[a_0 + a_1 \varepsilon_{t-1}^2 + a_2 h_{t-1}\right] + \left[b_0 + b_1 \varepsilon_{t-1}^2 + b_2 h_{t-1}\right] F(\varepsilon_{t-1}), \end{aligned} \tag{15.7}$$

where $\mu_2 M_t$ captures the asymmetric effect of a prior positive and negative volatility shock on the level of the conditional mean return. In Model 3, we focus on whether the estimated value of τ is still statistically significant even with a presence of $\mu_2 M_t$.

15.4 Empirical Results

15.4.1 The Data

We employ the excess market returns as the expected market risk premiums. To generate the excess returns, we use the monthly nominal returns of the value- and equal-weighted market portfolio index of the NYSE, AMEX, and NASDAQ from the CRSP data files from 1926:01 to 2008:12. The monthly excess return series is constructed by subtracting the 1-month US T-bill returns reported by Ibbotson Associates from the monthly nominal index returns. The excess return series is computed as percentage returns. We employ three sample periods: the full period (1926:01–2008:12), the pre-87 Crash period (1926:01–1987:09), and the GJR period (1951:04–1989:12). Table 15.1 reports the summary statistics for the data. The descriptive statistics indicate that both the nominal and the excess returns series of the two market indexes exhibit significant excess kurtosis and positive first-order autocorrelation, characterizing the nonnormality of the short-horizon stock returns.

15.4.2 Estimation Results, Interpretations, and Diagnostics

We employ the maximum likelihood method with the analytical derivatives of each parameter provided in the Gauss code. All the statistical inferences are based on the Bollerslev-Wooldrige (1992) robust standard errors. Estimation results of

[6]Note that $\tau = 0$ supports the volatility irrelevance argument by Poterba and Summers (1986).

Table 15.1 Summary statistics for monthly nominal and excess returns

	Nominal value-weighted returns		1-monthT-bill rates		Excess value-weighted returns	
Statistics	Full period	Subperiod	Full period	Subperiod	Full period	Subperiod
Observations	996	741	996	741	996	741
Mean ($\times$100)	0.895	0.961	0.303	0.287	0.592	0.674
Std. dev. ($\times$100)	5.352	5.706	0.253	0.276	5.363	5.722
Skewness	0.18	0.368	1.104	1.261	0.227	0.406
Kurtosis	11.221	10.888	4.684	4.624	11.246	10.875
1st order Autocorrelation	0.112 (0.000)	0.107 (0.003)	0.966 (0.000)	0.967 (0.000)	0.115 (0.000)	0.111 (0.002)

The nominal return series are the monthly value-weighted market index returns for NYSE, AMEX, and NASDAQ stocks and were retrieved from the CRSP tapes for the period from 1926:01 to 2008:12. The monthly nominal return series are the value-weighted market indexes retrieved from the CRSP tapes for the same. The monthly excess return series is computed by subtracting 1-month treasury bill returns as reported by Ibbotson Associates from the nominal returns. All returns are computed as percentage value. The analysis employs the full period (1926:01–2008:12) and the pre-87 Crash period (1926:01–1987:09). The value in the parentheses is the p-value for the Ljung-Box Q test

Model 1 for the full period (1926:01–2008:12) and the pre-87 Crash period (1926:01–1987:09) are reported in Table 15.2. Model 1 examines the sign of the simple linear relation between the expected risk premium and the predictable market volatility. Estimation results for Model 1 show that the intertemporal relation in a simple linear form is significantly positive for both the full sample and subsample periods. The estimated value of δ is 0.077 (0.080 for the pre-87 Crash period) and statistically significant at the 1 % level for both sample periods.

With regard to the conditional variance equation, the asymmetric volatility response of a positive and negative return shock is well captured by $b_1 + b_2 < 0$ with a statistical significance. The average estimated value of h_t for the full sample period under a lower volatility regime is 36.37, while it is 24.00 under a higher volatility regime. Also, the estimation results show a high estimated value of the transition parameter γ, which indicates that the transition between volatility regimes occurs very quickly. This implies that the volatility regime is divided into only two extreme regimes: the upper and lower volatility regimes. The upper (lower) regime is induced by any negative (positive) return shock and exhibits a high volatility persistence (a low volatility persistence) in the conditional volatility process.

Estimation results of Model 2 are reported in Table 15.2. There are several notable findings. First, a positive volatility shock has a positive impact on the intertemporal relation. The results show that the estimated value of τ is significantly positive (0.052 for the full period and 0.041 the pre-87 Crash period) and statistically significant at the 1 % level for both sample periods. This implies that an unexpected volatility shock is priced such that, for a positive volatility shock, the expected risk premium for the full sample period increases by 5.2 % (4.1 % for the pre-87 Crash period) of the predicted conditional volatility $\sqrt{h_t}$. Second, the results show that the intertemporal coefficients are all positive for both sample periods. The estimated value for the full period $\delta + \tau = 0.0878$ is (0.093 for the pre-87 Crash period) under prior positive volatility shock, while it is $\delta = 0.035$ for the full period and 0.052 for the pre-87 Crash period under prior negative volatility shock. The results indicate that the sign of the intertemporal relation is indeed positive when the effect of an unexpected volatility shock is incorporated in estimation. Third, the estimation result of $\tau > 0$ indicates that the slope of the capital market line is relatively steeper under a positive volatility shock than under a negative volatility shock. This implies that a positive volatility shock increases the degree of risk aversion.

Estimation results of Model 3 are also reported in Table 15.2. In Model 3, we examine the possibility that the asymmetrical effect of a positive and negative volatility shock may disappear under the presence of asymmetry in the constant term allowed in the conditional mean equation. The estimation results show a positive value of μ_2 for both periods, indicating that a prior positive volatility shock raises the conditional mean returns. This asymmetrical effect of a prior volatility shock on the constant term is more profound for the pre-87 Crash period (μ_2 is 0.351 and statistically significant at the 1 % level). With respect to the intertemporal relation, τ is still positive and statistically

Table 15.2 Estimation results of models 1–3

	Model 1		Model 2		Model 3	
	Full period	Subperiod	Full period	Subperiod	Full period	Subperiod
μ_1	0.420 (3.315)	0.366 (1.485)	0.150 (2.233)	0.251 (2.356)	0.116 (1.224)	0.168 (1.152)
μ_2					0.109 (1.163)	0.351 (3.250)
ϕ	0.077 (6.006)	0.080 (3.732)	0.041 (3.015)	0.072 (8.384)	0.030 (3.013)	0.051 (2.893)
δ	0.073 (4.749)	0.080 (3.126)	0.035 (2.426)	0.052 (2.610)	0.033 (1.915)	0.048 (2.115)
τ			0.052 (5.756)	0.041 (4.015)	0.054 (6.913)	0.042 (2.533)
a_0	0.005 (0.779)	0.001 (1.813)	0.030 (0.300)	0.007 (0.186)	0.022 (0.567)	0.022 (1.306)
a_1	0.131 (3.324)	0.142 (2.689)	0.118 (3.352)	0.141 (3.586)	0.119 (3.438)	0.141 (3.491)
a_2	1.028 (25.817)	1.028 (25.009)	1.063 (26.978)	1.025 (24.875)	1.057 (26.108)	1.027 (25.225)
b_0	2.555 (2.306)	2.855 (2.457)	2.696 (2.816)	2.782 (3.233)	2.652 (2.776)	2.707 (3.367)
b_1	−0.055 (−0.820)	−0.087 (−0.951)	−0.038 (−0.682)	−0.085 (−1.487)	−0.038 (−0.695)	−0.081 (−0.834)
b_2	−0.358 (−3.539)	−0.368 (−3.846)	−0.408 (−4.643)	−0.354 (−3.916)	−0.397 (−4.329)	−0.364 (−3.142)
γ	137.382 (2.057)	97.516 (1.090)	123.324 (3.044)	85.377 (1.577)	107.355 (2.380)	99.011 (1.330)

This table presents the maximum likelihood estimates for models 1−3 for monthly excess returns of the value-weighted index for the NYSE, AMEX, and NASDAQ stocks. The estimation employs the full period (1926:01–2008:12) and the pre-87 Crash period (1926:01–1987:09). For the monthly excess return series r_t, each model is specified as follows

Model 1: $r_t = \mu_1 + \phi r_{t-1} + \delta\sqrt{h_t} + \varepsilon_t$

Model 2: $r_t = \mu_1 + \phi r_{t-1} + (\delta + \tau M_t)\sqrt{h_t} + \varepsilon_t$

Model 3: $r_t = \mu_1 + \mu_2 M_t + \phi r_{t-1} + (\delta + \tau M_t)\sqrt{h_t} + \varepsilon_t$,

where the indicator function M_t is specified to capture the asymmetric effect of an unexpected volatility shock on the intertemporal relation, such that $M_t = 1$ if $\hat{\varepsilon}^2_{t-1} > h_{t-1}$ or $M_t = 0$ otherwise. The conditional variance equation for the models is specified as $h_t = [a_0 + a_1\varepsilon^2_{t-1} + a_2 h_{t-1}] + [b_0 + b_1\varepsilon^2_{t-1} + b_2 h_{t-1}]F(\varepsilon_{t-1})$, where the transition function $F(\varepsilon_{t-1})$ is defined as $F(\varepsilon_{t-1})=\{1 + \exp[-\gamma(\varepsilon_{t-1})]\}^{-1}$. The values in parentheses are the Bollerslev-Wooldridge robust t-statistics significant at the 1 % level for both periods (τ_1 is 0.054 for the full period and 0.048 for the pre-87 Crash period). This implies that, even with the asymmetry allowed in the constant term, there still exists a significant asymmetrical effect of a positive and negative volatility shock on the intertemporal relation. Also, δ is still significantly positive for both periods, confirming the positive full intertemporal relation.

Table 15.3 reports the summary of diagnostics for the estimation results, such as skewness, kurtosis, the Jarque-Bera normality test, and the Ljung-Box Q test on the normalized and the squared normalized residuals. The Ljung-Box Q statistics on the normalized residuals checks serial correlation in the residuals. Rejection of the null of no autocorrelation up to a certain lag length indicates that either the dynamics of the conditional mean or the lag structure of the conditional variance process is not well specified or that both equations are not well specified by the model. The Ljung-Box statistics on the *squared* normalized residuals ascertains if the serial dependence in the conditional variance is well captured by the equation.

We also perform the *negative sign bias test* (NSBT) suggested by Engle and Ng (1993) to examine the ability of the model to capture the so-called leverage effect of a negative return shock on the conditional variance process. The negative sign bias test shows insignificant t-values for all estimations and indicates that the asymmetric volatility response to a positive and negative return shock is well captured by the ANST-GARCH model.[7] The Ljung-Box $Q(10)$ test indicates that the serial dependence of the conditional mean and variance is well captured by models 1–3.

15.4.3 GJR Sample Period

Using modified (E)GARCH-M models, Glosten et al. (1993) report a strong negative intertemporal relation for their (GJR) sample period of 1951:04–1989:12. However, they do not consider the effect of an unanticipated volatility shock on the estimation of the relation, so that their results reflect only a partial intertemporal relation. We thus evaluate the full intertemporal relation for their sample period by incorporating the effect of an unexpected volatility shock in the estimation of the relation. Allowing a time dummy in the conditional mean equation to capture the characteristic of the GJR sample period, we specify Model 5 as follows:

[7] The negative sign bias test is performed with the regression equation $v_t^2 = a + bS_{t-1}^{-}\varepsilon_{t-1} + \pi' z_t^* + e_t$, where $v_t^2 = \left(\varepsilon_t/\sqrt{h_t}\right)^2$. $S_{t-1}^{-} = 1$ if $\varepsilon_{t-1} < 0$, and $S_{t-1}^{-} = 0$ otherwise. Also, $z_t^* = \widetilde{h}(\Psi)/h_t$, where $\widetilde{h}(\Psi) = \partial h_t/\partial \Psi$ evaluated at the values of maximum likelihood estimates of parameter Ψ. The test statistic of the NSBT is defined as the t-ratio of the coefficient b in the regression. A statistically significant t-value implies the failure of the model to absorb the effect of sign bias and indicates that the volatility model considered is misspecified.

Table 15.3 Diagnostics of models 1–3

	Model 1		Model 2		Model 3	
	Full period	Subperiod	Full period	Subperiod	Full period	Subperiod
Skewness of v_t	−0.610	−0.372	−0.631	−0.375	−0.628	−0.366
Kurtosis of v_t	4.961	3.909	5.013	3.877	4.983	3.850
JB-Normality	220.845 (0.000)	42.531 (0.000)	233.794 (0.000)	41.003 (0.000)	228.080 (0.000)	38.743 (0.000)
Q(10) on v_t	11.209 (0.341)	17.801 (0.058)	13.368 (0.204)	18.010 (0.055)	13.139 (0.216)	17.663 (0.061)
Q(10) on v_t^2	7.156 (0.711)	12.905 (0.229)	9.660 (0.471)	13.148 (0.215)	9.516 (0.484)	13.639 (0.190)
NSBT on h_t	−1.328 (0.184)	−0.579 (0.563)	−0.526 (0.599)	−1.275 (0.203)	−0.688 (0.492)	−1.365 (0.173)

This table presents a summary of diagnostics on the normalized residuals and the squared normalized residuals from the estimations. The normalized residual series is defined as $v_t = \varepsilon_t/\sqrt{h_t}$. JB-Normality refers to the Jarque-Bera normality test statistic, which is distributed as χ^2 with two degrees of freedom under the null hypothesis of normally distributed residuals. Q(10) is the Ljung-Box Q(10) test statistic for checking serial dependence in the normalized residuals and the squared normalized residuals from the estimations. NSBT refers to the negative sign bias test suggested by Engle and Ng (1993). It is a diagnostic test that examines the ability of the specified model to capture the so-called leverage effect of a negative return shock on the conditional volatility process. The test is performed with the regression equation $v_t^2 = a + bS_{t-1}^{-}\varepsilon_{t-1} + \pi' z_t^* + e_t$, where $v_t^2 = \left(\varepsilon_t/\sqrt{h_t}\right)^2$. S_{t-1}^{-}=1 if $\varepsilon_{t-1} < 0$, and $S_{t-1}^{-} = 0$ otherwise. Also, $z_t^* = \widetilde{h}(\Psi)/h_t$, where $\widetilde{h}(\Psi) = \partial h_t/\partial \Psi$, is evaluated at the values of the maximum likelihood estimates of parameter Ψ. The test statistic of the NSBT is defined as the t-ratio of the coefficient b in the regression. The value in the parentheses is the p-value of the individual test statistics considered

Model 4 (Modified Model 1 for GJR Period):

$$\begin{aligned} r_t &= \mu + \phi r_{t-1} + \left(\delta + \delta^G G_t\right)\sqrt{h_t} + \varepsilon_t \\ h_t &= \left[a_0 + a_1\varepsilon_{t-1}^2 + a_2 h_{t-1}\right] + \theta rf_t \\ &\quad + \left[b_0 + b_1\varepsilon_{t-1}^2 + b_2 h_{t-1}\right]F(\varepsilon_{t-1}), \end{aligned} \tag{15.8}$$

where G_t is a dummy variable that takes a value 1 for the GJR sample period or 0 otherwise. The coefficient δ^G captures the differential effect of the GJR period, if any, on the intertemporal relation, while δ measures the relation for the full period. The negative intertemporal relation reported by Glosten et al. (1993) can be confirmed by $\delta + \delta^G < 0$. One of the important features of the above modified Model 1 is its capacity to distinguish the relation for the GJR sample period from that for the entire sample period. We also estimate the same model with and without the 1-month T-bill returns rf_t included in the conditional variance equation.[8]

Estimation results of Model 4 are reported in Table 15.4. A notable finding is that the estimated value of δ^G is strongly negative (–0.103 with rf_t and –0.092 without rf_t) and highly significant with $\delta + \delta^G < 0$. This result implies that, comparing to the full period, the GJR sample period is especially characterized by a strong negative intertemporal relation, and this result is consistent with that of Glosten et al. (1993).

As mentioned earlier, however, this result does not consider the effect of an unexpected volatility shock on the relation, reflecting only a simple partial intertemporal relation. In order to examine the full relation for the GJR sample period, we specify Model 5 as follows:

Model 5 (Modified Model 2 for GJR Period):

$$\begin{aligned} r_t &= \mu + \phi r_{t-1} + \left[(\delta + \tau M_t) + \left(\delta^G + \tau^G M_t\right)G_t\right]\sqrt{h_t} + \varepsilon_t \\ h_t &= \left[a_0 + a_1\varepsilon_{t-1}^2 + a_2 h_{t-1}\right] + \theta rf_t + \left[b_0 + b_1\varepsilon_{t-1}^2 + b_2 h_{t-1}\right]F(\varepsilon_{t-1}), \end{aligned} \tag{15.9}$$

where G_t is a time dummy variable that takes the value 1 for the GJR period or 0 otherwise. The full intertemporal relation for the GJR period is measured by $\delta + \delta^G + \tau + \tau^G$ under a positive volatility shock and by $\delta + \delta^G$ under a negative volatility shock, such that the differential effect of the GJR period on the full relation is captured by $\delta^G + \tau^G$. We also estimate the same model with and without rf_t included in the conditional variance equation.

Estimation results of Model 5 for the GJR sample period are reported in Table 15.4. It shows that the estimated value of all four important parameters (δ, τ, δ^G, and τ^G) capturing the intertemporal relation are statistically significant at the 1 % level. There are several notable findings. First, the estimation result of

[8]Several studies show that the estimation results are sensitive to the inclusion of 1-month T-bill return in the conditional variance equation. See Campbell (1987), Glosten et al. (1993), and Scruggs (1998).

Table 15.4 Estimation results of models 4 and 5 for the GJR sample period

	Model 4 for GJR sample period		Model 5 for GJR sample period	
Coeff.	With *rf*	Without *rf*	With *rf*	Without *rf*
μ	0.546 (9.944)	0.443 (9.288)	0.531 (6.136)	0.543 (5.166)
ϕ	0.056 (5.930)	0.072 (7.909)	0.051 (5.143)	0.051 (5.394)
δ	0.059 (5.925)	0.080 (6.392)	0.103 (4.164)	0.100 (3.678)
τ			0.046 (4.312)	0.046 (3.778)
δ^G	−0.103 (−12.369)	−0.092 (−8.882)	−0.155 (−12.540)	−0.156 (−11.720)
τ^G			0.067 (6.560)	0.091 (6.903)
a_0	0.001 (0.766)	0.001 (0.581)	0.007 (0.238)	0.014 (0.202)
a_1	0.110 (3.199)	0.110 (3.291)	0.104 (3.359)	0.102 (3.303)
a_2	1.060 (25.142)	1.072 (28.727)	1.061 (26.070)	1.071 (29.469)
b_0	2.482 (2.564)	3.007 (3.464)	2.270 (2.538)	2.689 (2.971)
b_1	−0.029 (−0.457)	−0.042 (−0.795)	−0.019 (−0.337)	−0.021 (−0.399)
b_2	−0.412 (−4.672)	−0.428 (−4.883)	−0.404 (−4.420)	−0.416 (−4.840)
θ	0.885 (1.065)		0.901 (1.128)	
γ	121.735 (1.944)	217.354 (2.441)	287.546 (2.128)	254.235 (1.963)

The GJR sample period (1951:04–1989:12) is evaluated by incorporating the period as a dummy variable in the mean equation, with the 1-month T-bill rates included or excluded in the conditional variance equation. The modified models to estimate the GJR sample period are as follows
Model 4 (modified Model 1 for GJR period): $r_t = \mu + \phi r_{t-1} + \left(\delta + \delta^G G_t\right)\sqrt{h_t} + \varepsilon_t$
Model 5 (modified Model 2 for GJR period): $r_t = \mu + \phi r_{t-1} + \left[(\delta + \tau M_t) + \left(\delta^G + \tau^G M_t\right) G_t\right]\sqrt{h_t} + \varepsilon_t$
where the indicator function G_t is a dummy variable for the GJR sample period, and rf_t is the yield on 1-month T-bill from Ibbotson Associates. The function M_t is specified to capture the asymmetric effect of an unexpected volatility shock on the relation, such that $M_t = 1$ if $\hat{\varepsilon}^2_{t-1} > h_{t-1}$; otherwise, $M_t = 0$. The conditional variance equation for the two estimation models is specified as $h_t = [a_0 + a_1\varepsilon^2_{t-1} + a_2 h_{t-1}] + [b_0 + b_1\varepsilon^2_{t-1} + b_2 h_{t-1}]F(\varepsilon_{t-1})$, where $F(\varepsilon_{t-1}) = \{1 + \exp[-\gamma(\varepsilon_{t-1})]\}^{-1}$. The values in parentheses are the Bollerslev-Wooldridge robust *t*-statistics

$\delta^G + \tau^G = -0.088$ (–0.065 without rf_t) confirms that the GJR period is indeed characterized by a significant negative relation. Second, the estimation result of $\tau + \tau^G = 0.113$ (0.137 without rf_t) implies that, even for the GJR sample period, there still exists a significant positive effect of a positive volatility shock on the full intertemporal relation. Third, the estimation results of $\delta + \delta^G + \tau + \tau^G = 0.061$ (0.081 without rf_t) and $\delta + \delta^G = -0.052$ (–0.056 without rf_t) imply that the GJR sample period exhibits a positive intertemporal relation under a positive volatility shock and a negative relation under a negative volatility shock. Note the above results are not sensitive to the inclusion of rf_t in estimation.

Diagnostic tests in Table 15.5 indicate that all the estimations pass the Ljung-Box Q(10) test on the normalized and squared normalized residuals. This result implies that there is no serial dependence remaining in the conditional mean and variance processes. The negative sign bias test shows insignificant *t*-values for all estimations, indicating that the estimated conditional variance process well captures the excess volatility response caused by a negative return shock.

Table 15.5 Diagnostics of the models 4 and 5 for the GJR sample period

	Model 4 for GJR sample period		Model 5 for GJR sample period	
	Without *rf*	With *rf*	Without *rf*	Without *rf*
Skewness of v_t	−0.636	−0.623	−0.646	−0.626
Kurtosis of v_t	5.251	5.083	5.216	5.018
JB-Normality	246.664 (0.000)	217.536 (0.000)	242.89 (0.000)	208.29 (0.000)
Q(10) on v_t	10.190 (0.424)	10.397 (0.406)	10.111 (0.431)	10.846 (0.370)
Q(10) on v_t^2	10.514 (0.397)	10.418 (0.405)	9.7167 (0.466)	10.122 (0.430)
NSBT on h_t	−1.198 (0.231)	−1.162 (0.246)	−1.205 (0.228)	−1.105 (0.269)

This table presents a summary of diagnostics on the normalized residuals and the squared normalized residuals from the estimations. The normalized residual series is defined as $v_t = \varepsilon_t/\sqrt{h_t}$. JB-Normality refers to the Jarque-Bera normality test statistic, which is distributed as χ^2 with two degrees of freedom under the null hypothesis of normally distributed residuals. Q(10) is the Ljung-Box Q(10) test statistic for checking serial dependence in the normalized residuals and the squared normalized residuals from the estimations. NSBT refers to the negative sign bias test suggested by Engle and Ng (1993). It is a diagnostic test that examines the ability of the specified model to capture the so-called leverage effect of a negative return shock on the conditional volatility process. The test is performed with the regression equation $v_t^2 = a + bS_{t-1}^-\varepsilon_{t-1} + \pi' z_t^* + e_t$, where $v_t^2 = \left(\varepsilon_t/\sqrt{h_t}\right)^2$. $S_{t-1}^- = 1$ if $\varepsilon_{t-1} < 0$, and $S_{t-1}^- = 0$ otherwise. Also, $z_t^* = \widetilde{h}(\Psi)/h_t$, where $\widetilde{h}(\Psi) = \partial h_t/\partial\Psi$, is evaluated at the values of the maximum likelihood estimates of parameter Ψ. The test statistic of the NSBT is defined as the t-ratio of the coefficient b in the regression. The value in the parentheses is the p-value of the individual test statistics considered

15.4.4 The Link Between Asymmetric Mean Reversion and Intertemporal Relation

It has been known that the expected market returns exhibit an asymmetric mean-reverting pattern that negative returns are more likely to revert to positive returns than positive returns reverting to negative returns. However, the quicker reversion of negative returns is hardly justified under a positive intertemporal relation. Under a positive intertemporal relation, a negative return shock raises the risk premium to compensate for the excess volatility, and an increase in risk premium reduces the current stock price, which in turn reduces the concurrent stock price. Thus, if a positive intertemporal relation is correct, a negative return should be more likely to be accompanied by another negative return for the subsequent periods.

Nam et al. (2001) suggest that the quicker reversion of a negative return is attributed to a reduction in the expected risk. They show that the intertemporal relation is significantly negative under a prior negative return shock. They argue that a negative return shock generates an optimistic expectation by investors, of the future performance of a stock experiencing a recent price drop, thereby reducing the expected market risk premium. As a reduction in risk premium in turn raises the current stock price, negative returns are more likely to revert to positive returns.

While successfully explaining the link between the asymmetric mean-reverting property and the intertemporal relation, Nam et al. (2001) do not consider the effect of an unexpected volatility shock on the link. In this section, we investigate the impact of an unexpected volatility shock on the observed link. First, we specify the following nonlinear autoregressive model to confirm the asymmetric mean reversion of the expected market risk premium:

Model 6:

$$\begin{aligned} r_t &= [\mu_1 + \mu_2 F(\varepsilon_{t-1})] + [\phi_1 + \phi_2 F(\varepsilon_{t-1})] r_{t-1} + \varepsilon_t \\ h_t &= \left[a_0 + a_1 \varepsilon_{t-1}^2 + a_2 h_{t-1}\right] + \left[b_0 + b_1 \varepsilon_{t-1}^2 + b_2 h_{t-1}\right] F(\varepsilon_{t-1}), \end{aligned} \tag{15.10}$$

where the asymmetry is allowed in both the conditional mean and variance, such that the asymmetry in both processes is controlled by a prior return shock. The return serial correlation varies between ϕ_1 and $\phi_1 + \phi_2$ depending on the value of the transition function $F(\varepsilon_{t-1})$. Under an extreme negative prior return shock that causes $F(\varepsilon_{t-1}) = 0$, serial correlation is measured by ϕ_1, while it is measured by $\phi_1 + \phi_2$ for an extreme positive return shock yielding $F(\varepsilon_{t-1}) = 1$.[9] A quicker reversion of a negative return is captured by $\phi_1 > 0$ (or $\phi_1 + \phi_2 > \phi_1$). Note that the condition $\phi_1 < 0$ (a negative serial correlation under $\varepsilon_{t-1} < 0$) indicates a stronger reverting tendency of a negative return.

Estimation results of Model 6 are reported in Table 15.6. It shows that the estimated value of ϕ_2 is positive and highly significant for both the full period and the pre-87 Crash period. The measured serial correlation is negative ($\phi_1 = -0.078$ and –0.076, respectively, for the two periods) under a prior negative return shock, while it is positive ($\phi_1 + \phi_2 = 0.056$ and 0.094, respectively, for the two periods) under a prior positive return shock. This result confirms the asymmetrical reverting pattern of the expected returns that a negative return reverts more quickly, while a positive return tends to persist.

Secondly, we specify Model 7 to examine if there is a link between the asymmetric mean reversion and the intertemporal relation.

Model 7:

$$\begin{aligned} r_t &= [\mu_1 + \mu_2 F(\varepsilon_{t-1})] + \phi_1 r_{t-1} + [\delta_1 + \delta_2 F(\varepsilon_{t-1})]\sqrt{h_t} + \varepsilon_t \\ h_t &= \left[a_0 + a_1 \varepsilon_{t-1}^2 + a_2 h_{t-1}\right] + \left[b_0 + b_1 \varepsilon_{t-1}^2 + b_2 h_{t-1}\right] F(\varepsilon_{t-1}). \end{aligned} \tag{15.11}$$

The partial intertemporal relation is measured by the estimated value of δ_1 under a prior negative return shock causing $F(\varepsilon_{t-1}) = 0$, while it is measured by $\delta_1 + \delta_2$ under a prior positive return shock causing $F(\varepsilon_{t-1}) = 1$. Thus, δ_2 measures the differential impact of a positive and negative return shock on the partial intertemporal relation.

The estimation results presented in Table 15.6 show two notable findings. First, δ_2 is positive and statistically significant at the 1 % level ($\delta_2 = 0.258$ for the full period and 0.334 for the pre-87 Crash period). This result implies that there is

[9] Stationarity condition of r_t is satisfied with $|\phi_1 + \phi_2 F(\varepsilon_{t-1})| < 1$, i.e., $|\phi_1| < 1$ for $\varepsilon_{t-1} < 0$ or $|\phi_1 + \phi_2| < 1$ for $\varepsilon_{t-1} > 0$.

Table 15.6 Estimation results of models 6–8

	Model 6		Model 7		Model 8	
Coef.	Full period	Subperiod	Full period	Subperiod	Full period	Subperiod
μ_1	0.464 (13.257)	0.489 (9.673)	0.486 (2.569)	0.614 (4.531)	0.972 (6.584)	0.879 (2.429)
μ_2	0.074 (1.667)	−0.169 (−3.267)	−0.034 (−0.147)	−0.229 (−0.946)	−0.181 (−1.172)	0.092 (0.264)
ϕ_1	−0.078 (−8.476)	−0.076 (−11.278)	−0.058 (−7.059)	−0.070 (−5.170)	−0.023 (−2.424)	−0.058 (−5.008)
ϕ_2	0.134 (5.972)	0.170 (10.188)				
δ_1			−0.155 (−3.251)	−0.182 (−5.013)	−0.260 (−4.592)	−0.244 (−2.610)
τ_1					0.079 (2.277)	0.024 (1.111)
δ_2			0.258 (4.070)	0.334 (5.502)	0.229 (3.489)	0.225 (2.329)
τ_2					−0.028 (−0.659)	0.052 (1.983)
a_0	0.001 (0.259)	0.000 (1.115)	0.002 (1.149)	0.001 (0.997)	0.009 (0.982)	0.012 (1.833)
a_1	0.122 (3.656)	0.130 (3.581)	0.133 (3.674)	0.133 (3.774)	0.123 (3.762)	0.139 (3.776)
a_2	1.012 (31.079)	1.011 (29.957)	1.063 (18.379)	1.081 (24.110)	1.060 (35.832)	1.066 (22.949)
b_0	2.154 (2.598)	2.331 (2.998)	2.279 (3.102)	2.339 (3.943)	1.914 (3.139)	1.966 (3.483)
b_1	−0.030 (−0.526)	−0.076 (−1.218)	−0.054 (−1.036)	−0.070 (−1.329)	−0.028 (−0.503)	−0.061 (−1.039)
b_2	−0.312 (−4.200)	−0.306 (−4.391)	−0.383 (−3.406)	−0.405 (−4.893)	−0.368 (−5.506)	−0.379 (−4.390)
γ	126.008 (3.864)	131.391 (2.818)	157.246 (3.609)	216.327 (3.387)	148.326 (2.095)	67.239 (3.027)

This table presents the maximum likelihood estimates of models 6–8 for the monthly excess returns of the value-weighted index for the NYSE, AMEX, and NASDAQ stocks over the period of 1926:01–2008:12 and the period of 1926:01–1987:09. For the monthly excess return series r_t, each model is specified as follows

Model 6: $r_t = [\mu_1 + \mu_2 F(\varepsilon_{t-1})] + [\phi_1 + \phi_2 F(\varepsilon_{t-1})] r_{t-1} + \varepsilon_t$

Model 7: $r_t = [\mu_1 + \mu_2 F(\varepsilon_{t-1})] + \phi r_{t-1} + [\delta_1 + \delta_2 F(\varepsilon_{t-1})]\sqrt{h_t} + \varepsilon_t$

Model 8: $r_t = [\mu_1 + \mu_2 F(\varepsilon_{t-1})] + \phi_1 r_{t-1} + [(\delta_1 + \tau_1 M_t) + (\delta_2 + \tau_2 M_t) F(\varepsilon_{t-1})]\sqrt{h_t} + \varepsilon_t$

where the indicator function M_t is specified to capture the asymmetric effect of an unexpected volatility shock on the intertemporal relation, such that $M_t = 1$ if $\hat{\varepsilon}_{t-1}^2 > h_{t-1}$ or $M_t = 0$ otherwise. The conditional variance equation for Models 6–8 is specified as $h_t = [a_0 + a_1 \varepsilon_{t-1}^2 + a_2 h_{t-1}] + [b_0 + b_1 \varepsilon_{t-1}^2 + b_2 h_{t-1}] F(\varepsilon_{t-1})$. The values in parentheses are the Bollerslev-Wooldridge robust t-statistics a strong asymmetry in the partial intertemporal relation under a prior positive and negative return shock. Second, the partial intertemporal relation is indeed negative under a prior negative return shock ($\delta_1 = -0.155$ for the full period and –0.182 for the pre-87 Crash period), while it is positive under a prior positive return shock ($\delta_1 + \delta_2 = 0.103$ for the full period and 0.152 for the pre-87 Crash period). This result implies that while a positive return shock complies with the conventional positive intertemporal relation, a negative return shock indeed induces a negative intertemporal behavior of the expected market returns. More importantly, the results of Models 6 and 7 confirm the highly significant asymmetric link between the mean reversion and the intertemporal relation. The quicker reversion of negative returns is attributed to the negative intertemporal relation under a prior negative return shock.

Model 7 does not incorporate the asymmetrical impact of a positive and negative volatility shock in the estimation of the intertemporal relation. To get empirically more reliable results on the link between the mean reversion and the full intertemporal relation, we propose Model 8 in the following specification.

Model 8:

$$
\begin{aligned}
r_t &= [\mu_1 + \mu_2 F(\varepsilon_{t-1})] + \phi_1 r_{t-1} + [(\delta_1 + \tau_1 M_t) + (\delta_2 + \tau_2 M_t) F(\varepsilon_{t-1})]\sqrt{h_t} + \varepsilon_t \\
h_t &= \left[a_0 + a_1 \varepsilon_{t-1}^2 + a_2 h_{t-1}\right] + \left[b_0 + b_1 \varepsilon_{t-1}^2 + b_2 h_{t-1}\right] F(\varepsilon_{t-1}),
\end{aligned}
\tag{15.12}
$$

where the full intertemporal relation is measured separately under a positive and negative return shock. When a negative return shock causing $F(\varepsilon_{t-1}) = 0$ is realized, the full intertemporal relation is measured by $\delta_1 + \tau_1$, for which τ_1 captures an asymmetrical impact of a prior unexpected volatility shock on the relation. With a prior positive return shock, the relation is measured by $\delta_1 + \delta_2 + \tau_1 + \tau_2$, for which $\tau_1 + \tau_2$ measures an asymmetrical effect of a prior volatility shock on the relation, if any.

The estimation results of Model 8 are reported in Table 15.6. There are several notable findings. First, the results show that three out of the four important parameters (δ_1, τ_1, δ_2, and τ_2) to capture the full intertemporal relation are statistically significant at the 5 % level. Second, the result of $\delta_1 + \tau_1 = -0.181$ (–0.220 for the pre-87 Crash period) and $\delta_1 + \delta_2 + \tau_1 + \tau_2 = 0.020$ (0.057 for the pre-87 Crash period) implies that the full intertemporal relation is still negative (positive) under a prior negative (positive) return shock. Third, more importantly, the results of $\tau_1 = 0.079$ and $\tau_1 + \tau_2 = 0.051$ (0.024 and 0.076 for the pre-87 Crash period) indicate that a positive volatility shock has a positive impact on the intertemporal relation, regardless of the sign of a prior return shock. This implies that a positive unexpected volatility shock consistently induces a positive impact on the intertemporal risk-return relation.

Table 15.7 reports the results of diagnostic tests. The Ljung-Box Q(10) test results on v_t and v_t^2 indicate that serial dependence is well captured by the specified conditional mean and variance processes. The negative sign bias test shows insignificant *t*-values for all estimations, confirming the capability of the ANST-GARCH model to capture the excess volatility response caused by a negative return shock.

Table 15.7 Diagnostics of models 6–8

	Model 6		Model 7		Model 8	
	Full period	Subperiod	Full period	Subperiod	Full period	Subperiod
Skewness of v_t	−0.627	−0.368	−0.620	−0.343	−0.592	−0.320
Kurtosis of v_t	5.071	3.917	5.111	3.764	4.775	3.725
JB-Normality	242.843 (0.000)	42.563 (0.000)	248.414 (0.000)	32.511 (0.000)	188.671 (0.000)	28.843 (0.000)
Q(10) on v_t	14.059 (0.170)	20.323 (0.026)	10.852 (0.369)	17.070 (0.073)	9.669 (0.470)	15.935 (0.101)
Q(10) on v_t^2	8.379 (0.592)	9.902 (0.449)	10.005 (0.440)	13.429 (0.201)	10.095 (0.432)	13.714 (0.186)
NSBT on h_t	−1.415 (0.157)	−0.769 (0.442)	−1.058 (0.290)	−0.547 (0.585)	−1.248 (0.212)	−0.334 (0.738)

This table presents a summary of diagnostics on the normalized residuals and the squared normalized residuals from the estimations. The normalized residual series is defined as $v_t = \varepsilon_t/\sqrt{h_t}$. JB-Normality refers to the Jarque-Bera normality test statistic, which is distributed as χ^2 with two degrees of freedom under the null hypothesis of normally distributed residuals. Q(10) is the Ljung-Box Q(10) test statistic for checking serial dependence in the normalized residuals and the squared normalized residuals from the estimations. NSBT refers to the negative sign bias test suggested by Engle and Ng (1993). It is a diagnostic test that examines the ability of the specified model to capture the so-called leverage effect of a negative return shock on the conditional volatility process. The test is performed with the regression equation $v_t^2 = a + bS_{t-1}^-\varepsilon_{t-1} + \pi' z_t^* + e_t$, where $v_t^2 = \left(\varepsilon_t/\sqrt{h_t}\right)^2$. $S_{t-1}^- = 1$ if $\varepsilon_{t-1} < 0$, and $S_{t-1}^- = 0$ otherwise. Also, $z_t^* = \widetilde{h}(\Psi)/h_t$, where $\widetilde{h}(\Psi) = \partial h_t/\partial\Psi$, is evaluated at the values of the maximum likelihood estimates of parameter Ψ. The test statistic of the NSBT is defined as the t-ratio of the coefficient b in the regression. The value in the parentheses is the p-value of the individual test statistics considered

15.5 Conclusions

We suggest that the intertemporal risk-return relation is driven not only by the underlying market volatility but also by an unexpected volatility shock. Most of the previous literature on this topic ignores the effect of an unexpected volatility shock on the relation. Thus, their results reflect only a partial intertemporal relation. With the effect of an unexpected volatility shock incorporated in the estimation of the relation, we find a strong positive intertemporal relation for the US monthly excess returns for the period of 1926:01–2008:12.

We also reexamine the relation for the GJR sample period with the effect of a volatility shock incorporated in estimation. The estimation results show that the GJR sample period is indeed characterized by a strong positive (negative) relation under a positive (negative) volatility shock. This implies that the negative relation reported by Glosten et al. (1993) is attributed to ignoring the effect of an unexpected volatility shock on the relation.

Lastly, we examine the observed link between the mean reversion property and the intertemporal relation under a consideration of the impact of an unexpected volatility shock on the link. We confirm that the quicker reversion of negative returns is attributed to the negative intertemporal relation under a prior negative return shock. We interpret this negative intertemporal relation as reflective of strong optimistic expectations as perceived by investors of the future performance of a stock experiencing a recent price drop.

Appendix 1: Method to Derive the Variance of the Parameters with Restrictions

In estimations, we employ the parameter restrictions for the positivity of the conditional variance process. To get the true variance of the restricted parameter, we apply the following functional transformations:

Case 1. Logistic Function Transformation

Suppose that a parameter $\hat{\beta}$ is restricted to a logistic function of $\hat{b}$ in order to guarantee $0 < \hat{\beta} < 1$.

Hence,

$$\hat{\beta} = f\left(\hat{b}\right) = \frac{1}{1 + e^{-\hat{b}}}, \tag{15.13}$$

where $\hat{b}$ is the actual coefficient estimated from computation, and $\hat{\beta}$ is the true parameter estimate we want to transform from Eq. 15.13. Using Taylor expansion, Eq. 15.13 can be expressed as

$$f\left(\hat{b}\right) = f(b) + f'(b)\left(\hat{b} - b\right), \tag{15.14}$$

or

$$f(\hat{b}) - f(b) = f'(b)(\hat{b} - b). \tag{15.15}$$

Using $f(b) = \frac{1}{1+e^{-b}} = \beta$ and Eq. 15.15 can be expressed as

$$\hat{\beta} - \beta = f'(b) \cdot (\hat{b} - b). \tag{15.16}$$

Then, from Eq. 15.16 we can get $\text{var}(\hat{\beta})$ as follows:

$$\text{var}(\hat{\beta}) = E(\hat{\beta} - \beta)^2 = [f'(b)]^2 \cdot E(\hat{b} - b)^2, \tag{15.17}$$

where the true value of $f'(b)$ is not known. We thus use the MLE of b, $\hat{b}$, to get the value of $f'(b)$. Then the variance of $\hat{\beta}$ can be calculated as follows:

$$\text{var}(\hat{\beta}) = [f'(\hat{b})]^2 \cdot \text{var}(\hat{b}), \tag{15.18}$$

where $f'(\hat{b}) = f(\hat{b})[1 - f(\hat{b})]$.

Case 2. Exponential Function Transformation

Suppose that a parameter $\hat{\beta}$ is restricted to an exponential function of $\hat{b}$ in order to guarantee $0 < \hat{\beta} < 1$. Hence,

$$\hat{\beta} = f(\hat{b}) = e^{-\hat{b}}. \tag{15.19}$$

Using Eq. 15.18,

$$\text{var}(\hat{\beta}) = [f'(\hat{b})]^2 \cdot \text{var}(\hat{b}), \tag{15.20}$$

where $f'(\hat{b}) = [f(\hat{b})]^2 = \hat{\beta}^2$.

Appendix 2: News Impact Curve

We derive the functions of the news impact curve (NIC) for the ANST-GARCH and the GJR models as follows:
ANST-GARCH (1,1) Model:

$$\text{Model}: \quad h_t = (a_0 + a_1\varepsilon_{t-1}^2 + g_1 h_{t-1}) + (b_0 + b_1\varepsilon_{t-1}^2 + b_2 h_{t-1}) \cdot F(\varepsilon_{t-1}) \tag{15.21}$$

$$\text{NIC}: \quad h_t = C + [a_1 + b_1 F(\varepsilon_{t-1})]\varepsilon_{t-1}^2$$

and

$$C = a_0 + b_0 F(\varepsilon_{t-1}) + [g_1 + b_2 F(\varepsilon_{t-1})]\sigma^2, \tag{15.22}$$

$$\text{where } F(\varepsilon_{t-1}) = \{1 + \exp[-\gamma(\varepsilon_{t-1})]\}^{-1}.$$

GJR (1, 1) Model:

$$\text{Model}: \quad h_t = a_0 + a_1\varepsilon_{t-1}^2 + b_1 h_{t-1} + \left(a_2\varepsilon_{t-1}^2\right) \cdot I(\varepsilon_{t-1} > 0), \tag{15.23}$$

$$\text{where } I(\varepsilon_{t-1} > 0) = \left\{\begin{matrix} 1, if\, \varepsilon_{t-1} > 0 \\ 0, if\, \varepsilon_{t-1} < 0 \end{matrix}\right\}.$$

$$\text{NIC}: \quad h_t = \left\{\begin{matrix} (a_0 + b_1\sigma^2) + a_1\varepsilon_{t-1}^2, & if\, \varepsilon_{t-1} > 0 \\ (a_0 + b_1\sigma^2) + (a_1 + a_2)\varepsilon_{t-1}^2, & if\, \varepsilon_{t-1} < 0 \end{matrix}\right\}. \tag{15.24}$$

Appendix 3: Sign Bias Tests

Based on the news impact curve, we perform three diagnostic tests for examining the ability of a model to capture asymmetric effect of news on conditional variance: the *sign bias test* (SBT), the *negative sign bias test* (NSBT), and the *positive sign bias test* (PSBT). These tests are performed by the t-statistic on the coefficient b under the following regression equations:

$$v_t^2 = a + bS_{t-1}^- + \beta' z_t^* + e_t \tag{15.25}$$

$$v_t^2 = a + bS_{t-1}^- \varepsilon_{t-1} + \beta' z_t^* + e_t \tag{15.26}$$

$$v_t^2 = a + bS_{t-1}^+ \varepsilon_{t-1} + \beta' z_t^* + e_t, \tag{15.27}$$

where $v_t^2 = \left(\varepsilon_t/\sqrt{h_t}\right)^2$, $S_{t-1}^- = 1$ if $\varepsilon_{t-1} < 0$ and $S_{t-1}^- = 0$ otherwise, and $S_{t-1}^+ = 1 - S_{t-1}^-$. $z_t^* = \widetilde{h}(\theta)/h_t^*$, where $\widetilde{h}(\theta) = \partial h_t/\partial\theta$ evaluated at the values of maximum likelihood estimates of parameter θ, and h_t^* is the estimated conditional variance by a model considered.

References

Abel, A. B. (1988). Stock prices under time-varying dividend risk: An exact solution in an infinite-horizon general equilibrium model. *Journal of Monetary Economics, 22*, 373–393.

Ang, A., Hodrick, R., Xing, Y., & Zhang, X. (2006). The cross-section of volatility and expected returns. *Journal of Finance, 61*, 259–299.

Backus, D., & Gregory, A. W. (1993). Theoretical relations between risk premiums and conditional variances. *Journal of Business and Economic Statistics, 11*, 177–185.

Baillie, R., & DeGennaro, R. P. (1990). Stock returns and volatility. *Journal of Financial and Quantitative Analysis, 25*, 203–214.

Bali, T. G. (2008). The intertemporal relation between expected returns and risk. *Journal of Financial Economics, 87*, 101–131.

Ball, R., & Kothari, S. (1989). Nonstationary expected returns: Implications for tests of market efficiency and serial correlation in returns. *Journal of Financial Economics, 25*, 51–74.

Barky, R. B. (1989). Why don't the prices of stocks and bonds move together? *American Economic Review, 79*, 1132–1145.

Bollerslev, T., & Wooldridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time-varying covariances. *Econometric Review, 11*, 143–172.

Boudoukh, J., Richardson, M., & Whitelaw, R. F. (1997). Nonlinearities in the relation between the equity risk premium and the term structure. *Management Science, 43*, 371–385.

Brandt, W., & Kang, Q. (2004). On the relationship between the conditional mean and volatility of stock returns: A latent VAR approach. *Journal of Financial Economics, 72*, 217–257.

Breeden, D. T. (1979). An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics, 7*, 265–296.

Breen, W., Glosten, L., & Jagannathan, R. (1989). Economic significance of predictable variations in stock index returns. *Journal of Finance, 44*, 1177–1189.

Campbell, J. Y. (1987). Stock returns and the term structure. *Journal Financial Economics, 18*, 373–399.

Campbell, J. Y., & Hentschel, L. (1992). No news is good news: An asymmetric model of changing volatility in stock returns. *Journal Financial Economics, 31*, 281–318.

Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (1997). *The econometrics of financial markets*. Princeton: Princeton University Press.

Cecchetti, S. G., Lam, P., & Mark, N. C. (1990). Mean reversion in equilibrium asset prices. *American Economic Review, 80*, 398–418.

Darrat, A. F., Gilley, O. W., Li, B., & Wu, Y. (2011). Revisiting the risk/return relations in the Asian pacific markets: New evidence from alternative models. *Journal of Business Research, 64*, 199–206.

Engle, R. F., & Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *Journal of Finance, 48*, 1749–1778.

Evans, M. D. (1994). Expected returns, time-varying risk, and risk premia. *Journal of Finance, 49*, 655–679.

Fama, E. F., & French, K. R. (1988). Permanent and temporary components of stock prices. *Journal of Political Economy, 96*, 246–273.

Fama, E. F., & French, K. R. (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics, 25*, 23–49.

Ferson, W., & Harvey, C. (1991). The variation of economic risk premiums. *Journal of Political Economy, 99*, 385–415.

Fornari, F., & Mele, A. (1997). Sign- and volatility-switching ARCH models: Theory and applications to international stock markets. *Journal of Applied Econometrics, 12*, 49–65.

French, K. R., Schwert, G. W., & Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of Financial Economics, 19*, 13–29.

Friedman, M. (1957). *A theory of the consumption function*. Princeton: Princeton University Press.

Ghysel, E., Santa-Clara, P., & Valkanov, R. (2005). There is a risk-return trade-off after all. *Journal of Financial Economics, 76*, 509–548.

Glosten, L., Jagannathan, R., & Runkle, D. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance, 48*, 1779–1801.

Gonzalez-Rivera, G. (1998). Smooth transition GARCH models. *Studies in Nonlinear Dynamics and Econometrics, 3*, 61–78.

Gourieroux, C., & Monfort, A. (1997). *Time series and dynamic models*. Cambridge, UK: Cambridge University Press.

Hamilton, J. D. (1994). *Time series analysis*. Princeton: Princeton University Press.

Hamilton, J. D., & Susmel, R. (1994). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics, 64*, 307–333.
Harvey, C. R. (1989). Time-varying conditional covariances in tests of asset pricing models. *Journal of Financial Economics, 24*, 289–317.
Harvey, A. C. (1993). *The econometric analysis of time series*. Cambridge, MA: The MIT Press.
Harvey, C. R. (2001). The specification of conditional expectations. *Journal of Empirical Finance, 8*, 573–638.
Haugen, R. A., Talmor, E., & Torous, W. N. (1991). The effect of volatility changes on the level of stock prices and subsequent expected returns. *Journal of Finance, 46*, 985–1007.
Kim, C., Morley, J. C., & Nelson, C. R. (2001). Does an intertemporal tradeoff between risk and return explain mean reversion in stock price? *Journal of Empirical Finance, 8*, 403–426.
Lucas, R. E. (1978). Asset prices in an exchange economy. *Econometrica, 46*, 1429–1446.
Ludvigson, S. C., & Ng, S. (2007). The empirical risk-return relation: A factor analysis approach. *Journal of Financial Economics, 83*, 171–222.
Lütkepol, H. (1993). *Introduction to multiple time series analysis*. New York: Springer.
Merton, R. C. (1973). An intertemporal asset pricing model. *Econometrica, 41*, 867–888.
Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics, 8*, 323–361.
Modigliani, F., & Brumberg, R. (1955). Utility analysis and the consumption function. In K. Kurihara (Ed.), *Post Keynesian economics*. London: G. Allen.
Müller, G., Durand, R. B., & Maller, R. A. (2011). The risk–return tradeoff: A COGARCH analysis of Merton's hypothesis. *Journal of Empirical Finance, 18*, 306–320.
Nam, K., Pyun, C., & Avard, S. (2001). Asymmetric reverting behavior of short-horizon stock returns: An evidence of stock market overreaction. *Journal of Banking and Finance, 25*, 805–822.
Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica, 59*, 347–370.
Nyberga, H. (2012). Risk-return tradeoff in U.S. stock returns over the business cycle. *Journal of Financial and Quantitative Analysis, 47*, 137–158.
Pagan, A. R., & Hong, Y. S. (1989). Nonparametric estimation and the risk premium. In W.A. Barnett, J. Powell, & G. E. Tauchen (Eds.), *Semiparametric and nonparametric methods in econometrics and statistics* (pp. 51–75). Cambridge: Cambridge University Press.
Pastor, L., Sinha, M., & Swaminthan, B. (2008). Estimating the intertemporal risk-return tradeoff using the implied cost of capital. *Journal of Finance, 63*, 2859–2897.
Pindyck, R. (1984). Risk, inflation, and the stock market. *American Economic Review, 74*, 335–351.
Poterba, J. M., & Summers, L. H. (1986). The persistence of volatility and stock market fluctuations. *American Economic Review, 76*, 1142–1151.
Rothman, P. (1999). *Nonlinear time series analysis of economic and financial data*. Boston: Kluwer.
Scruggs, J. T. (1998). Resolving the puzzling intertemporal relation between the market risk premium and conditional market variance: A two-factor approach. *Journal of Finance, 52*, 575–603.
Turner, C. M., Startz, R., & Nelson, C. R. (1989). A Markov model of heteroskedasticity, risk, and learning in the stock market. *Journal of Financial Economics, 25*, 3–22.
Whitelaw, R. F. (1994). Time variations and covariations in the expectation and volatility of stock market returns. *Journal of Finance, 49*, 515–541.
Whitelaw, R. F. (2000). Stock market risk and return: An equilibrium approach. *Review of Financial Studies, 13*, 521–547.
Yu, J., & Yuan, Y. (2011). Investor sentiment and the mean–variance relation. *Journal of Financial Economics, 100*, 367–381.

Combinatorial Methods for Constructing Credit Risk Ratings

16

Alexander Kogan and Miguel A. Lejeune

Contents

A. Kogan (✉)
Rutgers Business School, Rutgers, The State University of New Jersey, Newark–New Brunswick, NJ, USA

Rutgers Center for Operations Research (RUTCOR), Piscataway, NJ, USA
e-mail: kogan@rutgers.edu

M.A. Lejeune
George Washington University, Washington, DC, USA
e-mail: mlejeune@gwu.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_16,

Abstract

This study uses a novel method, the Logical Analysis of Data (LAD), to reverse engineer and construct credit risk ratings which represent the creditworthiness of financial institutions and countries. LAD is a data mining method based on combinatorics, optimization, and Boolean logic that utilizes combinatorial search techniques to discover various combinations of attribute values that are characteristic of the positive or negative character of observations. The proposed methodology is applicable in the general case of inferring an objective rating system from archival data, given that the rated objects are characterized by vectors of attributes taking numerical or ordinal values. The proposed approaches are shown to generate transparent, consistent, self-contained, and predictive credit risk rating models, closely approximating the risk ratings provided by some of the major rating agencies. The scope of applicability of the proposed method extends beyond the rating problems discussed in this study and can be used in many other contexts where ratings are relevant.

We use multiple linear regression to derive the logical rating scores.

Keywords

Credit risk rating • Reverse engineering • Logical Analysis of Data • Combinatorial optimization • Data mining • Creditworthiness • Financial strength • Internal rating • Preorder • Logical rating score

16.1 Introduction

16.1.1 Importance of Credit Risk Ratings

Credit ratings published by such agencies as Moody's, Standard & Poor's, and Fitch are considered important indicators for financial markets, providing critical information about the likelihood of future default. The importance of credit ratings is recognized by international bodies, as manifested by the Basel Capital Accord (2001, 2006). The progressively increasing importance of credit risk ratings is driven by the dramatic expansion of investment opportunities associated with the globalization of the world economies. Since these opportunities are often risky, the internationalized financial markets have to rely on agencies' ratings for the assessment of credit risk.

The importance of credit risk rating systems is manifold, as shown below:

- **Credit approval**: the credit risk rating plays a major role in the credit preapproval decisions. Indeed, a binary decision model in which the credit risk rating of the obligor is an explanatory variable is typically used to preapprove the decision to grant or not the credit.
- **Pricing**: in case of loan preapproval, the credit risk rating impacts the conditions (interest rate, covenants, collaterals, etc.) under which the final credit is granted (Treacy and Carey 2000). The credit risk rating is also utilized in the operations subsequent to the preapproval stage.

- **Provisioning**: the credit risk rating is a variable used for calculating the expected losses, which, in turn, are used to determine the amount of economic capital that a bank must keep to hedge against possible defaults of its borrowers. The expected loss of a credit facility is a function of the probability of default of the borrower to which the credit line is granted, as well as the exposure at default and the loss given default associated with the credit facility. Since there is usually a mapping between credit risk rating and the borrower's probability of default, and provided that the rating of a borrower is a key predictor to assess the recovery rate associated with a credit facility granted to this borrower, the importance of the credit risk rating in calculating the amounts of economic capital is evident.
- **Moral hazard**: the reliance upon an objective and accurate credit risk rating system is a valuable tool against some moral hazard situations. Some financial institutions do not use a credit risk rating model but instead let lending officers assign credit ratings based on their judgment. The lending officers are in charge of the marketing of banking services, and their performance and therefore compensation are determined with respect to the "profitability" of the relationships between the bank and its customers. Clearly, the credit risk ratings assigned by the lending officers will affect the volume of the approved loans and the compensation of the officer who, as a result, could have an incentive to assign ratings in a way that is not consistent with the employer's interests. Thus, the use of a reliable credit risk rating system could lead to avoiding such perverse incentive situations.
- **Basel compliance**: the New Basel Capital Accord (Basel II) requires banks to implement a robust framework for the evaluation of credit risk exposures of financial institutions and the capital requirements they must bear. This involves the construction and the cross-validation of accurate and predictive credit risk rating systems.

Financial institutions, while taking into account external ratings (e.g., those provided by Fitch, Moody's, S&P's), have increasingly been developing efficient and refined internal rating systems over the past years. The reasons for this trend are multiple.

First, the work of rating agencies, providing external, public ratings, has recently come under intense scrutiny and criticism, justified partly by the fact that some of the largest financial collapses of the decade (Enron Corp, etc.) were not anticipated by the ratings. At his March 2006 testimony before the Senate Banking Committee, the president and CEO of the CFA Institute[1] highlighted the conflicts of interest in credit ratings agencies (Wall Street Letter 2006). He regretted that credit rating agencies "have been reluctant to embrace any type of regulation over the services they provide" and reported that credit rating agencies "should be held to the highest standards of transparency, disclosure and professional conduct. Instead, there are no standards." The problem is reinforced by the fact that rating agencies, charging fees

[1]Wall Street Letter. 2006. CFA To Senate: Follow Our Lead On Credit Rating.

to rated countries, can be suspected of reluctance to downgrade them, because of the possibility of jeopardizing their income sources. This is claimed, for example, by Tom McGuire, an executive vice-president of Moody's, who states that "the pressure from fee-paying issuers for higher ratings must always be in a delicate balance with the agencies' need to retain credibility among investors."[2] The necessity to please the payers of the ratings opens the door to many possible issues. Kunczik (2001) notes that the IMF fears the danger that "issuers and intermediaries could be encouraged to engage in *rating shopping* – a process in which the issuer searches for the least expensive and/or least demanding rating." The reader is referred to Hammer et al. (2006) for a discussion of some other criticisms (lack of comprehensibility, procyclicality, black box, lack of predictive and crisis-warning power, regional bias, etc.) commonly addressed to rating agencies.

Second, an internal rating system provides autonomy to a bank's management in defining credit risk in line with that bank's core business and best international practices. Internal ratings are also used to report to senior management various key metrics such as risk positions, loan loss reserves, economic capital allocation, and employee compensation (Treacy and Carey 2000).

Third, while the Basel Committee on Banking Supervision of the Bank for International Settlements favored originally (i.e., in the 1988 Capital Accord) the ratings provided by external credit ratings agencies, it is now encouraging the Internal Ratings-Based (IRB) approach under which banks use their own internal rating estimates to define and calculate default risk, on the condition that the robust regulatory standards are met and the internal rating system is validated by the national supervisory authorities. The committee has defined strict rules for credit risk models used by financial institutions and requires them to develop and cross-validate these models in order to comply with the Basel II standard (Basel Committee on Banking Supervision 2001, 2006).

Fourth, aside from the Basel II requirements, banks are developing internal risk models to make the evaluation of credit risk exposure more accurate and transparent. Credit policies and processes will be more efficient, and the quality of data will be improved. This is expected to translate into substantial savings on capital requirements. Today's eight cents out of every dollar that banks hold in capital reserves could be reduced in banks with conservative credit risk policies, resulting in higher profitability.

16.1.2 Contribution and Structure

The above discussion outlines the impact of the credit risk ratings, the possible problems of objectivity and transparency of external (i.e., provided by rating agencies) credit risk ratings, and the importance and need for financial institutions to develop their own, internal credit risk rating systems.

[2] The Economist, July 15, 1995, 62.

In this chapter, we use the novel combinatorial pattern extraction method called the logical analysis of data (LAD) to learn a given credit risk rating system and to develop on its basis a rating system having the following characteristics:

- **Self-containment**: the rating system does not use as predictor variables any other credit risk rating information (non-recursiveness). Clearly, this requirement precludes the use of lagged ratings as independent variables. It is important to note that this approach is in marked contrast with that of the current literature (see Hammer et al. 2006, 2011 for a discussion). The significant advantage of the non-recursive nature of the rating system is its applicability to not-yet-rated obligors.
- **Objectivity**: the rating system only relies on measurable characteristics of the rated entities.
- **Transparency**: the rating system has formal explicit specification.
- **Accuracy**: it is in close agreement with the learned, opaque rating system.
- **Consistency**: the discrepancies between the learned rating system and the constructed one are resolved by subsequent changes in the learned rating system.
- **Generalizability**: applicability of the rating system to evaluate the creditworthiness of obligors at subsequent years or of obligors that were not previously rated.
- **Basel compliance**: it satisfies the Basel II Accord requirements (cross-validation, etc.).

In this study, we derive a rating system for two types of obligors:

- **Countries**: Eliasson (2002) defines **country risk** as the "risk of national governments defaulting on their obligations," while Afonso et al. (2007) state that "sovereign credit ratings are a condensed assessment of a government's ability and willingness to repay its public debt both in principal and in interests on time." Haque et al. (1996) define **country credit risk ratings** compiled by commercial sources as an attempt "to estimate country-specific risks, particularly the probability that a country will default on its debt-servicing obligations."
- **Financial institutions**: bank financial strength ratings represent the "bank's intrinsic safety and soundness" (Moody's 2006).

Two different approaches developed in this chapter for reverse engineering and constructing credit risk ratings will be based on the following:

- **Absolute creditworthiness**, which evaluates the riskiness of individual obligors.
- **Relative creditworthiness**, which first evaluates the comparative riskiness of pairs of obligors; the absolute riskiness of entities is then derived from their relative riskiness using the combinatorial techniques of partially ordered sets.

As was noted in the literature (de Servigny and Renault 2004), banks and other financial service organizations are not fully utilizing the opportunities provided by the significant increase in the availability of financial data. More specifically, the area of credit rating and scoring is lacking in up-to-date methodological advances (Galindo and Tamayo 2000; de Servigny and Renault 2004; Huang et al. 2004). This provides a tremendous opportunity for the application of modern data mining and machine learning techniques, based on statistics (Jain et al. 2000) and combinatorial pattern extraction (Hammer 1986). The above described approaches are implemented using the LAD combinatorial pattern extraction method, and it turns out to be a very conclusive case for the application and the contribution of data mining to the credit risk industry.

This chapter is structured as follows. Section 16.2 provides an overview of LAD (Hammer 1986), which is used to develop a new methodology for reverse engineering and rating banks and countries with respect to their creditworthiness. Section 16.3 details the absolute creditworthiness model constructed for rating financial institutions and describes the obtained results. Section 16.4 is devoted to the description of two relative creditworthiness models for rating countries and contains an extensive description of the results. Section 16.5 provides concluding remarks.

16.2 Logical Analysis of Data: An Overview

The *logical analysis of data* (LAD) is a modern data mining methodology based on combinatorics, optimization, and Boolean logic[3]. LAD can be applied for the analysis and classification of archives containing both binary (Hammer 1986; Crama et al. 1988) and numerical (Boros et al. 1997) data. The novelty of LAD consists in utilizing combinatorial search techniques to discover various combinations of attribute values that are characteristic of the *positive* or *negative* character of observations (such as whether a bank is solvent or not or whether a patient is healthy or sick). Then LAD selects (usually small) subsets of such combinations (usually optimizing a certain quality objective) to construct what is called a model (Boros et al. 2000). We briefly describe below the basic concepts of LAD, referring the reader for a more detailed description to Boros et al. (2000) and Alexe et al. (2007).

Observations in archives analyzed by LAD are represented by n-dimensional real-valued vectors which are called positive or negative based on the value of the additional binary (0,1) attribute called the *outcome* or the *class* of the observation. Consider a dataset as a collection of M points (a^j, z^j) where the outcome z^j of observation j has value 1 for a positive outcome and 0 for a negative one and a^j is an n-dimensional vector. Figure 16.1 illustrates a dataset containing five observations described by three variables. Each component $a^j[i]$ of the $[5 \times 3]$-matrix in

Observation	Variables			Outcome
j	$a[1]$	$a[2]$	$a[3]$	z^j
1	3.5	3.8	2.8	1
2	2.6	1.6	5	1
3	1	2.2	3.7	1
4	3.5	1.4	3.9	0
5	2.3	2.1	1	0

Fig. 16.1 Set of observations

[3]The presentation in this section is partially based on Hammer et al. (2006).

Variables	Observations	a[1]			a[2]		a[3]			Outcome
Cutpoints		$c_{1,1}$	$c_{1,2}$	$c_{1,3}$	$c_{2,1}$	$c_{2,2}$	$c_{3,1}$	$c_{3,2}$	$c_{3,3}$	
		3	2.4	1.5	3	2	4	3	2	
Binary Variables		$y^j_{i,k}$								z^j
	1	1	1	1	1	1	0	0	1	1
	2	0	1	1	0	0	1	1	1	1
	3	0	0	0	0	1	0	1	1	1
	4	1	1	1	0	0	0	1	1	0
	5	0	0	1	0	1	0	0	0	0

Fig. 16.2 Binarized dataset

Fig. 16.1 gives the value taken by variable i in observation j. We will use $a[i]$ to denote the variables corresponding to the components of this dataset. The rightmost column provides the outcome of the observation.

LAD discriminates positive and negative observations by constructing a binary-valued function f depending on the n input variables, in such a way that it closely approximates the unknown actual discriminator. LAD constructs this function f as a weighed sum of combinatorial patterns.

In order to specify how such a function f is found, we first transform the original dataset into a *binarized dataset* in which the variables can only take the values 0 and 1. We shall achieve this goal by using *indicator variables* which show whether the values the variables take in a particular observation are "large" or "small"; more precisely, each indicator variable shows whether the value of a numerical variable does or does not exceed a specified level. This is achieved by defining, for each variable $a[i]$, a set of $K(i)$ values $\{c_{i,k} \mid k = 1,..., K(i)\}$, called *cut points* to which binary variables $\{y_{i,k} \mid k = 1,...,K(i)\}$ are associated. The values of these binary variables for each observation (z^j, a^j) are then defined as:

$$y^j_{i,k} = \begin{cases} 1 & \text{if } a^j[i] \geq c_{i,k} \\ 0 & \text{otherwise} \end{cases}$$

Figure 16.2 provides the binarized dataset corresponding to the data displayed in Fig. 16.1 and shows the values $c_{i,k}$ of the cut points k for each variable $a[i]$ and those of the binary variables $y^j_{i,k}$ associated with any cut point k of variable i in observation j. For example, $y^1_{1,1} = 1$ since $a^1[1 = 3.5$ is greater than $c_{1,1} = 3$.

Positive (*negative*) *patterns* are combinatorial rules obtained as conjunctions of binary variables and their negations, which, when translated to the original variables, constrain a subset of input variables to take values between identified upper and lower bounds, so that:

- All the pattern conditions are satisfied by a sufficiently high proportion of the positive (negative) observations in the dataset.

- At least one of the pattern conditions is violated by a sufficiently high proportion of the negative (positive) observations.

The number of variables the values of which are constrained in the definition of a pattern is called the *degree* of the pattern. The fraction of positive (negative) observations covered by a positive (negative) pattern is called the *prevalence* of it. The fraction of positive (negative) observations among those covered by a positive (negative) pattern is called the *homogeneity* of it. Considering the above example, the pattern

$$y_{1,3} = 0 \text{ and } y_{3,1} = 1$$

is a positive one of degree 2 covering one positive observation and no negative observation. Therefore, its prevalence is equal to 100/3 % and its homogeneity is equal to 100 %. In terms of the original variables, it imposes a strict upper bound (3) on the value of $a[1]$ and a strict lower bound (4) on the value of $a[3]$.

LAD starts its analysis of a dataset by generating the *pandect*, that is, the collection of all patterns in a dataset. Note that the pandect of a dataset of typically occurring dimension can contain exponentially large number of patterns, but many of these patterns are either subsumed by other patterns or similar to them. It is therefore important to impose a number of limitations on the set of patterns to be generated, by restricting their degrees (to low values), their prevalence (to high values), and their homogeneity (to high values); these bounds are known as LAD *control parameters*. The quality of patterns satisfying these conditions is usually much higher than that of patterns having high degrees, or low prevalence, or low homogeneity. Several algorithms have been developed for the efficient generation of large subsets of the pandect corresponding to reasonable values of the control parameters (Boros et al. 2000). The collections of patterns sufficient for classifying the observations in the dataset are called *models*. A model includes sufficiently many positive (negative) patterns to guarantee that each of the positive (negative) observations in the dataset is "covered" by (i.e., satisfies the conditions of) at least one of the positive (negative) patterns in the model. Good models tend to minimize the number of points in the dataset covered simultaneously by both positive and negative patterns in the model. LAD models can be constructed using the Datascope software (Alexe 2002).

LAD classifies observations on the basis of model's evaluation of them in the following way. An observation (contained in the given dataset or not), satisfying the conditions of some of the positive (negative) patterns in the model and not satisfying the conditions of any of the negative (positive) patterns in the model, is classified as *positive* (*negative*). To classify an observation that satisfies both positive and negative patterns in the model, LAD utilizes a discriminant that assigns specific weights to the patterns in the model (Boros et al. 2000). To define the simplest discriminant that assigns equal weights to all positive (negative) patterns, let p and q represent the number of positive and negative patterns in a model, and let h and k represent the numbers of those positive, respectively, negative patterns in the model which cover a new observation ω. Then the discriminant $\Delta(\omega)$ is calculated as

Table 16.1 Classification matrix

	Classification of observations		
Observation classes	Positive	Negative	Unclassified
Positive	a	c	e
Negative	b	d	f

$$\Delta(\omega) = h/p - k/q \tag{16.1}$$

and the corresponding classification is determined by the sign of this expression. LAD leaves *unclassified* any observation for which $\Delta(\omega) = 0$, since in this case either the model does not provide sufficient evidence or the evidence it provides is contradictory. Computational experience with real-life problems has shown that the number of unclassified observations is usually small. The results of classifying the observations in a dataset can be represented in the form of a *classification matrix* (Table 16.1).

Here, the percentage of positive (negative) observations that are correctly classified is represented by a (respectively d). The percentage of positive (negative) observations that are misclassified is represented by c (respectively b). The percentage of positive (negative) observations that remain unclassified is represented by e (respectively f). Clearly, $a + c + e = 100\ \%$ and $b + d + f = 100\ \%$. The quality of the classification is defined by

$$Q = \frac{1}{2}(a + d) + \frac{1}{4}(e + f) \tag{16.2}$$

16.3 Absolute Creditworthiness: Credit Risk Ratings of Financial Institutions

16.3.1 Problem Description

The capability of evaluating the credit quality of banks has become extremely important in the last 30 years given the increase in the number of bank failures: during the period from 1950 to 1980, bank failures averaged less than seven per year, whereas during the period from 1986 to 1991, they averaged 175 per year (Barr and Siems 1994).[4] Curry and Shibut (2000) report that the so-called savings and loan crisis cost around \$123.8 billion. Central banks are afraid of widespread bank failures since they could exacerbate cyclical recessions and result in more severe financial crises (Basel Committee on Banking Supervision 2004). More accurate credit risk models for banks could enable the identification of problematic banks early, which is seen as a necessary condition by the Bank for International Settlements (2004) to avoid failure, and could serve the regulators in their efforts to minimize bailout costs.

[4]The presentation in this section is based on Hammer et al. (2012).

The evaluation and rating of the creditworthiness of banks and other financial organizations is particularly challenging, since banks and insurance companies appear to be more opaque than firms operating in other industrial sectors. Morgan (2002) attributes this to the fact that banks hold certain assets (loans, trading assets, etc.), the risks of which change fast and are very difficult to assess, and it is further compounded by banks' high leverage. Therefore, it is not surprising that the main rating agencies (Moody's and S&P's) disagree much more often about the ratings given to banks than about those given to obligors in other sectors. The difficulty of accurately rating those organizations is also due the fact that the rating migration volatility of banks is historically significantly higher than it is for corporations and countries and that banks tend to have higher default rates than corporations (de Servigny and Renault 2004). Another distinguishing characteristic of the banking sector is the external support (i.e., from governments) that banks receive and other corporate sectors do not (Fitch Rating 2006). A thorough review of the literature pertaining to the rating and evaluation of credit risk of financial institutions can be found in Hammer et al. (2012).

In the next sections, we shall:

- Identify a set of variables that provides sufficient information to accurately replicate the Fitch bank ratings.
- Construct LAD patterns to discriminate between banks with high and low ratings.
- Construct an optimized model utilizing (some of) the LAD patterns which is capable of distinguishing between banks with high and low ratings.
- Define an accurate bank rating system on the basis of the discriminant values provided by the constructed model.
- Cross-validate the proposed rating system.

16.3.2 Data

16.3.2.1 External Credit Risk Ratings of Financial Institutions

This section starts with a brief description of the Fitch individual bank rating system. Long- and short-term credit ratings provided by Fitch constitute an opinion on the ability of an entity to meet financial commitments (interest, preferred dividends, or repayment of principal) on a timely basis (Fitch Ratings 2001). These ratings are comparable worldwide and are assigned to countries and corporations, including banks.

Fitch bank ratings include *individual* and *support* ratings. Support ratings comprise five rating categories which reflect the likelihood that a banking institution will receive support either from the owners or the governmental authorities if it runs into difficulties. The availability of support, though critical, does not reflect completely the likelihood that a bank will remain solvable in case of adverse situations. To complement a support rating, Fitch also provides an individual bank rating to evaluate credit quality separately from any consideration of outside support. This rating is commonly viewed as assessing a bank were it entirely independent and could not rely on external support. It supposedly takes into

Table 16.2 Fitch individual rating system (Fitch Ratings 2001)

Category	Numerical scale	Description
A	9	A very strong bank. Characteristics may include outstanding profitability and balance sheet integrity, management, or operating environment
B	7	A strong bank. There are no major concerns regarding the bank
C	5	An adequate bank which, however, possesses one or more troublesome aspects. There may be some concerns regarding its profitability, balance sheet integrity, management, operating environment or prospects
D	3	A bank which has weaknesses of internal and/or external origin. There are concerns regarding its profitability, management, balance sheet integrity, franchise, operating environment or prospects
E	1	A bank with very serious problems which either requires or is likely to require external support

consideration such factors as profitability and balance sheet integrity, franchise, management, operating environment, and prospects.

We present in Table 16.2 a detailed description of the nine rating categories characterizing the Fitch individual bank credit rating system. Since individual bank credit ratings are comparable across different countries, they will be used in the remaining part of this chapter. In addition, Fitch uses gradations among these five ratings: A/B, B/C, C/D, and D/E, the corresponding numerical values of which being, respectively, 8, 6, 4, and 2. This conversion of the Fitch individual bank ratings into a numerical scale is commonly used (see, e.g., Poon et al. 1999).

16.3.2.2 Variables and Observations

We use the following 14 financial variables (loans, other earning assets, total earning assets, nonearning assets, net interest revenue, customer and short-term funding, overheads, equity, net income, total liability and equity, operating income) and nine representative financial ratios as predictors in our model. The variables are as follows: ratio of equity to total assets (asset quality), net interest margin, ratio of interest income to average assets, ratio of other operating income to average assets, ratio of noninterest expenses to average assets, return on average assets (ROAA), return on average equity (ROAE), cost-to-income ratio (operations), and ratio of net loans to total assets (liquidity). The values of these variables were collected at the end of 2000 and come from the database Bankscope.

As an additional variable, we use in this study the S&P's risk rating of the country where the bank is located. The S&P's country risk rating scale comprises 22 different categories (from AAA to D). We convert these categorical ratings into a numerical scale, assigning the largest numerical value (21) to the countries with the highest rating (AAA). Similar numerical conversions of country risk ratings are also used by Ferri et al. (1999) and Sy (2004). Similarly, Bloomberg has also developed a standard cardinal scale for comparing Moody's, S&P's, and Fitch-BCA ratings.

Our dataset consists of 800 banks rated by Fitch and operating in 70 different countries (247 in Western Europe; 51 in Eastern Europe; 198 in Canada and the USA; 45 in developing Latin American countries; 47 in the Middle East; 6 in Oceania; 6 in Africa; 145 in developing Asian countries; 55 in Hong Kong, Japan, and Singapore).

16.3.3 An LAD Model for Bank Ratings

Our design of an objective and transparent bank rating system on the basis of the LAD methodology is guided by the properties of LAD as a classification system. Therefore, we define a classification problem associated to the bank rating problem, construct an LAD model for it, and then define a bank rating system rooted in this LAD model. We define as *positive observations* the banks which have been rated by Fitch as A, A/B, or B and as *negative observations* those whose Fitch rating is D, D/E, or E.

In the binarization process, cut points were introduced for the 19 of the 24 numerical variables shown in Table 16.3. Actually, the other five numerical variables were also binarized, but since it turned out that they were redundant, only the variables shown in Table 16.3 were retained for constructing the model. Table 16.3 provides all the cut points used in pattern and model construction. For example, two cut points (24.8 and 111.97) are used to binarize the numerical variable "profit before tax" (PbT), that is, two binary indicator variables replace PbT, one telling whether PbT exceeds 24.8 and the other telling whether PbT exceeds 111.97.

The first step of applying the LAD technique to the problem binarized with the help of these variable cut points was the identification of a collection of powerful patterns. One example of such a powerful negative pattern is (i) the country risk rating is strictly lower than A and (ii) the profits before tax are at most equal to €111.96 millions. One can see that these conditions describe a *negative pattern*, since none of the positive observations (i.e., banks rated A, A/B, or B) satisfy both of them, while no less than 69.11 % of the negative observations (i.e., those banks rated D, D/E, or E) do satisfy both conditions. This pattern has degree 2, prevalence 69.11 %, and homogeneity 100 %.

The model we have developed for bank ratings is very parsimonious, consisting of only 11 positive and 11 negative patterns, and is built on a *support set* of only 19 out of the 24 original variables. All the patterns in the model are of degree at most 3, have perfect homogeneity (100 %), and very substantial prevalence (averaging 50.9 % for the positive and 37.7 % for the negative patterns).

16.3.4 LAD Model Evaluation

16.3.4.1 Accuracy and Robustness of the LAD Model

The classification of the banks whose ratings are A, A/B, B, D, D/E, or E with the above LAD model is 100 % accurate. We use ten two-folding experiments to cross-validate the model. In each of those experiments, we randomly assign the

Table 16.3 Cut points

Numerical variables	Cut points	Numerical variables	Cut points	Numerical variables	Cut points
Country risk rating	11, 16, 19.5, 20	Operating income (memo)	1,360.74	Noninterest expenses/ average assets	2.00, 2.77, 3.71, 4.93
Other earning assets	3,809, 9,564.9	Profit before tax	24.8, 111.97	Return on average assets	0.30, 0.80, 1.52
Nonearning assets	364	Equity/total assets	4.90	Return on average equity	11.82, 15.85, 19.23
Total assets	5,735.8	Equity	370.45	Overhead	127, 324
Customer & short-term funding	4,151.3, 9,643.8	Net interest margin	1.87	Net loans/total assets	44.95, 53.90, 59.67, 66.50
Cost-to-income ratio	50.76, 71.92	Net interest revenue/average assets	3.02		
Other operating income	46.8, 155.5, 470.5	Other operating income/ average assets	0.83, 1.28, 1.86, 3.10		

observations to a training and a testing sets of equal size. We use the training set to build the model, and we apply it to classify the observations in the testing set. In the second half of the experiment, we reverse the role of the two sets. The average accuracy and the standard deviation of the accuracy are, respectively, equal to 95.12 % and 0.03. These results highlight the predictive capability and the robustness of the derived LAD model.

We have also computed the correlation between the discriminant values of the LAD model and the bank credit risk ratings (represented on their numerical scale). Despite the fact that the correlation measure takes into account all the banks in the dataset (i.e., not only those which were used in creating the LAD model but also those rated B/C, C, or C/D, which were not used in the learning process), the correlation is very high, equal to 80.70 %, attesting to the strong predictive power of the LAD model. The ten two-folding experiments described above were then used to verify the stability of the correlation between the LAD discriminant values and the bank ratings. The average correlation is equal to 80.04 %, with a standard deviation of 0.04, another testimony of the stability of the close positive association between the LAD discriminant values and the bank ratings.

16.3.4.2 From LAD Discriminant Values to Ratings

The objective of this section is to map the numerical values of the LAD discriminant to the nine bank rating categories (A, A/B, ..., E). This will be accomplished using a nonlinear optimization problem that partitions the interval of the discriminant values into nine subintervals that we associate to the nine rating categories. The partitioning is determined through *cut points* x_i such that $-1 = x_0 \leq x_1 \leq x_2 \leq \ldots\ldots\ldots \leq x_8 \leq x_9 = 1$, where i indexes the rating categories (with one corresponding to E and nine corresponding to A). A bank should be rated i if its discriminant value is in the interval $[x_i, x_{i+1}]$. Since such a perfect partitioning may not exist, we replace the LAD discriminant values d_i by an *adjusted discriminant value* δ_i and find values of δ_i for which such a partitioning exists and which are "as close as possible" to the values d_i. The expression "as close as possible" involves the minimization of the mean square approximation error. Referring to the rating category of bank i by $j(i)$ and to the set of banks by N, we solve the convex nonlinear problem

$$
\begin{aligned}
\text{minimize} \quad & \sum_{i \in N} (\delta_i - d_i)^2 \\
\text{subject to} \quad & \delta_i \leq x_{j(i)+1}, i \in N \\
& x_{j(i)} < \delta_i, i \in N \\
& -1 = x_0 \leq x_1 \leq x_2, \ldots, \leq x_j, \ldots, x_8 \leq x_9 = 1 \\
& -1 \leq \delta_i \leq 1, i \in N
\end{aligned}
\tag{16.3}
$$

on the NEOS server (Czyzyk et al. 1998) with the solver Lancelot to determine the cut points x_j and the adjusted discriminant values δ_i that will be used for rating the banks not only in the training sample but also those which are not. In this case, the bank rating is defined by the particular subinterval containing the LAD discriminant value.

Table 16.4 Discrepancy analysis

	q_k		
k	$N = \{A, A/B, B, D, D/E, E\}$	$N = \{B/C, C, C/D\}$	$N = \{A, A/B, B, B/C, C, C/D, D, D/E, E\}$
0	29.60 %	29.36 %	29.50 %
1	53.28 %	47.71 %	51.00 %
2	12.05 %	18.96 %	14.88 %
3	4.23 %	3.36 %	3.88 %
4	0.85 %	0.61 %	0.75 %
5–6–7–8	0.00 %	0.00 %	0.00 %

As compared to ordered logistic regression, the LAD-based approach has the additional advantage that it can be used for generating any desired number of rating categories. The LAD model can take the form of a binary classification model and can be used to preapprove or not a loan to a bank. The LAD rating approach can also be used to derive models with higher granularity (i.e., more than nine rating categories), which are used by banks to further differentiate their customers and to tailor accordingly their credit pricing policies.

16.3.4.3 Conformity of Fitch and LAD Bank Ratings

In this section, the goal is to investigate to goodness of fit of our rating system and observe to which extent the original LAD discriminant values fit in the identified rating subintervals. We denote by q_k ($k = 0, \ldots, 8$) the fraction of banks whose rating category determined in this way differs from its Fitch rating by exactly k categories. It is worth recalling that the rating cut points were derived using all the banks in the sample but that the LAD discriminant values were obtained by only taking into account the banks rated A, A/B, B, D, D/E, and E. This is why we calculate separately the discrepancy counts for the banks rated A, A/B, B, D, D/E, and E and for these rated B/C, C, and C/D.

Table 16.4 highlights the high goodness of fit of the proposed LAD model. More than 95 % of the banks are rated within two categories of their Fitch rating, with about 30 % of the banks receiving exactly the same rating as in the Fitch rating system and another 51 % being off by exactly one rating category. The very high concordance between the LAD and the Fitch ratings is illustrated by the weighted average distances between the two ratings equal to (i) 0.93 for the categories A, A/B, B, D, D/E, and E; (ii) 0.98 for the categories B/C, C, and C/D; and (iii) 0.95 for all banks in the dataset. The stability of the proposed rating system and its suitability to evaluate the creditworthiness of "new" banks, that is, banks which are not rated by agencies or banks the rater has not dealt with before, are magnified by the fact that the goodness of fit of the ratings of the banks not used in deriving the LAD model is very close to the goodness of fit for the banks used for deriving the LAD model (i.e., those rated A, A/B, B, D, D/E, and E).

In order to appraise the robustness of the proposed rating system, we apply ten times the two-folding procedure described above to derive the average ($\overline{q}_k$)

Table 16.5 Cross-validated discrepancy analysis

	k								
	0	1	2	3	4	5	6	7	8
$\overline{q}_k$	29.01 %	51.18 %	14.10 %	4.36 %	1.14 %	0.20 %	0.01 %	0.00 %	0.00 %

discrepancy counts (Table 16.5) of the bank ratings predicted for the testing sets. The fact that on the average the difference between the Fitch and the LAD ratings is only 0.98 is a very strong indicator of the LAD model's stability and the absence of overfitting.

16.3.4.4 Importance of Variables

While the focus in LAD is on discovering how the interactions between the values of small groups of variables (as expressed in patterns) affect the outcome (i.e., the bank ratings), one can also use the LAD model to learn about the importance of individual variables. A natural measure of importance of a variable in an LAD model is the frequency of its appearance in the model's patterns. The three most important variables in the 22 patterns constituting the LAD model are the credit risk rating of the country where the bank is located, the return on average total assets, and the return on average equity. The importance of the country risk rating variable, which appears in 18 of the 22 patterns, can be explained by the fact that credit rating agencies are reluctant to give an entity a better credit risk rating than that of the country where it is located. Both the return on average assets and the return on average equity variables appear in six patterns. These two ratios, respectively, representing the efficiency of assets in generating profits and that of shareholders' equity in generating profits, are critical indicators of a company's prosperity and are presented by Sarkar and Sriram (2001) as key predictors to assess the wealth of a bank. The return on average equity is also found significant for predicting the rating of US banks (Huang et al. 2004).

16.3.5 Remarks on Reverse Engineering Bank Ratings

The results presented above demonstrate that the LAD approach can be used to derive rating models with varying granularity levels:

- A binary classification model to be used for the preapproval operations
- A model with the same granularity as the benchmarked rating model
- A model with higher discrimination power, that is, with higher granularity than that of the benchmarked rating model, to allow the bank to refine its pricing policies and the allocation of regulatory capital

We show that the LAD model cross-validates extremely well and therefore is highly generalizable and could therefore be used by financial institutions to develop internal, Basel-compliant rating models.

The availability and processing of data have been major obstacles in the way of using credit risk rating models. Until recently, many banks did not maintain such datasets and were heavily dependent on qualitative judgments. It is only after the currency crises of the 1990s and the requirements imposed by the Basel Accord that financial institutions have seen an incentive in collecting the necessary data and maintaining the databases.

The move towards a heavier reliance on rating models is based on the assumption that models produce more consistent ratings and that, over the long haul, operating costs will diminish since less labor will be required to produce ratings. The model proposed in this chapter will reinforce the incentives to develop and rely upon credit risk models in bank operations due to the following:

- The accuracy and predictive ability of the proposed model will guarantee dependable ratings.
- Its parsimony will alleviate the costs of extracting and maintaining large datasets.
- It will result in leaner loan approval operations and faster decisions and could thus reduce the overall operating costs.

16.4 Relative Creditworthiness: Country Risk Ratings

16.4.1 Problem Description

Country risk ratings have critical importance in the international financial markets since they are the primary determinants of the interest rates at which countries can obtain credit.[5] There are numerous examples of countries having to pay higher rates on their borrowing following their rating downgrade, an often cited example being Japan. As mentioned above, another critical aspect of country risk ratings concerns their influence on the ratings of national banks and companies that would make them more or less attractive to foreign investors. That is why the extant literature calls country risk ratings the "pivot of all other country's ratings" (Ferri et al. 1999) and considers them the credit risk ceiling for all obligors located in a country (Eliasson 2002; Mora 2006). Historical record shows the reluctance of raters to give a company a higher credit rating than that of the sovereign where the company operates. Contractual provisions sometimes prohibit institutional investors from investing in debt rated below a prescribed level. It has been demonstrated (Ferri et al. 1999) that variations in sovereign ratings drastically affect the ratings of banks operating in low-income countries, while the ratings of banks operating in high-income countries (Kaminsky and Schmukler 2002; Larrain et al. 1997) do not depend that much on country ratings. Banks, insurance companies, and public institutions frequently assess the amount of exposure they have in each country and establish lending limits taking into account the estimated level of country risk. Credit managers in multinational corporations have to assess evolving conditions in foreign countries in order to decide whether to request letters of credit for particular

[5]The presentation in this section is based on Hammer et al. (2006, 2012).

transactions. Country risk estimates and their updates are utilized on a real-time basis by multinational corporations facing constant fluctuation in international currency values and difficulties associated with moving capital and profits across national boundaries. Financial institutions have to rely on accurate assessments of credit risk to comply with the requirements of the Basel Bank for International Settlements. Feeling the pressure to get higher ratings could lead to fraud attempts. A notorious case was Ukraine's attempt to obtain IMF credits through misleading reporting of its reserve data on foreign exchanges.

The existing literature on country risk, defined by Bourke and Shanmugam (1990) as "the risk that a country will be unable to service its external debt due to an inability to generate sufficient foreign exchange," recognizes both financial/economic and political components of country risk. There are two basic approaches to the interpretation of the reasons for defaulting. The first one is the *debt-service capacity approach* which considers the deterioration of solvency of a country as preventing it from fulfilling its commitments. This approach views country risk as a function of various financial and economic country parameters. The second one is the *cost-benefit approach* which considers a default on commitments or a rescheduling of debt as a deliberate choice of the country. In this approach the country accepts possible long-term negative effects (e.g., the country's exclusion from certain capital markets (Reinhart 2002) as preferable to repayment). Being politically driven, this approach includes political country parameters in addition to the financial and economic ones in country risk modeling (Brewer and Rivoli 1990, 1997; Citron and Neckelburg 1987).

16.4.2 Data

16.4.2.1 Ratings

We analyze in this chapter Standard & Poor's foreign currency country ratings, as opposed to the ratings for local currency debt. The former is the more important problem, since the sovereign government has usually a lower capacity to repay external (as opposed to domestic) debt, and as an implication, the international bond market views foreign currency ratings as the decisive factor (Cantor and Packer 1996). This is manifested by much higher likelihood for international investors to acquire foreign currency obligations rather than domestic ones. In evaluating foreign currency ratings, one has to take into account not only the economic factors but also the country intervention risk, that is, the risk that a country imposes, for example, exchange controls or a debt moratorium. The evaluation of local currency ratings need not take into account this country intervention risk.

Table 16.16 lists the different country risk levels used by S&P's and also provides descriptions associated with these labels. A rating inferior to BB+ indicates that a country is non-investment grade (speculative). A rating of CCC+ or lower indicates that a country presents serious default risks. BB indicates the least degree of speculation and CC the highest. The addition of a plus or minus sign modifies the rating (between AA and CCC) to indicate relative standing within the major rating category. These subcategories are treated as separate ratings in our analysis.

16.4.2.2 Selected Variables

The selection of relevant variables is based on three criteria. The first criterion is the variable's significance (relevance) in assessing countries' creditworthiness. The review of the literature on predictors for country risk ratings (see Hammer et al. 2011) played an important role in defining the set of candidate variables to include in our model. The second criterion is the availability of complete and reliable statistics, to allow us to maintain the significance and the scope of our analysis. The third criterion is the uniformity of data across countries. This is why we decided against incorporating the unemployment rate statistics provided by the World Bank, since it is compiled according to different definitions: there are significant differences between countries in the treatment of temporarily laid off workers, those looking for their first job, and the criteria for being considered as unemployed.

After applying the criteria of relevance, availability, and uniformity described above, the following variables were incorporated in our model: *gross domestic product per capita* (*GDPc*), *inflation rate* (*IR*), *trade balance* (*TB*), *exports*' *growth rate* (*EGR*), *international reserves* (*RES*), *fiscal balance* (*FB*), *debt to GDP* (*DGDP*), *political stability* (*PS*), *government effectiveness* (*GE*), *corruption* (*COR*), *exchange rate* (*ER*), *and financial depth and efficiency* (*FDE*).

In addition to eight economic variables that have already been used in the country credit risk rating literature, a new one, called *financial depth and efficiency* (*FDE*), was added. This variable is measured as the ratio of the domestic credit provided by the banking sector to the GDP. It measures the growth of the banking system since it reflects the extent to which savings are financial. The financial depth and efficiency variable was first considered in Hammer et al. (2006, 2011) for the evaluation of country risk ratings. The reason to include this variable is that it captures some information that is relevant to the assessment of the creditworthiness of a country and that is not accounted for by other variables and in particular by the fiscal balance of a country.

Our analysis utilized the values of these nine economic/financial variables and three political variables taken at the end of 1998. Our dataset included the values of these 12 variables for the 69 countries considered: 24 industrialized countries, 11 Eastern European countries, eight Asian countries, ten Middle Eastern countries, 15 Latin American countries, and South Africa. The dependent variable is the S&P's country risk ratings for these countries at the end of 1998. The sources utilized for compiling the values of the economic/financial variables include the International Monetary Fund (World Economic Outlook Database), the World Bank (World Development Indicators database), and, for the ratio of debt to gross domestic product, Moody's publications. Values of political variables are taken from Kaufmann et al. (1999a, b), whose database is a joint product of the Macroeconomics and Growth, Development Research Group and Governance, Regulation and Finance Institutes affiliated with the World Bank.

16.4.3 Rating Methodologies

In the next subsections, we derive two new combinatorial models for country risk rating by reverse engineering and "learning" from past S&P's ratings. The models are developed using the novel combinatorial-logical technique of logical analysis of data which derives a new rating system only from the qualitative information representing pairwise comparisons of country riskiness. The approach is based on a *relative creditworthiness* concept, which posits that the knowledge of (pre)order of obligors with respect to their creditworthiness should be the sole source for creating a credit risk rating system. Stated differently, only the order relation between countries should determine the ratings. Thus, the inference of a model for the order relation between countries is the objective of this study. This is in perfect accordance with the general view of the credit risk rating industry. Altman and Rijken (2004) state that the objective of rating agencies is "to provide an accurate relative (i.e., ordinal) ranking of credit risk," which is confirmed by Fitch ratings (2006) saying that "Credit ratings express risk in relative rank order, which is to say they are ordinal measures of credit." Bhatia (2002) adds that: "Although ratings are measures of absolute creditworthiness, in practice, the ratings exercise is highly comparative in nature... On one level, the ratings task is one of continuously sorting the universe of rated sovereigns – assessed under one uniform set of criteria – to ensure that the resulting list of sovereigns presents a meaningful global order of credit standing. On another level, the sorting task is constrained by a parallel need to respect each sovereign's progression over time, such that shifting peer comparisons is a necessary condition – but not a sufficient one – for upward or downward ratings action."

A self-contained model of country risk ratings can hardly be developed by standard econometric methods since the dataset contains information *only* about 69 countries, each described by 12 explanatory variables. An alternative at hand possible with combinatorial techniques is to examine the *relative riskiness* of one country compared to another one, rather than modeling the riskiness of each individual country. This approach has the advantage of allowing the modeling to be based on a much richer dataset (2,346 pairs of countries), which consists of the comparative descriptions of all pairs of countries in the current dataset.

The models utilize the values of nine economic and three political variables associated to a country, but do not use directly or indirectly previous years' ratings. This is a very important feature, since the inclusion of information from past ratings (lagged ratings, rating history) does not allow the construction of a self-contained rating system and does not make possible to rate the creditworthiness of not-yet-rated countries. We refer to Hammer et al. (2006, 2011) for a more detailed discussion of the advantages of building a non-recursive country risk rating system. Moreover, the proposed LAD models completely eliminate the need to view the ratings as numbers. Section 16.4.3.1 presents the common features of the two developed LAD models, while Sects. 16.4.3.2 and 16.4.3.3 discuss the specifics of the LAD-based Condorcet ratings and the logical rating scores, respectively.

16.4.3.1 Commonalities

Pairwise Comparison of Countries: Pseudo-observations

Every country $i \in I = \{1, \ldots, 69\}$ in this study is described by the 13-dimensional vector C_i, whose first component is the country risk rating given by Standard & Poor's, while the remaining 12 components specify the values of the nine economic/financial and of the three political variables. A *pseudo-observation* P_{ij}, associated to every pair of countries i, $j \in I$, provides in a way specified below a comparative description of the two countries.

Every pseudo-observations is also described by a 13-dimensional vector. The first component is an indicator which takes the value 1 if the country i in the pseudo-observation P_{ij} has a higher rating than the country j, –1 if the country j has a higher rating than the country i, and 0 if the two countries have the same rating. The other components k, $k = 2, \ldots, 13$ of the pseudo-observation $P_{ij}[k]$ are derived by taking the differences of the corresponding components of C_i and C_j:

$$P_{ij}[k] = C_i[k] - C_j[k], k = 2, \ldots, 13 \tag{16.4}$$

Transformation (16.4) alleviates the problems related to the small size ($| I |$) of the original dataset by constructing a substantially larger dataset containing $| I |*(| I | -1)$ pseudo-observations. While the set of pseudo-observations is not independent, since $P_{hi} + P_{ij} = P_{hj}$, Hammer et al. (2006) show that this does not create any problems for the LAD-based combinatorial data analysis techniques.

We illustrate the construction of pseudo-observations with Japan and Canada. Rows 2 and 3 in Table 16.6 display the values of the 12 economic/financial and political variables, as well as the S&P's rating for Japan and Canada (at the end of December 1998), while rows 4 and 5 report the pseudo-observations $P_{\text{Japan, Canada}}$ and $P_{\text{Canada, Japan}}$ from the country observations C_{Japan} and C_{Canada}. Since Japan and Canada are rated, respectively, AAA and AA + by S&P's at the end of December 1998 and the rating AAA is better than rating AA+, the first component of the pseudo-observation vector is equal to 1. The set of pseudo-observations is antisymmetric.

An advantage of this transformation is that it allows us to avoid the problems posed by the fact that the original dataset contains only a small number ($| I |$) of observations. The transformation (16.4) provides a larger dataset containing $| I |*(| I | -1)$ pseudo-observations.

Construction of Relative Preferences

The LAD-based reverse engineering of Standard & Poor's rating system starts with deriving an LAD model (constructed as a weighed sum of patterns) from the archive of all those pseudo-observations P_{ij}, corresponding to pairs of countries i and j having different S&P's ratings. A model resulting from applying LAD to the 1998 dataset consists of 320 patterns. As an example, let us describe two of these patterns below. The positive pattern

$$FDE > 28.82; GDPc > 1539.135; GE > 0.553$$

Table 16.6 Examples of country and pseudo-observations

	S&P's rating	*FDE*	*RES*	*IR*	*TB*	*EGR*	*GDPc*	*ER*	*FB*	*DGDP*	*PS*	*GE*	*COR*
C_{Japan}	AAA	138.44	5.168	.65	21.7471	−2.54	24314.2	0.839	−7.7	0.47	1.153	0.839	0.724
C_{Canada}	AA+	94.69	1.01964	0.99	55.9177	8.79	24855.7	0.939	0.9	0.5	1.027	1.717	2.055
$P_{\text{Japan, Canada}}$	1	43.75	4.15	−0.34	−34.17	−11.33	−541.5	−0.1	−8.6	−0.03	0.126	−0.878	−1.331
$P_{\text{Canada, Japan}}$	−1	−43.75	−4.15	0.34	34.17	11.33	541.5	0.1	8.6	0.03	−0.126	0.878	1.331

can be interpreted in the following way. If country i is characterized by

(i) A financial depth and efficiency (FDE) exceeding that of country j by at least 28.82, and
(ii) A gross domestic product per capita ($GDPc$) exceeding that of country j by at least 1,539.135, and
(iii) A government efficiency (GE) exceeding that of country j by at least 0.553

then country i is perceived as more creditworthy than country j.

Similarly, the negative pattern

$$GDPc < -4,886.96; ER < 0.195; COR < -0.213$$

can be interpreted in the following way. If country j is characterized by

(i) A gross domestic product per capita ($GDPc$) exceeding that of country i by 4,886.96, and
(ii) An exports growth rate (EGR) exceeding that of country i by 0.195, and
(iii) A level of incorruptibility (CR) exceeding that of country i by 0.213

then country i is perceived as less creditworthy than country j.

The constructed LAD model allows us to compute the discriminant $\Delta(P_{ij})$ for each pseudo-observation $P_{ij}(i \neq j)$. We call the values $\Delta(P_{ij})$ of the discriminant the *relative preferences*. They can be interpreted as measuring how "superior" the country i's rating over that of country j. We call the [69 $\times$ 69]-dimensional antisymmetric matrix Δ, having the relative preferences as components, the *relative preference matrix*. Its components are relative preferences $\Delta(P_{ij})$ $(i \neq j)$ associated with every pair of countries, including those that have the same S&P's ratings, even though only those pseudo-observation P_{ij} for which i and j had different ratings were used in deriving the LAD model.

Classification of Pseudo-observations and Cross-Validation

The dataset used to derive the LAD model contains 4,360 pseudo-observations, with an equal number of positive and negative pseudo-observations, and is antisymmetric ($P_{ij} = -P_{ij}$). The results of applying the LAD model to classify the observations in the dataset are presented in Table 16.7. The overall classification quality of the LAD model (according to Eq. 16.2) is 95.425 %.

This very high classification accuracy may be misleading, since the constructed LAD model can be overfitting the data, that is, it is adapted excessively well to random noise in the training data and thus has an excellent accuracy on this training data, but could perform very poorly on new observations. We therefore utilize a statistical technique known as "jackknife" (Quenouille 1949) or "leave-one-out"

Table 16.7 Classification matrix

	Classified as			
	Positive	Negative	Unclassified	Total
Positive observations	93.90 %	3.72 %	2.38 %	100 %
Negative observations	3.72 %	93.90 %	2.38 %	100 %

Table 16.8 Classification matrix for Δ^{JK}

	Classified as			
	Positive	Negative	Unclassified	Total
Positive observations	93.49 %	3.67 %	2.84 %	100 %
Negative observations	3.67 %	93.49 %	2.84 %	100 %

to show that the achieved high classification accuracy is not due to overfitting. This technique removes from the dataset one observation at a time, learns a model from all the remaining observations, and evaluates the resulting model on the removed observation; then it repeats all these steps for each observation in the dataset. If on the average the predicted evaluations are "close to" the actual ones, then the model is not affected by overfitting.

In the case at hand, it would not be statistically sound to implement the jackknife technique in a straightforward way because of the dependencies among pseudo-observations (Hammer et al. 2006). Indeed, even after explicitly eliminating a single pseudo-observation P_{ij} from the dataset, it would still remain in the dataset implicitly, since $P_{ih} + P_{hj} + P_{ij}$. This problem can be resolved by modifying the above described procedure so that at each step, instead of just removing a single pseudo-observation P_{ij}, all the pseudo-observations which involve a particular country i are removed. Then, the LAD discriminant is derived on the basis of the remaining pseudo-observations and used to evaluate the relative preferences for every removed pseudo-observation, resulting in a row of relative preferences of all pseudo-observations P_{ij} which involve the country i. This modified procedure is repeated for every country in the dataset, and the obtained rows are combined into a matrix of relative preferences denoted by Δ^{JK}. The absence of overfitting is indicated by a very high correlation level of 96.48 % between the matrix Δ^{JK} and the original relative preference matrix Δ.

To further test for overfitting, we use the obtained matrix of relative preferences Δ^{JK} to classify the dataset of 4,360 pseudo-observations. The results of this classification are presented in Table 16.8, and the overall classification quality of the LAD model (according to formula (16.2)) is 95.10 %. The results in Table 16.8 are virtually identical to those shown in Table 16.7, thus proving the absence of overfitting.

Since the derived LAD model does not suffer from overfitting, it is tempting to interpret the signs of relative preferences as indicative of rating superiority and infer that a positive value for $\Delta(P_{i,j})$ indicates that country i is more creditworthy than country j, while a negative one for $\Delta(P_{i,j})$ justifies the opposite conclusion. However, this naïve approach towards relative rating superiority ignores the potential noise in the data and in relative preferences, which would make it difficult to transform classifications of pseudo-observations to a consistent ordering of countries by their creditworthiness. It is shown in Hammer et al. (2006) that the relationship based on the naïve interpretation of the relative preferences can violate the transitivity requirement of an order relation and therefore does not provide a consistent partially ordered set of countries.

The following section is devoted to overcoming this issue by relaxing the overly constrained search for (possibly nonexistent) country ratings whose pairwise orderings are in *precise* agreement with the signs of relative preferences. Instead, we utilize a more flexible search for *a partial order* on the set of countries, which satisfies the transitivity requirements and *approximates* well the set of relative preferences.

16.4.3.2 Condorcet Ratings

The ultimate objective of this section, which is based on the results of Hammer et al. (2006), is to derive a rating system from the LAD relative preferences. This is accomplished in two stages. First, we use the LAD relative preferences to define a partial order on the set of countries which represents their creditworthiness dominance relationship. Then, we extend the derived dominance relationship to two rating systems, respectively, based on the so-called weak Condorcet winners and losers. The number of rating categories in these rating systems is the same and will be determined by the structure of the partial order obtained in the first stage, that is, it is not a priory fixed.

From LAD Relative Preferences to a Partial Order on the Set of Countries

A partial order $\Pi(X)$ is a reflexive, antisymmetric, and transitive binary relation on a set X. Any two distinct elements x and y of X such that $(x, y) \in \Pi(X)$ are said to be comparable and denoted by $x \supset y$. If neither $(x, y) \in \Pi(X)$ nor $(y, x) \in \Pi(X)$, then x and y are called incomparable and denoted by $x \parallel y$.

As was mentioned above, the naïve rating superiority relation based on the LAD model is not transitive. This difficulty can be overcome by defining a strengthened version of the naïve rating superiority relation, to be called the *dominance relationship*. While the former only relies on the sign of the relative preference $\Delta(P_{ij})$, the definition of dominance of a country i over another country j takes into account not only the sign of the relative preference $\Delta(P_{ij})$ but also the values of the relative preferences of each of these two countries i and j over every other country k.

We will need the following notation to define the dominance relationship. Let

$$S_{ij}(k) = \Delta(P_{ik}) - \Delta(P_{jk}) \tag{16.5}$$

define the external preference of country i over country j with respect to country k, let

$$S_{ij} = \frac{\sum_{j \in I} S_{ij}(k)}{|I|} \tag{16.6}$$

define the *average external preference* of i over j, and let

$$\sigma_{ij} = \sqrt{\frac{\sum_{k \in C} \left[\left(\Delta(P_{ik}) - \Delta(P_{jk})\right) - S_{ij}\right]^2}{|I|}} \tag{16.7}$$

define the *standard deviation of the external preference of i over j*.

The dominance relationship of a country i over another country j will be defined by two conditions. The first one requires that $\Delta(P_{ij}) > 0$. It might also seem logical to require that $S_{ij}(k) > 0$ for every country k, $k \neq i, j$. However, this condition is so difficult to satisfy that the resulting partially ordered set is extremely sparse, with very few country pairs that are comparable. Thus, a more relaxed second condition will require that, only at a certain confidence level, the external preference of i over j should be positive. We parameterize the level of confidence by the multiplier η of the standard deviation σ_{ij}. More formally, for a given $\eta > 0$,

- A country i is said to dominate another country j if

$$\begin{cases} \Delta(P_{ij}) > 0 \\ S_{ij} - \eta\sigma_{ij} > 0 \end{cases} \tag{16.8}$$

- A country i is said to be dominated by another country j if

$$\begin{cases} \Delta(P_{ij}) < 0 \\ S_{ij} + \eta\sigma_{ij} < 0 \end{cases} \tag{16.9}$$

- In all the other cases, countries i and j are said to be not comparable; this can be due to the lack of evidence, or to conflicting evidence about the dominance of i over j.

Note that the larger the value of η is, the stronger the conditions are, and the fewer pairs of countries are comparable. If η is sufficiently large, then the dominance relationship is transitive and is a partial order, since it becomes progressively sparser until being reduced to empty. On the other hand, if η is small (and in particular for $\eta = 0$), then the dominance relationship becomes denser, but is not necessarily transitive.

We are looking for a "rich" dominance relationship (applying to as many country pairs as possible) which is transitive. Formally, our objective is to maximize the number of comparable country pairs subject to preserving the transitivity of the dominance relationship.

The richest dominance relationship defined by the two conditions (16.8) and (16.9) in the case of $\eta = 0$ will be called the *base dominance relationship*. If i and j are any two countries comparable in the base dominance relationship, then it is possible to calculate the smallest value of the parameter $\eta_{ij} = |\, S_{ij}/\sigma_{ij} \,|$ such that the i and j are not comparable in any dominance relationship defined by a parameter

Table 16.9 Correlation levels between LAD and canonical relative preferences

	Δ	$d^{S\&P}$	d^{M}	d^{II}	d^{LRS}
Δ	100 %	93.21 %	92.89 %	91.82 %	97.57 %
$d^{S\&P}$	93.21 %	100 %	98.01 %	96.18 %	95.54 %
d^{M}	92.89 %	98.01 %	100 %	96.31 %	95.20 %
d^{II}	91.82 %	96.18 %	96.31 %	100 %	94.11 %
d^{LRS}	97.57 %	95.54 %	95.20 %	94.11 %	100 %

value exceeding or equal to η_{ij}. This calculation is based on the fact that given a value of η, it is possible to check in polynomial time (Tarjan 1972) whether the corresponding dominance relationship is transitively closed. Then the algorithm that determines in polynomial time the minimum value η^* for which the corresponding dominance relationship is still transitive, sorts at most $|I|^2$ numbers η_{ij} in ascending order, and then checks one by one the transitivity of the corresponding dominance relationships. When the dominance relationship becomes transitive for the first time, the algorithm stops and outputs η^* equal to the corresponding value of the parameter η_{ij}. This study utilizes this value η^* and the corresponding dominance relationship between countries, called here the *logical dominance relationship* and denoted by the subscript of LAD (e.g., $\underset{LAD}{\succ}$).

The definition of dominance relationship between countries on the basis of average external preferences bears some similarities to the so-called “column sum methods” (Choo and Wendley 2004) utilized in reconciling inconsistencies resulting from the application of pairwise comparison matrix methods.

Extending Partially Ordered Sets to “Extreme” Linear Preorders

The information about country preferences contained in the economic and political attributes is represented most faithfully by the logical dominance relationship defined above. However, it is impractical to use, since this partial order requires a large amount of data to describe. On the other hand, country preferences can be expressed very compactly by country ratings, since the latter are a very special type of partial orders called linear preorders.

A partial order $\Pi(X)$ is called a *linear preorder* if there exists a mapping $M : X \to \{0,1, \ldots, k\}$ such that $x \supset y$ if and only if $M(x) > M(y)$. Therefore, a linear preorder is completely described by specifying its mapping M. Without loss of generality, one can assume that for every $i \in \{0,1, \ldots, k\}$, there exists $x \in X$ such that $M(x) = i$. Such a linear preorder is said to have $k + 1$ levels.

To make logical dominance of countries practically utilizable, this relationship should be transformed into a linear preorder preserving all the order relations between countries (i.e., is an *extension* of the partial order) and is as close as possible to it. The logical dominance relationship can be extended in a multitude of ways to a variety of linear preorders. In particular, two extreme linear preorders are constructed below that we call the *optimistic* and the *pessimistic* extensions, denoted by OE and PE, respectively. The names are justified since the former assigns to each country the highest level it can expect, while the latter assigns to each country the lowest level it can expect:

- In the first step of *OE* construction, those countries that are not dominated by any other country are assigned the highest level and are then removed from the set of countries under consideration. Iteratively, non-dominated countries in the remaining set of countries are assigned the highest remaining level until every country is assigned a level denoted by OE_i.
- In the first step of *PE* construction, those countries that do not dominate any other country are assigned the lowest level and are then removed from the set of countries under consideration. Iteratively, non-dominating countries in the

remaining set of countries are assigned the lowest remaining level until every country is assigned a level denoted by PE_i.

The method utilized above to construct OE and PE is known as the *Condorcet method*. It represents a specific type of voting system (Gehrlein and Lepelley 1998), and it is often used to determine the winner of an election. The Condorcet winner(s) of an election is generally defined as the candidate(s) who, when compared in turn with every other candidate, is preferred over each of them. Given an election with preferential votes, one can determine *weak Condorcet winners* (Ng et al. 1996) by constructing the *Schwartz set* as the union of all possible candidates such that (i) every candidate inside the set is pairwise unbeatable by any other candidate outside the set (ties are allowed) and (ii) no proper subset of the set satisfies the first property.

The Schwartz set consists exactly of all weak Condorcet winners. The *weak Condorcet losers* are the reverse of the weak Condorcet winners, that is, those losing pairwise to every other candidate. OE assigns the highest level to those countries that are the weak Condorcet winners, that is, better than or incomparable with every other country. PE assigns the lowest level to those countries that are the weak Condorcet losers, that is, worse than or are incomparable with every other country. Note that both OE and PE have the minimum possible number of levels. Indeed, in a directed graph whose vertices are the countries and whose arcs represent comparable countries in the dominance relationship, the length of the longest directed path bounds from below the number of levels in any linear preorder extending the dominance relationship. This length equals the number of levels in OE and PE.

16.4.3.3 Logical Rating Scores

The LAD relative preferences can be utilized in a completely different way (compared to OE or PE) to construct country risk ratings. We describe here how to derive new numerical ratings of all countries, called "logical rating scores" (LRS), by applying multiple linear regression. LRS were defined by Hammer et al. (2011) as numerical values whose pairwise differences approximate optimally the relative preferences over countries as expressed in their risk ratings. A common way to calculate the relative preferences is based on interpreting sovereign ratings as cardinal values (see, e.g., Ferri et al. 1999; Hu et al. 2002; Sy 2004). If the sovereign ratings β are viewed as cardinal values, then one can view the relative preferences Δ as differences of the corresponding ratings:

$$\Delta\left(P_{ij}\right) = \beta_i - \beta_j, \quad \text{for all } i,j \in I, i \neq j \tag{16.10}$$

Since the system (16.10) is not necessarily consistent, it should be relaxed in the following way:

$$\Delta\left(P_{ij}\right) = \beta_i - \beta_j + \varepsilon_{ij}, \quad \text{for all } i,j \in I, i \neq j \tag{16.11}$$

The values of the β's providing the best L_2 approximation of the Δ's can be found by solving the following multiple linear regression problem:

$$\Delta(\pi) = \sum_{k \in I} \beta_k * x_k(\pi) + \varepsilon(\pi), \tag{16.12}$$

where $\pi = \{(i,j)|i,j \in I, i \neq j\}$ and $x_k(i,j) = \begin{cases} 1, & \text{for } k = i \\ -1, & \text{for } k = j \\ 0, & \text{otherwise} \end{cases}$.

The logical rating scores β_k obtained by fitting the regression model are given in Column 8 of Table 16.16.

16.4.4 Evaluation of the Results

In this section, we analyze the results obtained with the proposed rating systems. The evaluation of the results involves the analysis of the following:

- The relative preferences
- The partially ordered set (i.e., the logical dominance relationship)
- The LRS scores and the Condorcet ratings (extremal linear extensions)

with respect to the rating system of S&P's as well as that of Moody's and The Institutional Investor.

16.4.4.1 Canonical Relative Preferences

In this section, we evaluate the quality of the informational content in the matrix Δ of relative preferences. To reach this goal, we first define, for any set of numerical scores s_i representing sovereign ratings, the *canonical relative preferences* $d_{ij} = s_i - s_j$ for each pair of countries. Second, we compare the LAD relative preferences $\Delta(P_{ij})$ with the canonical relative preferences $d^{S\&P}{}_{ij}$, $d^{M}{}_{ij}$, $d^{II}{}_{ij}$, and $d^{LRS}{}_{ij}$ associated, respectively, to the S&P's ratings, Moody's ratings, The Institutional Investor's scores, and the logical rating scores (Hammer et al. 2011). The corresponding matrices of relative preferences are denoted $d^{S\&P}$, d^{M}, d^{II}, and d^{LRS}, respectively. We evaluate the proximity between the LAD relative preferences (Hammer et al. 2006) and the canonical relative preferences based on their correlation levels (Table 16.9).

The high correlation levels show that the LAD relative preferences are in strong agreement with both the ratings of S&P's and those of the other agencies, as well as with the logical rating scores. The superiority of the logical rating scores compared to the relative preferences associated with the LAD discriminant is explained by the filtering of the noise done when deriving the LRS scores.

16.4.4.2 Preference Orders

The logical dominance relationship is compared to the preference orders derived from the S&P's and Moody's ratings (viewed as partially ordered sets) and the partially ordered sets associated with The Institutional Investor's scores and the logical rating scores. The partially ordered sets corresponding to The Institutional Investor's scores and the logical rating scores are obtained as follows:

$$\begin{array}{ll} i \succ j & \text{if } s_i - s_j > \theta \\ i \prec j & \text{if } s_i - s_j < -\theta \\ i \| j & \text{otherwise} \end{array}$$

where θ is the positive number chosen so as to obtain a partially ordered set of the same density as the dominance relationship and s_i represents the numerical score given to country i by the respective rating system. The incomparability between two countries means, for S&P's and Moody's, that the two countries are equally creditworthy, while for the logical dominance relationship, it means that the evidence about the relative creditworthiness of the two countries is either missing or conflicting. The concept of density of a partially ordered set is defined in Hammer et al. (2006) and represents the extent to which a partial order on a set of countries differentiates them by their creditworthiness.

To assess the extent to which the preference orders agree with each other, the following concepts are introduced (Hammer et al. 2006). Given a pair of countries (i, j), two partially ordered sets are:

- In *concordance* if one of the following relations $i \succ j$, $i \prec j$, or $i \| j$ holds for both partially ordered sets
- In *discordance* if $i \succ j$ for one of the partially ordered sets and $i \prec j$ for the other one
- *Incomparable* otherwise, that is, if $i \| j$ for one of the partially ordered sets, and either $i \succ j$ or $i \prec j$ for the other partially ordered set.

We measure the *levels of concordance*, *discordance*, or *incomparability* between two partially ordered sets by the fractions of pairs of countries for which the two partially ordered sets are, respectively, in concordance, discordance, or incomparable. Hammer et al. (2006) have shown that there is a very high level of agreement between the following:

- The logical dominance relationship and the preference orders associated with S&P's and Moody's ratings and The Institutional Investor scores
- The logical dominance relationship and the logical rating scores

16.4.4.3 Discrepancies with S&P's Logical Dominance Relationship

This section is devoted to the study of the discordance between the logical dominance relationship and the preference order of S&P's. We define as a *discrepancy* (Hammer et al. 2006) a country pair for which the logical dominance relationship and the preference order of S&P's are in discordance. The 2.17 % discordance level between the logical dominance relationship and the preference order of S&P's represents 51 discrepancies.

Next, in order to determine the minimum number of countries, for which the S&P's ratings must be changed so that the new *adjusted* S&P's preference order has a 0 % discordance level with the dominance relationship, we solve the integer program below:

$$\min \sum_{i \in I} a_i$$

subject to

$$\left|S_i^* - S_i\right| \le M^* a_i, \quad \text{for all } i \in I \tag{16.13}$$

$$S_i^* \ge S_j^* \quad \text{for every pair } (i, j) \text{ such that } i \underset{LAD}{\succ} j$$

$$a_i \in \{0, 1\}, \quad S_i^* \varepsilon \{0, 1, \ldots, 21\} \qquad \text{for all } i \in I$$

where a_i takes the value 1 if the S&P's rating of country i must be modified and the value 0 if otherwise, S_i is the original S&P's rating of country i, S_i^* is the adjusted S&P's rating of country i, and M is a sufficiently large positive number (e.g., $M = 22$).

The optimal solution of Eq. 16.13 shows that the 0 % discordance level can be achieved by adjusting the S&P's ratings of nine countries: France, India, Japan, Colombia, Latvia, Lithuania, Croatia, Iceland, and Romania. To check the relevance of the proposed rating adjustments, we examine the S&P's ratings posterior to December 1998. We observe that Romania, Japan, and Columbia's S&P's ratings have been modified in the direction suggested by our model. More precisely, Columbia was downgraded by S&P's twice, moving from BBB− in December 1998 to BB+ in September 1999 and then to BB in March 2000. Japan was downgraded to AA+ in February 2001 and AA in November 2001. Romania was upgraded to B in June 2001. The S&P's rating of the other countries (Iceland, France, India, Croatia, Latvia, and Lithuania) has remained unchanged.

Logical Rating Scores

We have already shown that the LRS and the S&P's ratings are in close agreement. However, since LRS and the S&P's ratings are not expressed on the same scale, the comparison of the two scores of an individual country presents a challenge. In order to bring the LRS and the S&P's ratings to the same scale, we apply a linear transformation $a^*\beta_i + c$ to the logical rating scores β_i in such a way that the mean square difference between the transformed LRS and the S&P's ratings is minimized. This is obtained by solving a series of quadratic optimization problems in which the decision variables are a and c. Clearly, the consistency of the LRS and S&P's ratings is not affected by this transformation.

In 1998, it appears that five countries (Columbia, Hong Kong, Malaysia, Pakistan, and Russia) have an S&P's rating that does not fall within the confidence interval of the transformed LRS. The 1-year modification of the S&P's ratings for Columbia, Pakistan, and Russia is in agreement with the 1998 LRS of these countries and highlights the prediction power of the LRS model. Moreover, the evolution of the S&P's ratings of Malaysia and Hong Kong is also in agreement with their 1998 LRS. Indeed, both Malaysia and Hong Kong have been upgraded shortly thereafter, the former moving from BBB- to BBB in November 1999 and the latter from A to A+ in February 2001.

Table 16.10 Correlation analysis

	S&P's	Moody's	II	OE	PE	LRS
S&P's	100 %	98.01 %	96.18 %	94.31 %	95.40 %	95.54 %
Moody	98.01 %	100 %	96.31 %	94.13 %	95.42 %	95.20 %
II	96.18 %	96.31 %	100 %	93.26 %	94.62 %	94.11 %
OE	94.31 %	94.13 %	93.26 %	100 %	99.15 %	99.24 %
PE	95.40 %	95.42 %	94.62 %	99.15 %	100 %	99.10 %
LRS	95.54 %	95.20 %	94.11 %	99.24 %	99.10 %	100 %

16.4.4.4 Optimistic and Pessimistic Extensions

The optimistic and pessimistic extensions of the logical dominance relationship (Table 16.16) comprise 21 levels, while the S&P's rating system contains 22 rating categories. Table 16.10 provides the correlation levels between all the ratings (scores). The analysis reconfirms the high level of agreement between the proposed rating models and that of S&P's. The high correlation levels attest that the LRS approximate very well the S&P's ratings, as well as those of the other rating agencies.

16.4.4.5 Temporal Validity

In this section, we apply the "out-of-time" or "walk-forward" validation approach (Sobehart et al. 2000; Stein 2002) to further verify the robustness and the relevance of our rating models. This involves testing how well the LAD model derived from the 1998 data performs when applied to the 1999 data. To evaluate the "temporal validity" of the proposed models, we proceed as follows: (i) we derive the LAD relative preferences, (ii) we build the logical dominance relationship and run the regression model for the LRS scores, (iii) we calculate the weak Condorcet ratings (pessimistic and optimistic extensions) and the logical rating scores, and (iv) we compare these to the rating systems of the rating agencies (S&P's, Moody's, and The Institutional Investor).

Relative Preferences

Table 16.11 shows that the LAD relative preferences are highly correlated with those of the S&P's rating system, as well as with the logical rating scores. The LAD relative preferences and the LRS were obtained by applying to the 1999 data the models derived from the 1998 data.

The high levels of pairwise correlations between the S&P's 1999 ratings, the relative preferences given by the LAD discriminant, and the canonical relative preferences corresponding to LRS show that the LRS model has a very strong temporal stability and indicate its high predictive power.

Preorders

The logical dominance relationship is compared to the 1999 preference order of the S&P's and the partially ordered set associated with the logical rating scores. Table 16.12 displays the concordance, discordance, and incomparability levels

Table 16.11 Correlation levels between relative preference matrices

	$d^{S\&P}$	Δ	d^{LRS}
$d^{S\&P}$	100 %	91.70 %	94.12 %
Δ	91.70 %	100 %	96.98 %
d^{LRS}	94.12 %	96.98 %	100 %

Table 16.12 Concordance, discordance, and incomparability levels with dominance relationship

	Logical dominance relationship		
	Concordance	Incomparability	Discordance
S&P's	84.21 %	13.15 %	2.64 %
Logical rating score	93.43 %	6.49 %	0.08 %
	Logical rating score		
	Concordance	Incomparability	Discordance
S&P's	83.46 %	12.69 %	3.85 %

between the logical dominance relationship, the preference order of S&P's, and the partial order associated with the logical rating scores and underlines their very strong level of agreement.

Discrepancies with S&P's

The discordance level between the logical dominance relationship obtained using the 1999 data and the preference order of the 1999 S&P's ratings is equal to 2.64 %. The solution of the above integer programming problem (16.13) reveals that the discrepancies would disappear if one modified the ratings of eight countries (France, Japan, India, Colombia, Latvia, Croatia, Iceland, and Hong Kong). The relevance of the ratings obtained with the logical dominance relationship is proven by observing the rating changes published by S&P's subsequent to December 1999.

The identification of the discrepancies between S&P's ratings and the LRS scores requires the derivation of the confidence intervals for the new, 1999 observations and therefore for the transformed LRS of each country (Hammer et al. 2007). We denote by n and p the number of observations and predictors, respectively. The expression $t(1 - \alpha/2, n - p)$ refers to the Student test with $(n - p)$ degrees of freedom, and with upper and lower tail areas of $\alpha/2$, X_j is the p-dimensional vector of the values taken by the observation Y_j on the p predictors, $X_p^{'}$ is the transposed of X_j, and $(X'X)^{-1}$ is the variance-covariance matrix, that is, the inverse of the $[p \times p]$-dimensional matrix $(X'X)$. Denoting by MSE the mean square of errors, the estimated variance $s^2\left[\hat{Y}_j\right] = MSE^*\left[1 + X_j^{'}\left(X^{'}X\right)^{-1}X_j\right]$ of the predicted rating $\hat{Y}_j$ and the $(1-\alpha)$ confidence interval for $\hat{Y}_{j,n}$ are given by

$$\left\{\hat{Y}_{j,n} - t(1 - \alpha/2, n - p)^* s[pred], \hat{Y}_{j,n} + t(1 - \alpha/2, n - p)^* s[pred]\right\} \quad (16.14)$$

Hammer et al. (2007b) say that there is a discrepancy between S&P's rating R_j^{SP} and the logical rating score if

$R_j^{SP} \notin \{\hat{Y}_{j,n} - t(1-\alpha/2, n-p)^* s[pred], \hat{Y}_{j,n} + t(1-\alpha/2, n-p)^* s[pred]\}$ for $\alpha = 0.1$.

Applying the 1998 LRS model to the 1999 data, only two countries (Russia and Hong Kong) have S&P's ratings that are outside the confidence intervals of the corresponding transformed LRS. The creditworthiness of these two countries seems to have been under-evaluated by S&P's in 1998: the ratings of both were upgraded in 1999.

Condorcet Ratings and LRS Scores

The optimistic and pessimistic extensions (Table 16.16) of the logical dominance relationship obtained using the 1999 data both comprise 20 levels, while the S&P's rating system contains 22 rating categories. Table 16.13 provides the correlation levels between the 1999 S&P's ratings, the optimistic and pessimistic extension levels, and the logical rating scores. The high levels of correlation and their comparison with those presented in Table 16.10 provide further evidence of the temporal validity of the proposed models.

16.4.4.6 Predicting Creditworthiness of Unrated Countries

Condorcet Approach

The application of the logical dominance relationship to predict the rating of countries not included in the original dataset, and for years subsequent to 1998, is an additional validation procedure, sometimes referred to as "out-of-universe" cross-validation (Sobehart et al. 2000; Stein 2002). We use the 1998 LAD model to calculate the relative preferences for all pseudo-observations involving one or two of the four "new" countries (Ecuador, Guatemala, Jamaica, Papua New Guinea), which allows us to derive the logical dominance relationship and the computation of the optimistic and pessimistic extensions of previously unrated countries. The levels assigned to these countries by the recalculated optimistic and pessimistic extensions are shown in Table 16.14.

It appears that:

- Guatemala's first S&P's rating (in 2001) was BB. Guatemala's OE/PE levels are the same as Morocco's (the only country with OE = PE = 5), and Morocco's S&P's rating in 1999 was BB.

Table 16.13 Correlation analysis

	S&P's	*OE*	*PE*	*LRS*
S&P's	100.00 %	95.09 %	95.15 %	94.12 %
OE	95.09 %	100.00 %	99.59 %	98.43 %
PE	95.15 %	99.59 %	100.00 %	98.24 %
LRS	94.12 %	98.43 %	98.24 %	100.00 %

Table 16.14 Out-of-universe validation

	Optimistic extension	Pessimistic extension	First S&P's rating	S&P's linear extension
Ecuador	3	2	SD (07/2000)	0
Guatemala	5	5	BB (10/2001)	10
Jamaica	3	2	B (11/1999)	7
Papua New Guinea	3	3	B+ (01/1999)	8

- Jamaica's first S&P's rating (1999) was B. Its OE/PE levels (OE = 3, PE = 2) are identical to these of Paraguay, Brazil, the Dominican Republic, and Bolivia, which had 1999 S&P's ratings of B, B+, B+, and BB-, respectively.
- The first S&P's rating for Papua New Guinea was B+. Its OE/PE levels (OE = 3, PE = 3) are the same as those of Peru and Mexico, which both had 1999 S&P's ratings of BB.
- Ecuador's OE/PE levels (OE = 3, PE = 2) are the same as those of Paraguay, Brazil, the Dominican Republic, and Bolivia, which had 1999 S&P's ratings of B, B+, B+, and BB-, respectively. Interestingly, while the initial S&P's rating of Ecuador was SD (in July 2000), it was upgraded in August 2000 (1 month later) to B.

The striking similarity between the initial S&P's rating of each of the four countries discussed above and the S&P's ratings of those countries which have the same OE/PE levels validates the proposed model, indicating its power to predict the creditworthiness of previously unrated countries.

LRS Approach

The LAD discriminant, which does not involve in any way the previous years' S&P's ratings, allows the rating of previously unrated countries in the following way. First, we construct all the pseudo-observations involving the new countries to be evaluated. Second, we calculate the relative preferences for these pseudo-observations, and we add the resulting columns and rows to the matrix of relative preferences. Third, we determine the new LRS for all the countries (new and old) by running the multiple linear regression model (16.12). Fourth, we apply the linear transformation defined above to the LRS so that the transformed LRS and the S&P's ratings are on the same scale.

The evaluation of the ability of LRS to accurately predict S&P's ratings is carried out by comparing the predicted LRS (obtained as described above) and the S&P's ratings (when they first become available). We compute the confidence intervals (16.14) for the transformed LRS of four countries never rated by S&P's by December 1998. The predictions for Guatemala, Jamaica, and Papua New Guinea correspond perfectly to the first time (subsequent) S&P's ratings. The comparison between the LRS and the first (July 2000) S&P's rating (SD) to Ecuador shows that S&P's rated it too harshly, since 1 month later S&P's raised its rating to B-, justifying the LRS prediction.

16.4.5 Importance of Variables

The methodology developed in this chapter permits the assessment of the importance of the variables in rating countries' creditworthiness. In LAD, the importance of variables is associated with their use in the patterns of the LAD model and is usually measured by the proportion of patterns containing a particular variable. The patterns of the 1998 LAD model show that the three most frequently used variables are financial depth and efficiency, political stability, and gross domestic product per capita (appearing in 47.5 %, 39.4 %, and 35.6 % of the LAD patterns, respectively).

The presence of political stability among the three most significant ones in the selected set justifies the inclusion of political variables in country risk rating models. This result is in agreement with the cost-benefit approach to country risk (i.e., the risk of defaulting is heavily impacted by the political environment, see Brewer and Rivoli 1990, 1997; Citron and Neckelburg 1987; Afonso 2003) which is not a view shared by all (Haque et al. 1996, 1998).

The fact that the LAD approach identifies gross domestic product per capita as significant was expected, for most studies on country risk ratings acknowledge its crucial importance in evaluating the creditworthiness of a country. A key new result is the identification of the *financial depth and efficiency* variable as a major factor for the prediction of country risk ratings.

16.5 Conclusions

The central objective of this study is to develop transparent, consistent, self-contained, and stable credit risk rating models, closely approximating the risk ratings provided by some of the main rating agencies. We use the combinatorial optimization method called LAD and develop a relative creditworthiness approach for assessing the credit risk of countries, while an absolute creditworthiness approach is used for financial institutions.

The evaluation of the creditworthiness of financial organizations is particularly important due to the growing number of banks going bankrupt and the magnitude of losses caused by such bankruptcies, and it is challenging due to the opaqueness of the banking sector and the high variability of banks' creditworthiness. We use the logical analysis of data (LAD) to reverse engineer the Fitch bank credit ratings. The LAD method identifies strong combinatorial patterns distinguishing banks with high and low ratings. These patterns constitute the core of the rating model developed here for assessing the credit risk of banks. The results show that the LAD ratings are in very close agreement with the Fitch ratings. In that respect, it is important to note that the critical component of the LAD rating system – the LAD discriminant – is derived utilizing only information about whether a bank's rating is "high" or "low," without the exact specification of the bank's rating category. Moreover, the LAD approach uses only a fraction of the observations in the dataset. The higher classification accuracy of LAD appears even more clearly when performing cross-validation and applying the LAD model derived by using the banks in the training set to those in the testing one.

This study also shows that the LAD-based approach to reverse engineering bank ratings provides a model that is *parsimonious* and *robust*. This approach allows to derive rating models with varying levels of granularity that can be used at different stages in the credit granting decision process and can be employed to develop internal rating systems that are Basel 2 compliant. Besides impacting the credit risk of a financial institution, the use of the generalizable and accurate credit risk rating system proposed here will also be critical in mitigating the financial institution's operational risk due to breakdowns in established processes and risk-management operations or to inadequate process mapping within business lines. In particular, the reliance on such risk rating system will reduce the losses due to mistakes made in executing transactions, such as settlement failures, failures to meet capital regulatory requirements, or untimely debt collections, or the losses due to the offering of inappropriate financial products or credit conditions, or giving incorrect advice to counterparty.

The evaluation of the creditworthiness of countries is also of utmost importance, since the country's risk rating is generally viewed as the upper bound on the rating that entities within a given country can be assigned. This study proposes an LAD methodology for inducing a credit risk system from a set of country risk rating evaluations. It uses nine economic and three political variables to construct the relative preferences of countries on the basis of their creditworthiness. Two methods are then developed to construct countries' credit rating systems on the basis of their relative creditworthiness. The first one is based on extending the preorder of countries using the Condorcet voting technique and provides two rating systems (weak Condorcet winners and losers), while the second one uses linear regression to determine the logical rating scores.

The proposed rating systems correlate highly with those of the utilized rating system (S&P's) and those of other rating agencies (Moody's and The Institutional Investor) and are shown to be stable, having an excellent classification accuracy when applied to the following years' data or to the ratings of previously unrated countries. Rating changes implemented by the S&P's in subsequent years have resolved most of the (few) discrepancies between the constructed partially ordered set and S&P's initial ratings. This study provides new insights on the importance of variables by supporting the necessity of including in the analysis, in addition to economic variables, also political variables ("political stability"), and by identifying "financial depth and efficiency" as a new critical factor in assessing country risk.

The rating systems proposed here for banks as well as countries are as follows:

- Avoid overfitting as attested by the back-testing analysis (i.e., extremely high concordance between in- and out-of-sample rating predictions calculated using the k-folding and jackknife cross-validation methods).
- Distinguish themselves from the rating models in the existing literature by their *self-contained* nature, that is, by their non-reliance on any information derived from lagged ratings. Therefore, the high level of correlation between predicted

Table 16.15 Standard & Poor's country rating system

	Level	Description
Investment rating	AAA	An obligor rated AAA has extremely strong capacity to meet its financial obligations. AA is the highest issuer credit rating assigned by S&P's
	AA	An obligor rated AA has very strong capacity to meet its financial commitments. It differs from the highest rated obligors only in small degree
	A	An obligor rated A has strong capacity to meet its financial commitments but is somewhat more susceptible to the adverse effects of changes in circumstances and economic conditions than obligors in higher-rated categories
	BBB	An obligor rated BBB has adequate capacity to meet its financial commitments. However, adverse economic conditions or changing circumstances are more likely to lead to a weakened capacity of the obligor to meet its financial commitments
Speculative rating	BB	An obligor rated BB is less vulnerable in the near term than other lower-rated obligors. However, it faces major ongoing uncertainties and exposure to adverse financial or economic conditions which could lead to its inability to meet financial commitments
	B	An obligor rated B is more vulnerable than the obligors rated BB, but, at the time of the rating, it has the capacity to meet financial commitments. Adverse business, financial conditions could likely impair its capacity to meet financial commitments
Default rating	CCC	An obligor rated CCC is vulnerable at the time of the rating and is dependent upon favorable business, financial, and economic conditions to meet financial commitments
	CC	An obligor rated CC is highly vulnerable at the time of the rating
	C	An obligor rated C is vulnerable to nonpayment at the time of the rating and is dependent upon favorable business, financial, and economic conditions to meet financial commitments
	D	An obligor rated D is predicted to default
	SD	An obligor rated SD (selected default) is presumed to be unwilling to repay

and actual ratings cannot be attributed to the reliance on lagged ratings and is a reflection of the predictive power of the independent variables included in these models. An important advantage of the non-recursive nature of the proposed models is their applicability to not-yet-rated obligors.

The scope of the proposed methodology extends beyond the rating problems discussed in this study and can be used in many other contexts where ratings are relevant. The proposed methodology is applicable in the general case of inferring an objective rating system from archival data, given that the rated objects are characterized by vectors of attributes taking numerical or ordinal values.

Appendix

See Tables 16.15 and 16.16.

Table 16.16 1998 Ratings

Countries	S&P's ratings (1998)	S&P's preorder (1998)	Moody's ratings (1998)	The Institutional Investor ratings (1998)	Optimistic extension ratings (1998)	Pessimistic extension ratings (1998)	LRS scores (1998)	S&P's ratings (1999)	S&P's preorder (1999)	Optimistic extension ratings (1999)	Pessimistic extension ratings (1999)	LRS scores (1999)
Argentina	BB	10	9	42.7	7	6	−0.2768	BB	10	7	6	−0.263
Australia	AA	19	19	74.3	16	16	−0.0289	AA+	20	16	16	−0.0128
Austria	AAA	21	21	88.7	17	17	−0.0094	AAA	21	17	16	0.0038
Belgium	AA+	20	20	83.5	15	14	−0.0476	AA+	20	15	15	−0.0439
Bolivia	BB−	9	8	28	3	3	−0.366	BB−	9	3	2	−0.3518
Brazil	BB−	9	7	37.4	2	2	−0.3744	B+	8	3	2	−0.4016
Canada	AA+	20	20	83	16	16	−0.0241	AA+	20	16	16	−0.0112
Chile	A−	15	14	61.8	11	11	−0.191	A−	15	11	11	−0.1841
China	BBB+	14	15	57.2	10	9	−0.2159	BBB	13	10	10	−0.224
Colombia	BBB−	12	12	44.5	2	2	−0.3854	BB+	11	2	2	−0.3964
Costa Rica	BB	10	11	38.4	7	6	−0.2748	BB	10	7	6	−0.257
Croatia	BBB−	12	12	39.03	5	5	−0.297	BBB−	12	6	6	−0.3202
Cyprus	A+	17	16	57.3	12	12	−0.1081	A	16	12	12	−0.1021
Czech Republic	A−	15	14	59.7	10	10	−0.2088	A−	15	10	10	−0.1904
Denmark	AA+	20	20	84.7	15	15	−0.048	AA+	20	15	15	−0.0492
Dominican Rep	B+	8	10	28.1	3	2	−0.3568	B+	8	3	2	−0.3431
Egypt	BBB−	12	11	44.4	6	5	−0.2915	BBB−	12	6	6	−0.3067
El Salvador	BB	10	12	31.2	4	4	−0.3379	BB+	11	5	4	−0.3301
Estonia	BBB+	14	14	42.8	9	8	−0.2518	BBB+	14	8	8	−0.245
Finland	AA	19	21	82.2	14	14	−0.064	AA+	20	14	13	−0.0458

(*continued*)

Table 16.16 (continued)

Countries	S&P's ratings (1998)	S&P's preorder (1998)	Moody's ratings (1998)	The Institutional Investor ratings (1998)	Optimistic extension ratings (1998)	Pessimistic extension ratings (1998)	LRS scores (1998)	S&P's ratings (1999)	S&P's preorder (1999)	Optimistic extension ratings (1999)	Pessimistic extension ratings (1999)	LRS scores (1999)
France	AAA	21	21	90.8	13	13	−0.0828	AAA	21	13	13	−0.0614
Germany	AAA	21	21	92.5	18	18	−0.001	AAA	21	18	17	0.0126
Greece	BBB	13	14	56.1	9	9	−0.2255	A−	15	9	9	−0.1917
Hong Kong	A	16	15	61.8	17	13	−0.017	A	16	16	15	0.0213
Hungary	BBB	13	13	55.9	9	8	−0.2442	BBB	13	9	8	−0.247
Iceland	A+	17	18	67	15	15	−0.047	A+	17	15	14	−0.0378
India	BB	10	10	44.5	1	1	−0.4063	BB	10	2	1	−0.3994
Indonesia	CCC+	5	6	27.9	0	0	−0.4576	CCC+	5	0	0	−0.4316
Ireland	AA+	20	21	81.8	17	16	−0.0179	AA+	20	17	16	−0.0249
Israel	A−	15	15	54.3	11	9	−0.2215	A−	15	10	9	−0.2189
Italy	AA	19	18	79.1	12	12	−0.1064	AA	19	12	12	−0.1122
Japan	AAA	21	20	86.5	15	14	−0.0604	AAA	21	15	14	−0.0506
Jordan	BB−	9	9	37.3	4	3	−0.323	BB−	9	4	3	−0.2818
Kazakhstan	B+	8	9	27.9	1	1	−0.4095	B+	8	1	1	−0.4048
Korea Rep.	BB+	11	11	52.7	9	6	−0.2649	BBB	13	10	8	−0.2182
Latvia	BBB	13	13	38	5	5	−0.3026	BBB	13	5	5	−0.3039
Lebanon	BB−	9	8	31.9	4	4	−0.3223	BB−	9	4	4	−0.3121
Lithuania	BBB−	12	11	36.1	4	4	−0.3247	BBB−	12	4	4	−0.3233
Malaysia	BBB−	12	12	51	11	9	−0.1676	BBB	13	11	11	−0.1712

Malta	A+	17	15	61.7	13	12	−0.0999	A	16	12	12	−0.2402
Mexico	BB	10	10	46	3	3	−0.3608	BB	10	3	3	−0.3284
Morocco	BB	10	11	43.2	6	5	−0.2952	BB	10	5	5	−0.2881
Netherlands	AAA	21	21	91.7	19	19	0.0251	AAA	21	19	18	0.0337
New Zealand	AA+	20	19	73.1	18	17	0.0001	AA+	20	17	16	0.0001
Norway	AAA	21	21	86.8	19	18	0.0125	AAA	21	18	18	−0.0076
Pakistan	CC	2	5	20.4	0	0	−0.4563	B−	6	0	0	−0.4501
Panama	BB+	11	11	39.9	7	6	−0.2712	BB+	11	6	6	−0.2487
Paraguay	BB−	9	7	31.3	2	2	−0.3865	B	7	3	2	−0.4066
Peru	BB	10	9	35	3	3	−0.3536	BB	10	3	3	−0.3644
Philippines	BB+	11	11	41.3	4	4	−0.3242	BB+	11	4	3	−0.349
Poland	BBB−	12	12	56.7	7	6	−0.2772	BBB	13	7	7	−0.2743
Portugal	AA	19	19	76.1	14	13	−0.0742	AA	19	14	13	−0.0706
Romania	B−	6	6	31.2	1	1	−0.3987	B−	6	1	1	−0.3942
Russia	CCC−	3	6	20	0	0	−0.4428	SD	0	0	0	−0.4197
Singapore	AAA	21	20	81.3	19	18	0.0073	AAA	21	18	18	0.0225
Slovak Republic	BB+	11	11	41.3	7	6	−0.2814	BB+	11	7	6	−0.269
Slovenia	A	16	15	58.4	11	9	−0.1922	A	16	10	9	−0.1878
South Africa	BB+	11	12	45.8	8	6	−0.2523	BB+	11	7	6	−0.2386
Spain	AA	19	19	80.3	13	12	−0.0924	AA+	20	13	13	−0.0798
Sweden	AA+	20	19	79.7	18	17	0.0106	AA+	20	17	17	0.0143
Switzerland	AAA	21	21	92.7	20	20	0.071	AAA	21	19	19	0.0613

(*continued*)

Table 16.16 (continued)

Countries	S&P's ratings (1998)	S&P's preorder (1998)	Moody's ratings (1998)	The Institutional Investor ratings (1998)	Optimistic extension ratings (1998)	Pessimistic extension ratings (1998)	LRS scores (1998)	S&P's ratings (1999)	S&P's preorder (1999)	Optimistic extension ratings (1999)	Pessimistic extension ratings (1999)	LRS scores (1999)
Thailand	BBB−	12	11	46.9	8	8	−0.2452	BBB−	12	8	8	−0.2383
Trinidad & Tob	BB+	11	11	43.3	6	6	−0.2824	BBB−	12	6	6	−0.248
Tunisia	BBB−	12	12	50.3	9	8	−0.2488	BBB−	12	8	8	−0.242
Turkey	B	7	8	36.9	0	0	−0.4458	B	7	0	0	−0.4177
UK	AAA	21	21	90.2	18	18	−0.0057	AAA	21	18	18	0.0062
United States	AAA	21	21	92.2	19	19	0.0205	AAA	21	19	18	0.0264
Uruguay	BBB−	12	12	46.5	7	7	−0.2695	BBB−	12	7	7	−0.2409
Venezuela	B+	8	7	34.4	0	0	−0.4444	B	7	1	0	−0.3921

We have converted the Standard & Poor's rating scale (columns 1 and 4) into a numerical scale (columns 2 and 5). Such a conversion is not specific to us. Bouchet et al. (2003), Ferri et al. (1999), and Sy (2004) proceed similarly. Moreover, Bloomberg, a major provider of financial data services, developed a standard cardinal scale for comparing Moody's, S&P's, and Fitch-BCA ratings (Kaminsky and Schmukler 2002). A higher numerical value denotes a higher probability of default. The numerical scale is referred to in this chapter as Standard & Poor's preorder

References

Afonso, A. (2003). Understanding the determinants of sovereign debt ratings: Evidence for the two leading agencies. *Journal of Economics and Finance, 27*(1), 56–74.

Afonso, A., Gomes, P., & Rother, P. (2007). *What "hides" behind sovereign debt ratings?* (European Central Bank Working Paper Series 711). https://www.repository.utl.pt/bitstream/10400.5/2134/1/ecbwp711.pdf

Alexe, S. (2002). *Datascope – A new tool for logical analysis of data. DIMACS Mixer Series*. DIMACS, Rutgers University.

Alexe, G., Alexe, S., Bonates, T. O., & Kogan, A. (2007). Logical analysis of data – The vision of Peter L. Hammer. *Annals of Mathematics and Artificial Intelligence, 49*, 265–312.

Altman, E. I., & Rijken, H. A. (2004). How rating agencies achieve rating stability. *Journal of Banking and Finance, 28*, 2679–2714.

Barr, R. S., & Siems, T. F. (1994). Predicting bank failure using DEA to quantify management quality. *Federal Reserve Bank of Dallas Financial Industry Studies, 1*, 1–31.

Basel Committee on Banking Supervision. (2001). *The internal ratings based approach*. http://www.bis.org/publ/bcbsca05.pdf

Basel Committee on Banking Supervision. (2004). *Bank failures in mature economies* (Working Paper 13).

Basel Committee on Banking Supervision. (2006). *International convergence of capital measurement and capital standards: A revised framework (Basel II)*.

Bhatia, A.V. (2002). *Sovereign credit risk ratings: An evaluation* (IMF Working Paper WP/03/170).

Boros, E., Hammer, P. L., Ibaraki, T., & Kogan, A. (1997). Logical analysis of numerical data. *Mathematical Programming, 79*, 163–190.

Boros, E., Hammer, P. L., Ibaraki, T., Kogan, A., Mayoraz, E., & Muchnik, I. (2000). An implementation of logical analysis of data. *IEEE Transactions on Knowledge and Data Engineering, 12*(2), 292–306.

Bouchet, M. H., Clark, E., & Groslambert, B. (2003). *Country risk assessment: A guide to global investment strategy*. Chichester: Wiley.

Bourke, P., & Shanmugam, B. (1990). *An introduction to bank lending* (Addison-Wesley Business Series). Sydney: Addison-Wesley.

Brewer, T. L., & Rivoli, P. (1990). Politics and perceived country creditworthiness in international banking. *Journal of Money, Credit and Banking, 22*, 357–369.

Brewer, T. L., & Rivoli, P. (1997). Political instability and country risk. *Global Finance Journal, 8*(2), 309–321.

Cantor, R., & Packer, F. (1996). Determinants and impact of sovereign credit ratings. *FRBNY Economic Policy Review, 2*, 37–53.

Choo E.U., & Wedley W.C., (2004). A Common Framework for Deriving Preference Values from Pairwise Comparison Matrices. *Computers and Operational Research 31,* 893–908.

Citron, J. T., & Neckelburg, G. (1987). Country risk and political instability. *Journal of Development Economics, 25*, 385–395.

Crama, Y., Hammer, P. L., & Ibaraki, T. (1988). Cause-effect relationships and partially defined boolean functions. *Annals of Operations Research, 16*, 299–326.

Curry, T., & Shibut, L. (2000). The Cost of the Savings and Loan Crisis: Truth and Consequences. *FDIC Banking Review, 2*(2), 26–35.

Czyzyk, J., Mesnier, M. P., & Moré, J. J. (1998). The NEOS server. *IEEE Computer Science Engineering, 5*(3), 68–75.

de Servigny, A., & Renault, O. (2004). *Measuring and managing credit risk*. New York: McGraw-Hill.

Eliasson, A. (2002). *Sovereign credit ratings* (Working Papers 02-1). Deutsche Bank.

Ferri, G., Liu, L.-G., & Stiglitz, J. (1999). The procyclical role of rating agencies: Evidence from the east asian crisis. *Economic Notes, 3*, 335–355.

Fitch Ratings. (2001). Fitch simplifies bank rating scales. Technical Report.
Fitch Ratings. (2006). The role of support and joint probability analysis in bank ratings. Fitch Special Report.
Galindo, J., & Tamayo, P. (2000). Credit risk assessment using statistical and machine learning/basic methodology and risk modeling applications. *Computational Economics, 15*, 107–143.
Gehrlein, W. V., & Lepelley, D. (1998). The condorcet efficiency of approval voting and the probability of electing the condorcet loser. *Journal of Mathematical Economics, 29*, 271–283.
Hammer, P. L. (1986). *Partially defined boolean functions and cause-effect relationships. International conference on multi-attribute decision making via OR-based expert systems*. Passau: University of Passau.
Hammer, P. L., Kogan, A., & Lejeune, M. A. (2006). Modeling country risk ratings using partial orders. *European Journal of Operational Research, 175*(2), 836–859.
Hammer, P. L., Kogan, A., & Lejeune, M. A. (2011). Reverse-engineering country risk ratings: Combinatorial non-recursive model. *Annals of Operations Research, 188*, 185–213.
Hammer, P. L., Kogan, A., & Lejeune, M. A. (2012). A logical analysis of banks' financial strength ratings. *Expert Systems with Applications, 39*(9), 7808–7821.
Haque, N. U., Kumar, M. S., Mark, N., & Mathieson, D. (1996). The economic content of indicators of developing country creditworthiness. *International Monetary Fund Working Paper, 43*(4), 688–724.
Haque, N. U., Kumar, M. S., Mark, N., & Mathieson, D. (1998). The relative importance of political and economic variables in creditworthiness ratings. *International Monetary Fund Working Paper, 46*, 1–13.
Hu, Y.-T., Kiesel, R., & Perraudin, W. (2002). The estimation of transition matrices for sovereign credit ratings. *Journal of Banking and Finance, 26*(7), 1383–1406.
Huang, Z., Chen, H., Hsu, C.-J., Chen, W.-H., & Wu, S. (2004). Credit rating analysis with support vector machines and neural networks: A market comparative study. *Decision Support Systems, 37*, 543–558.
Jain, K., Duin, R., & Mayo, J. (2000). Statistical pattern recognition: A review. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22*, 4–37.
Kaminsky, G., & Schmukler, S. L. (2002). Emerging market instability: Do sovereign ratings affect country risk and stock returns? *World Bank Economic Review, 16*, 171–195.
Kaufmann, D., Kraay A., & Zoido-Lobaton, P. (1999a). *Aggregating governance indicators*. (World Bank Policy Research Department Working Paper, 2195).
Kaufmann, D., Kraay, A., & Zoido-Lobaton, P. (1999b). *Governance matters* (World Bank Policy Research Department Working Paper, 2196).
Kunczik, M. (2001). *Globalization: News media, images of nations and the flow of international capital with special reference to the role of rating agencies*. Paper presented at the IAMCR Conference, Singapore, pp. 1–49.
Larrain, G., Reisen, H., & von Maltzan, J. (1997). *Emerging market risk and sovereign credit ratings* (pp. 1–30). OECD Development Centre Working Papers 124, OECD Publishing.
Moody's. (2006). Bank financial strength ratings: Revised methodology. Moody's Global Credit Research Report.
Mora, N. (2006). Sovereign credit ratings: Guilty beyond reasonable doubt? *Journal of Banking and Finance, 30*(7), 2041–2062.
Morgan, D. P. (2002). Rating banks: Risk and uncertainty in an opaque industry. *The American Economic Review, 92*(4), 874–888.
Ng, W.-Y., Choi, K.-W., & Shum, K.-H. (1996). Arbitrated matching: Formulation and protocol. *European Journal of Operational Research, 88*(2), 348–357.
Poon, W. P. H., Firth, M., & Fung, H.-G. (1999). A multivariate analysis of the determinants of Moody's bank financial strength ratings. *Journal of International Financial Markets, Institutions & Money, 9*(3), 267–283.

Quenouille, M. (1949). Approximate tests of correlation in time series. *Journal of the Royal Statistical Society, Series B, 11*, 18–84.
Reinhart, C. M. (2002). Default, currency crises, and sovereign credit ratings. *World Bank Economic Review, 16*, 151–170.
Sarkar, S., & Sriram, R. S. (2001). Bayesian models for early warning of bank failures. *Management Science, 47*(11), 1457–1475.
Sobehart, J. R., Keenan, S. C., & Stein, R. M. (2000). *Benchmarking quantitative default risk models: A validation methodology*. New York: Moody's Investors Service.
Stein, R. M. (2002). *Benchmarking default prediction models: Pitfalls and remedies in model validation*. New York: Moody's KMV.
Sy, A. N. R. (2004). Rating the rating agencies: Anticipating currency crises or debt crises? *Journal of Banking and Finance, 28*(11), 2845–2867.
Tarjan, R. E. (1972). Depth-first search and linear graph algorithms. *SIAM Journal on Computing, 1*, 146–160.
Treacy, W. F., & Carey, M. S. (2000). Credit risk rating systems at large US banks. *Journal of Banking & Finance, 24*(1–2), 167–201.

Dynamic Interactions Between Institutional Investors and the Taiwan Stock Returns: One-Regime and Threshold VAR Models

17

Bwo-Nung Huang, Ken Hung, Chien-Hui Lee, and Chin W. Yang

Contents

B.-N. Huang
National Chung-Cheng University, Minxueng Township, Chiayi County, Taiwan
e-mail: ecdbnh@ccu.edu.tw

K. Hung (✉)
Division of International Banking & Finance Studies, Texas A&M International University, Laredo, TX, USA
e-mail: ken.hung@tamiu.edu

C.-H. Lee
National Kaohsiung University of Applied Sciences, Kaohsiung, Taiwan
e-mail: chlee@cc.kuas.edu.tw

C.W. Yang
Clarion University of Pennsylvania, Clarion, PA, USA
e-mail: yang@mail.clarion.edu; ecdycw@ccu.edu.tw

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_17,

Abstract

This paper constructs a six-variable VAR model (including NASDAQ returns, TSE returns, NT/USD returns, net foreign purchases, net domestic investment companies (dic) purchases, and net registered trading firms (rtf) purchases) to examine: (i) the interaction among three types of institutional investors, particularly to test whether net foreign purchases lead net domestic purchases by dic and rtf (the so-called demonstration effect); (ii) whether net institutional purchases lead market returns or vice versa; and (iii) whether the corresponding lead-lag relationship is positive or negative? The results of unrestricted VAR, structural VAR, and multivariate threshold autoregression models show that net foreign purchases lead net purchases by domestic institutions and the relation between them is not always unidirectional. In certain regimes, depending on whether previous day's TSE returns are negative or previous day's NASDAQ returns are positive, we find ample evidence of a feedback relation between net foreign purchases and net domestic institutional purchases. The evidence also supports a strong positive-feedback trading by institutional investors in the TSE. In addition, it is found that net dic purchases negatively lead market returns in Period 4. The MVTAR results indicate that net foreign purchases lead market returns when previous day's NASDAQ returns are positive and have a positive influence on returns.

Readers are well advised to refer to chapter appendix for detailed discussion of the unrestricted VAR model, the structural VAR model, and the threshold VAR analysis.

Keywords

Demonstration effect • Multivariate threshold autoregression model • Foreign investment • Lead-lag relationship • Structural VAR • Block Granger causality • Institutional investors • Domestic investment companies • Registered trading firms • Qualified foreign institutional investors

17.1 Introduction

As financial markets are gradually liberalized in emerging economies, capital investments have been flowing into these countries at increasing rates. Needless to say, such capital movements in terms of bringing in direct investment along with technological know-how can be instrumental in raising a nation's productivity. On the downside, capital inflows directed at investing in the host country's security markets can be disruptive. From a macroeconomics perspective, foreign capital inflows are beneficial in that they provide much-needed capital. From a microeconomics perspective, they also lower the cost of capital and enhance competitiveness. Nevertheless, capital inflows can also be disruptive if arbitrage is their main purpose. When capital flight occurs, as was the case in the 1997 Asian financial debacle and the 1994 Mexican Peso crisis, movements in foreign capital could be extremely disruptive and damaging.

Given the potentially negative impact of foreign investments, Taiwan's Ministry of Finance exercised caution by implementing foreign capital policies in a three-stage

Table 17.1 Inflows and outflows of foreign capital investment in the TSE Unit: millions of US dollars

Period	Securities investment companies		Foreign institutional investors		Natural person		Foreign total
	Inflow	Outflow	Inflow	Outflow	Inflow	Outflow	Net inflow
1991	263	53	448	0	0	0	658
1992	57	61	447	17	0	0	426
1993	653	93	1,859	97	0	0	2,322
1994	451	207	2,279	634	0	0	1,889
1995	664	457	3,509	1,506	0	0	2,210
1996	565	477	6,213	3,881	334	8	2,747
1997/1 ~ 4	12	644	4,442	2,529	261	78	1,465
Cumulative amount	2,664	1,992	19,198	8,664	595	85	11,716

Source of data: Securities and Futures Commission, Ministry of Finance, Taiwan

process. First, it directed a number of local trust companies to issue investment funds abroad for the local stock market. Second, qualified foreign investment companies were allowed to invest in Taiwan's stock market. Third, foreign individual as well as institutional investors were permitted to directly participate in trading in the Taiwan Stock Exchange Corporation (hereafter TSE). The first stage took effect from September 1983 until December 29, 1990, when the second stage replaced the first stage with a maximum investment limit of $2.5 billion. In response to the strong demand to invest in the TSE, the ceiling was raised to $5 billion and to $7.5 billion in August 1993 and March 1994, respectively. Foreign investments later increased substantially (see Table 17.1), and as a result the ceiling was lifted entirely in 1996 except for a maximum investment limit to each individual stock.

The relatively slow pace in allowing foreign investment in Taiwan was the result of ongoing debates between the Central Bank of China and the Securities and Futures Commission of the Ministry of Finance regarding the stability of foreign investment. The focus of the discussion was on the following three questions. First, are there differences in the trading behaviors of different type of institutions? One objective of opening up the domestic market to foreign investments is to utilize its advantages in information acquisition, information processing, and trade execution to improve the overall performance of local institutional investors.[1] Second, will

[1]This is because foreign institutions may well have better research teams and buy stocks according to the fundamentals such as the firm's future profitability. In contrast, local institutions and individual investors usually choose stocks based on insider information or what the newspapers write about. However, if stocks bought by foreign investors have a better performance than that of local institutions and individual investors, the latter tend to buy the stocks bought by successful foreign investors the previous day. Hence this gives rise to the so-called "demonstration effect." Such a concept is similar to the "herding" which Lakonishok et al. (1992) referred to as correlated trading across institutional investors. Nonetheless, their definition is close to the contemporaneous correlation rather than the lead-lag relation under study, and as such this leads to the demonstration effect which we documented.

the trades of the three groups of classified institutional investors affect stock returns of the TSE? Third, will the lifting of restrictions on foreign capital contribute to the volatility of both the foreign exchange and stock markets in Taiwan?

The purpose of this paper is to provide answers to two of the aforementioned issues: (1) Are there differences in the trading behaviors of different types of institutions? Typically, we examine interaction among the three types of institutions: qualified foreign institution investors (*qfii*), domestic investment companies (*dic*), and registered trading firms (*rtf*). (2) Does institutional trading "cause" stock returns or do institution investments follow movements in stock prices? The distinction between this paper and previous literature lies in following aspects: (i) Most previous studies uses low-frequency – yearly, quarterly, or at best weekly – data (Nofsinger and Sias 1999; Cai and Zheng 2002; Karolyi 2002). In this paper we employ daily data to help explore these issues in detail and to provide new evidence on the short-term dynamics among institutional investors and stock returns.[2] (ii) Unlike previous studies that used a bivariate VAR model (Froot et al. 2001), we use a six-variable VAR model, which includes three types of institutional trades, stock returns in the TSE, NT/USD exchange rate changes, and NASDAQ index returns to test related relationships. (iii) Improving on the conventional linear VAR analysis in the previous studies, we employ the threshold concept and split data into two regimes based on whether previous trading day's market returns are positive or negative.

A number of studies have examined the relationship between investment flows and stock returns. Brennan and Cao (1997) develop a theoretical model of international equity flows that relies on the informational difference between foreign and domestic investors. The model predicts that if foreign and domestic investors are differently informed, portfolio flows between two countries will be a linear function of the contemporaneous return on all national market indices. Moreover, if domestic investors have a cumulative information advantage over foreign investors about domestic securities, the coefficient of the host market return is expected to be positive.

Nofsinger and Sias (1999) use US annual data to investigate the relationship between stock returns and institutional and individual investors. They identify a strong positive correlation between changes in institutional ownership and stock returns measured over the same period. Their results suggest that (i) institutional investors practiced positive-feedback trade more than individual investors and (ii) institutional herding impacted price more than that by individual investors. Nofsinger and Sias show that institutional herding is positively correlated with lag returns and appears to be related to stock return momentum.

Choe et al. (1999) use order and trade data from November 30, 1996 to the end of 1997 to examine the impact of foreign investors on stock returns. They found strong

[2]Froot et al. (2001) also use daily data, but they examine the behavior of capital flows across countries. In addition, our models and approaches used in the estimation differ drastically from theirs.

evidence of positive-feedback trading and herding by foreign investors before South Korea's economic crisis. During the crisis period, herding lessened and positive- feedback trading by foreign investors mostly disappeared.[3]

Grinblatt and Keloharju (2000) use a Finland data set to analyze it to the extent that past returns determine the propensity to buy and sell. They find that foreign investors tend to be momentum investors, buying past winning stocks and selling past losing ones. Domestic investors, particularly households, tend to be contrarians.

Froot et al. (2001) by making use of daily international portfolio flows (44 countries from 1994 to 1998) along with a bivariate unrestricted VAR model find that lagged returns are statistically significant in predicting future flows. The evidence of the predictability of returns by flows is, however, ambiguous. In developed markets, there is no statistical evidence of stock predictability. For emerging markets, the evidence for predictability is strong, although less so for the Emerging European region. However, they estimate a restricted VAR model that assumes current inflows will affect current prices, and the causality does not run from contemporaneous returns to the flows. They provide the evidence of a positive contemporaneous correlation between current inflows and returns in emerging markets.

Hamao and Mei (2001) investigate the impact of foreign investment on Japan's financial markets. Using monthly data from July 1974 to June 1992, they find that (1) trades of foreign investors tend to increase market volatility more than that by domestic investors; (2) foreign investors have more sophisticated investment technology than do domestic investors; and (3) foreign investors seem to make investment decisions on the basis of not only short-term gains but also long-term fundamentals.

Cai and Zheng (2002) use institutional holding data from the third quarter of 1981 to the last quarter of 1996 in order to examine the lead-lag relationship between portfolio excess returns and the institutional trading. Beyond it, they compute institutional trading as the change of institutional holdings from last quarter to the current quarter. The unrestricted VAR analysis indicates that stock returns Granger-cause institutional trading on quarterly basis, not vice versa. This implies the institutions "herd" on past price behavior instead of being dominant price-setters in the market.

Using weekly data of Japan, Karolyi (2002) find consistent positive-feedback trading among foreign investors before, during and after the Asian financial debacle. Japanese banks, financial institutions, investment trusts and companies are, on the other hand, aggressive contrarian investors. There is no evidence that the trading activity by foreigners destabilized the markets during the crisis.[4]

Griffin et al. (2003) study the daily and intraday relationship between stock returns and trading of institutional as well as individual investors for NASDAQ 100 securities. The daily unrestricted VAR results indicate that the institutional

[3]Lakonishok et al. (1992) refer to the positive-feedback trading or trend chasing as buying winners and selling losers and the negative feedback trading or contrarian as buying losers and selling winners. Cai and Zheng (2002) point out that feedback trading occurs when lagged returns act as the common signal that the investors follow.

[4]Karolyi (2002) reaches such a conclusion because there is little evidence of any impact of foreign net purchases on future Nikkei returns or currency returns.

buy-sell imbalances are positively related to previous day's returns and the institutional buy-sell imbalances (previous day) are not associated with current return. The results are consistent with the finding by Sias and Starks (1997) using US data. Griffin et al. (2003) estimate a structural VAR with the contemporaneous returns in the institutional imbalance equation and discover a strong contemporaneous relationship between daily returns and institutional buy-sell imbalances.

Kamesaka et al. (2003) use Japanese weekly investment flow data over 18 years to investigate the investment patterns and performance of foreign investors, individual investors, and five types of institutional investors. Not surprisingly, they find individual investors perform poorly, while securities firms, banks, and foreign investors perform admirably over the sample period.

Several related studies focus mainly on Taiwan's stock market. Huang and Hsu (1999) detect decreased volatility in the weighted TSE using Levene's F-statistic following market liberalization. Lee and Oh (1995), implementing a vector autoregression (VAR) model, find a reduction in the explanatory power of macroeconomics variables. Wang and Shen (1999) indicate that foreign investments exert a positive impact on the exchange rate with only a limited effect on the TSE. In addition, by using the turnover rate as a proxy for non-fundamental factors and earnings per share for fundamental factors within the framework of a panel data model, Wang and Shen are able to identify that (i) the non-fundamental factors impacted the returns of the TSE before market liberalization and (ii) both the fundamental and non-fundamental factors exerted an impact following market liberalization.

Lee et al. (1999) investigate interdependence and purchasing patterns among institutional investors, large, and small individual investors. Their results, based on 15-min intraday transaction data (3 months for 30 companies), highlight the important role played by large individual investors, whose trading affects not only stock returns but also small individual investors. However, net buys (i.e., the difference between total buy and total sell) by institutional investors have no effect on the TSE returns, and vice versa.

The previous literature is predominantly focused on the relationship between institutional trading and stock returns, rarely on the interaction among institutional investors. For example, the majority of prior studies find evidence of positive-feedback trading by institutions, with the exception of Froot et al. (2001), who discover that in Latin America and emerging East Asian markets, the trading by institutions positively predicts future returns. Karolyi (2002) also detects that foreign investors in Japan are positive-feedback traders, while Japanese financial institution and companies are contrarian investors.

Most of the studies to date on these issues have been on the USA and Japanese markets despite that some of the literature gives scant attention to Taiwan's stock market. When investigating the related issues in large countries such as the USA and Japan, the influence of the foreign sector on the domestic market could be neglected; however, for a small country such as Taiwan, it should not be ignored. This is because the electronics industry in Taiwan is closely connected to the US companies listed on the NASDAQ. Ultimately, after examining the interaction among institutional investors and the dynamic relationship between stock returns and institutional investors, the conclusion may also be affected by whether returns of domestic market are positive or negative.

To circumvent the above problems, this paper employs a six-variable VAR model, which takes into account trades of three types of institutional investors (*qfii*, *dic*, and *rtf*), foreign returns, domestic returns, and changes in the NT/USD exchange rate to jointly test hypotheses under different market conditions. Using daily data, we find that net foreign purchases lead net domestic purchases. However, such a relation is not unidirectional. Under certain conditions (either when previous day's TSE returns are negative or previous day's NASDAQ returns are positive), we identify a feedback relation between net foreign purchases and net domestic purchases. It highlights the well-known argument in Taiwan regarding foreign investors: The demonstration effect on domestic institutional investors is not entirely correct. As for the lead-lag relation between market returns and institutional trading, we find that in most cases market returns at least lead both net foreign and *dic* purchases; however, market returns also lead net *rtf* purchases if the relationship between contemporaneous returns and institutional trading is considered. On the other hand, our results also indicate that net *dic* purchases lead market returns and are negatively associated with market returns in the fourth period. The MVTAR analysis shows that when previous day's NASDAQ returns are positive, net foreign purchases positively lead stock returns.

The remainder of this paper is organized as follows. Section 17.2 describes the sample data and the basic statistics. Section 17.3 investigates the lead-lag relation for three groups of institutional investors in order to explore the issue of whether foreign investments give rise to demonstration effects in Taiwan's stock market and examines the relationship between institutional trading activity and stock returns of the TSE. To further explore the interaction among three types of market participants, the sample is divided into two regimes based on either the sign of the market returns or that of the NASDAQ index returns of previous trading day, respectively. The last section provides a conclusion.

17.2 Sample Data and Basic Statistics

This paper employs daily data from December 13, 1995 to May 13, 2004 for a large sample analysis.[5] The variables considered include purchases (*qfiibuy*) and sales (*qfiisell*) by *qfii*, purchases (*dicbuy*) and sales (*dicsell*) by *dic*, purchases (*rtfbuy*) and sales (*rtfsell*) by *rtf*, TSE daily weighted stock index (p_t), NASDAQ stock index ($nasp_t$), and the NT/USD exchange rate (e_t). The data are from the Taiwan Economic Journal (TEJ). The changes in the exchange rate and the logarithmic returns on TSE and NASDAQ indices are defined as

[5]The data began on December 13, 1995, since the inception of the TEJ.

Table 17.2 Unit root tests for the six time series

Test	$nasr_t$	r_t	Δe_t	$qfiibs_t$	$dicbs_t$	$rtfbs_t$
PP	-44.04^*	-42.76^*	-43.32^*	-28.47^*	-31.12^*	-29.90^*

The sample period starts from December 13, 1995, to May 13, 2004, a total of 1989 observations. $qfiibuy_t$ and $qfiisell_t$ = purchases and sales by qualified foreign institutional investors, and $qfiibs_t = qfiibuy_t\text{-}qfiisell_t$. $dicbuy_t$ and $dicsell_t$ = purchases and sales by domestic investment companies, and $dicbs_t = dicbuy_t\text{-}dicsell_t$. $rtfbuy_t$ and $rtfsell_t$ = purchases and sales by registered trading firms, and $rtfbs_t = rtfbuy_t\text{-}rtfsell_t$. $nasr_t$ are the NASDAQ index returns. r_t is the TSE index return. Δe_t is changes in the NT/USD exchange rate. $qfiibs_t$ is net purchases by *qfii*. $dicbs_t$ is net purchases by *dic*, and $rtfbs_t$ is net purchases by *rtf*. y denotes the level of the variable. Δy denotes the first difference of the variable. * denotes statistical significance at 1 % level

Table 17.3 Summary statistics for the net institutional purchases, stock returns and NT/USD currency returns

	Mean	Median	Maximum	Minimum	Std. dev.
Δe_t	0.0105	0.0000	3.4014	−2.9609	0.3331
r_t	0.0083	−0.0438	8.5198	−12.6043	1.7839
$nasr_t$	0.0302	0.1300	13.2546	−10.4078	2.0210
$qfiibuy_t$	5,386.30	4,130	31,415	151	4,513.78
$qfiisell_t$	4,675.17	3,601	44,000	74	4,027.41
$qfiibs_t$	711.11	372	19,408	−23772	3,171.87
$dicbuy_t$	3,191.66	2,924	14,980	192	1,643.79
$dicsell_t$	3,306.13	3,096	11,854	141	1,556.99
$dicbs_t$	−114.49	−120	10,070	−8,876	1,277.43
$rtfbuy_t$	1,814.00	1,418	10,972	49	1,433.95
$rtfsell_t$	1,837.15	1,504	18,024	32	1,384.79
$rtfbs_t$	−23.15	−36	6,379	−11,177	934.55

For variable definitions, see Table 17.2

$$\Delta e_t = (\log e_t - \log e_{t-1}) \times 100\%,$$
$$r_t = (\log p_t - \log p_{t-1}) \times 100\%$$

$$nasr_t = (\log nasp_t - \log nasp_{t-1}) \times 100\%$$

Net foreign purchases ($qfiibs_t$) are computed as the daily purchases ($qfiibuy_t$) less sales ($qfiisell_t$) of Taiwan stocks by foreigners. Similarly, net *dic* purchases ($dicbs_t$) are computed as the daily purchases ($dicbuy_t$) less sales ($dicsell_t$) of Taiwan stocks by *dic*, and net *rtf* purchases ($rtfbs_t$) are computed as the daily purchases ($rtfbuy_t$) less sales ($rtfsell_t$) of Taiwan stocks by *rtf*. The VAR analysis used here depends on whether the time series are stationary; hence, a unit root test is to be performed in advance to avoid spurious regression. The Phillips and Perron test is applied and the results are illustrated in Table 17.2.

The Phillips and Perron test results indicate that all time series are statistically significant at 1 % level. There is no further differencing needed before applying VAR. Table 17.3 presents the summary statistics for the time series used in this paper.

The average percentage change in the exchange rate on the daily basis is 0.011 %; the average daily TSE return is 0.0083 %; and the average daily NASDAQ return equals 0.0302 %. Overall, *qfii* are net purchasers on average and two other domestic institutional investors are net sellers of equity over the sample period, reflecting different trading strategies adopted by foreign and domestic institutional investors. Such distinct trading activities among institutional investors can also be seen in Fig. 17.1.

Figure 17.1 presents the cumulative net purchases and daily net purchases by *qfii*, *dic*, and *rtf* and how they are associated with the TSE returns, NASDAQ returns, and NT/USD exchange rate. Over the entire period, the cumulative net purchases by *qfii* suggest an upward trend in general, while those of *dic* and *rtf* tend to present a downward trend. Overall, the NASDAQ index is more volatile than the TSE index (1995/12/13 = 100), and it seems that there exists some correlation between the two indices. During the Asia financial crisis in 1997, the NT/USD exchange rate suffered a great upward swing (depreciation of New Taiwan Dollar) followed by a slight downward slide in 1999 and then rose again from 2002 onwards. Over the sample period, the volatility of net purchases by foreigners seemed to have increased since 2002. As for the relationship between net purchases by institutions and stock returns, no clear correlation could be detected as shown in Fig. 17.1. To grasp a better understanding on their linkages, the contemporaneous correlation of net purchases by the three types of institutional investors, stock returns, and currency returns are displayed in Table 17.4.

An inspection of Table 17.4 points out that returns on the NT/USD exchange rate (currency returns) are negatively correlated with both the TSE returns and net purchases by the three types of institutional investors, especially by foreign investors. Such relations are very much in line with the expectation. When stock prices rise following the influx of foreign capital, the local currency is expected to appreciate to a degree and as such negative correlations among them is expected. In addition, we find that returns on the TSE and NASDAQ are positively correlated. It is noteworthy that there exists a positive contemporaneous correlation between net purchases by the three types of institutional investors and the TSE returns with the correlation coefficients ranging from about 0.3 ($qfiibs_t$ and r_t) to 0.4419 ($rtfbs_t$ and r_t). In short, this finding largely echoes the previous results (e.g., Froot et al. 2001; Karolyi 2002).

The greatest correlation between institutional trading and NASDAQ returns is that of $rtfbs_t$ and $nasr_t$ at 0.0938 followed by $dicbs_t$ and $qfiibs_t$, respectively. Owing to the time difference, the TSE returns may be influenced by NASDAQ index returns. If $nasr_{t-1}$ is used instead, a higher correlations between $nasr_{t-1}$ and net purchases by institutions are found: 0.3441 for $qfiibs_t$, 0.2369 for $dicbs_t$, and 0.1428 for $rtfbs_t$, respectively. It implies that previous day's NASDAQ returns exert a greater impact on net purchases by each institutional investor in the TSE than do the current NASDAQ returns.

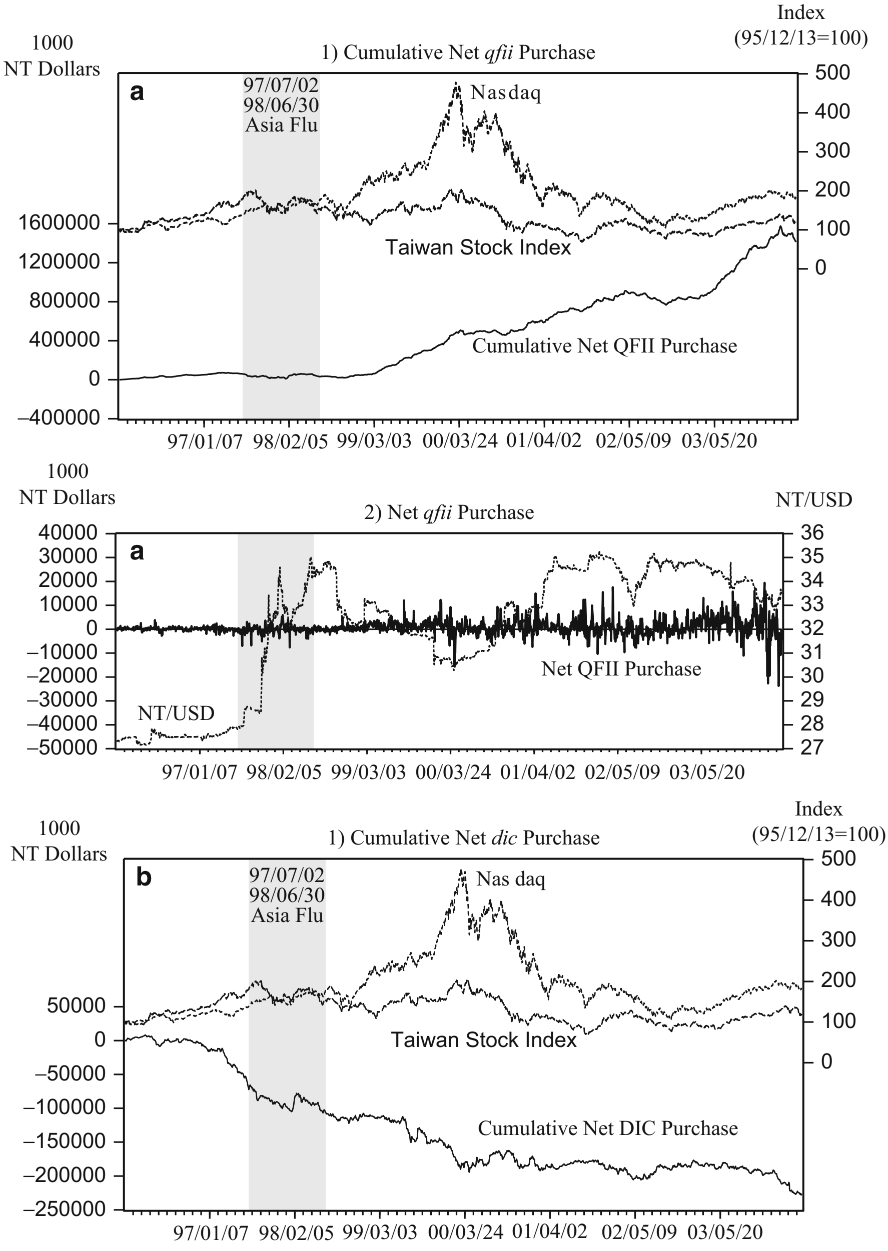


Fig. 17.1 (continued)

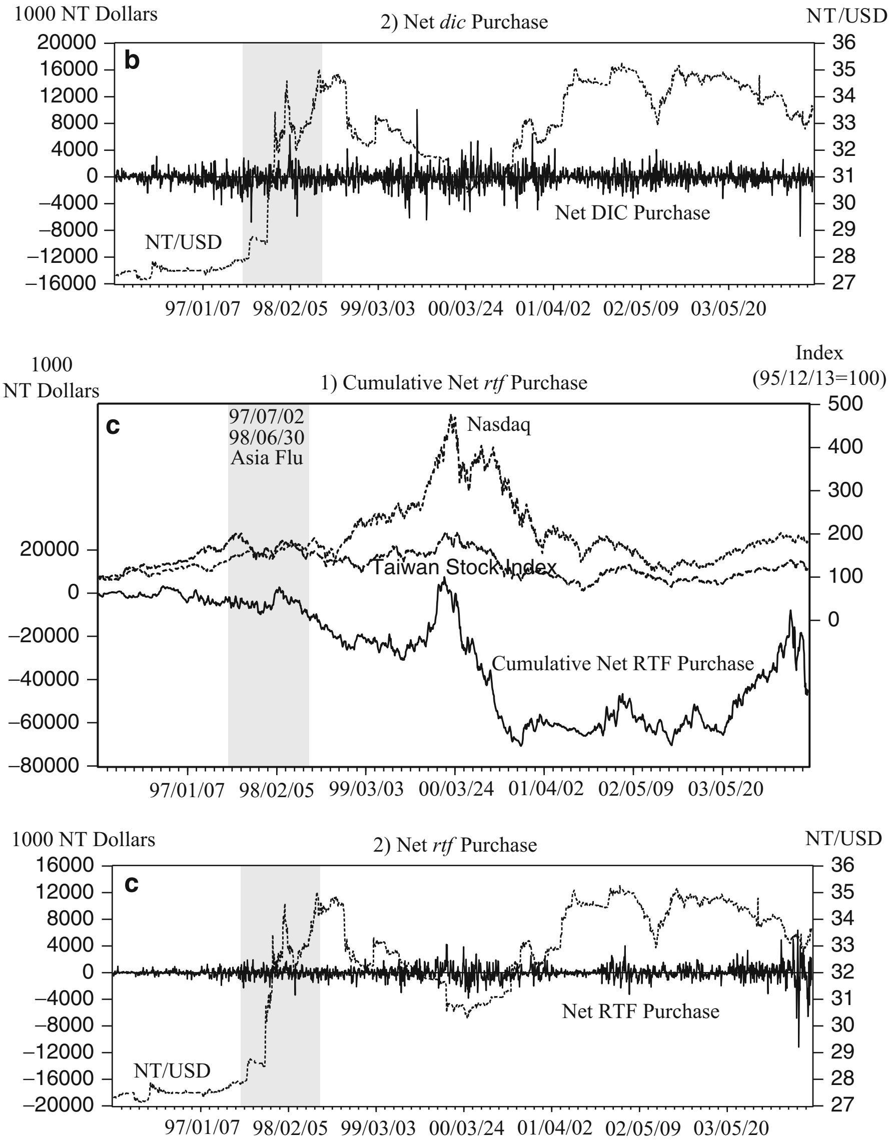


Fig. 17.1 Trends of cumulative net purchase, net purchases by the three types of institutions, Taiwan stock prices, NASDAQ index, and the NT/USD exchange rate. **a-1** Cumulative net *qfii* purchase, **a-2** Net *qfii* purchase, **b-1** Cumulative net *dic* purchase, **b-2** Net *dic* purchase, **c-1** Cumulative net *rtf* purchase, **c-2** Net *rtf* purchase. Notes: *qfii* qualified foreign institutional investors, *dic* domestic investment companies, *rtf* registered trading firms

Table 17.4 Correlation matrix of net purchases by institutions, stock returns, and exchange rate changes

	Δe_t	r_t	$nasr_t$	$qfiibs_t$	$dicbs_t$	$rtfbs_t$
Δe_t	1.0000					
r_t	−0.1338	1.0000				
$nasr_t$	0.0124	0.1322	1.0000			
$qfiibs_t$	−0.1270	0.2976	0.0482	1.0000		
$dicbs_t$	−0.0991	0.3750	0.0586	0.2400	1.0000	
$rtfbs_t$	−0.0767	0.4419	0.0938	0.3559	0.3779	1.0000

See also Table 17.3

17.3 Lead-Lag Relation Among the Three Groups of Institutional Investors in the TSE

17.3.1 The Unrestricted VAR Model

Note that prices of many Taiwanese electronics securities are affected by the NASDAQ returns and hence foreign portfolio inflows may induce fluctuations of exchange rate. To investigate interactions emanated from the three types of institutional investors and the relationship between institutional trading activity and stock returns in Taiwan, we employ a six-variable VAR model using the NASDAQ returns ($nasr_t$), currency returns (Δe_t), TSE returns (r_t), net foreign purchases ($qfiibs_t$), net *dic* purchases ($dicbs_t$), and net *rtf* purchases ($rtfbs_t$) as the underlying variables. We attempt to answer the issues pertaining to (i) the interaction among trading activities of the three types of institutions and (ii) the relationship between stock returns and institutional trading. First, we propose a six-variable unrestricted VAR model shown below:

$$\begin{bmatrix} nasr_t \\ \Delta e_t \\ r_t \\ qfiibs_t \\ dicbs_t \\ rtfbs_t \end{bmatrix} = \begin{bmatrix} \phi_{11}(L) & \phi_{12}(L) & \phi_{13}(L) & \phi_{14}(L) & \phi_{15}(L) & \phi_{16}(L) \\ \phi_{21}(L) & \phi_{22}(L) & \phi_{23}(L) & \phi_{24}(L) & \phi_{25}(L) & \phi_{26}(L) \\ \phi_{31}(L) & \phi_{32}(L) & \phi_{33}(L) & \phi_{34}(L) & \phi_{35}(L) & \phi_{36}(L) \\ \phi_{41}(L) & \phi_{42}(L) & \phi_{43}(L) & \phi_{44}(L) & \phi_{45}(L) & \phi_{46}(L) \\ \phi_{51}(L) & \phi_{52}(L) & \phi_{53}(L) & \phi_{54}(L) & \phi_{55}(L) & \phi_{56}(L) \\ \phi_{61}(L) & \phi_{62}(L) & \phi_{63}(L) & \phi_{64}(L) & \phi_{65}(L) & \phi_{66}(L) \end{bmatrix} \times \begin{bmatrix} nasr_{t-1} \\ \Delta e_{t-1} \\ r_{t-1} \\ qfiibs_{t-1} \\ dicbs_{t-1} \\ rtfbs_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \\ \varepsilon_{4t} \\ \varepsilon_{5t} \\ \varepsilon_{6t} \end{bmatrix} \quad (1)$$

where $\phi_{ij}(L)$ is the polynomial lag of the jth variable in the ith equation. To investigate the lead-lag relation among three types of institutional investors, we need to test the hypothesis that each off-diagonal element in the sub-matrix $\begin{bmatrix} \phi_{44}(L) & \phi_{45}(L) & \phi_{46}(L) \\ \phi_{54}(L) & \phi_{55}(L) & \phi_{56}(L) \\ \phi_{64}(L) & \phi_{65}(L) & \phi_{66}(L) \end{bmatrix}$ is zero.

On the other hand, to determine whether the TSE returns of the previous day lead net purchases by the three types of institutional investors, we test the hypothesis that each polynomial lag in the vector $[\,\phi_{43}(L)\phi_{53}(L)\phi_{63}(L)\,]'$ is zero. Conversely, if we want to determine whether previous day's net purchases by institutional investors lead current market returns, we test the hypothesis that each element in the vector $[\phi_{34}(L)\phi_{35}(L)\phi_{36}(L)]$ is zero. Before applying the VAR model, an appropriate lag structure needs to be specified. A 3-day lag is selected based on the Akaike information criterion (AIC). Table 17.5 presents the lead-lag relation among the six time series using block exogeneity tests.

The $\begin{bmatrix} \phi_{44}(L) & \phi_{45}(L) & \phi_{46}(L) \\ \phi_{54}(L) & \phi_{55}(L) & \phi_{56}(L) \\ \phi_{64}(L) & \phi_{65}(L) & \phi_{66}(L) \end{bmatrix}$ block represents the potential interaction among three types of institutional investors. The results indicate that net foreign purchases lead net *dic* purchases and the dynamic relationship between these two variables can be provided by the impulse response function (IRF) in Fig. 17.2a. Clearly, a one-unit standard error shock to net foreign purchases leads to an increase in net *dic* purchases, but this effect dissipates quickly by Period 2. Figure 17.2b, c indicate a feedback relation between net purchases by *qfii* and *rtf*. A one-unit standard error shock to net foreign purchases results in a positive response to net *rtf* purchases over the next two periods, which become negative in Period 3, followed by a positive response again after Period 4. Furthermore, a one-unit standard error shock to net *rtf* purchases also gives rise to an increase in net foreign purchases, which decays slowly over ten-period horizon.

Figure 17.2d shows that net purchases by *dic* lead net *rtf* purchases. A one-unit standard error shock to net *dic* purchases produces an increase in net *rtf* purchases in the first three periods and then declines thereafter. Overall, these impulse responses suggest that previous day's net foreign purchases exert a noticeable impact on net *rtf* purchases, while previous day's net *rtf* purchases also has an impact on net foreign purchases. It implies that not only do foreign capital flows affect the trading activity of domestic institutional investors but also the relation is not unidirectional. To be specific, there is a feedback relation between net *rtf* purchases and net foreign purchases.

As for the effect of the three types of institutional trading activity on stock returns in the TSE, Table 17.5 reveals that net *dic* purchases on previous day lead the TSE returns. We can also see in Fig. 17.2e that after Period 3, net *dic* purchases exert a negative (and thus destabilizing) effect on market returns, while the other two institutional investors do not have such an effect over the sample period. Examining the relationship between market returns and trading activity of the three types of institutional investors, we find that either the net foreign purchases or net *dic* purchases on previous trading day are affected by the previous day's TSE

Table 17.5 Results of Granger causality tests using the unrestricted VAR models

All	$nasr_t$	Δe_t	r_t	$qfiibs_t$	$dicbs_t$	$rtfbs_t$
$\sum_{i=1}^{3} nasr_{t-i}$		5.23 (0.16)	103.73^{*} (0.00)	329.58^{*} (0.00)	103.46^{*} (0.00)	40.92^{*} (0.00)
$\sum_{i=1}^{3} \Delta e_{t-i}$	2.77 (0.43)		2.66 (0.45)	2.01 (0.57)	5.98 (0.11)	4.66 (0.20)
$\sum_{i=1}^{3} r_{t-i}$	0.04 (1.00)	17.43^{*} (0.00)		14.22^{*} (0.00)	97.30^{*} (0.00)	0.27 (0.97)
$\sum_{i=1}^{3} qfiibs_{t-i}$	12.67^{*} (0.01)	5.81 (0.12)	4.38 (0.22)		9.55^{**} (0.02)	33.95^{*} (0.00)
$\sum_{i=1}^{3} dicbs_{t-i}$	9.38^{**} (0.02)	0.85 (0.84)	15.97^{*} (0.00)	5.83 (0.12)		12.70^{*} (0.01)
$\sum_{i=1}^{3} rtfbs_{t-i}$	0.66 (0.88)	0.81 (0.85)	2.33 (0.51)	52.84^{*} (0.00)	0.85 (0.84)	

*, **, and *** denote statistical significance at 1 %, 5 %, and 10 % levels, respectively. Values in parentheses are p values. The optimal lag length of three is selected based on the Akaike information criterion

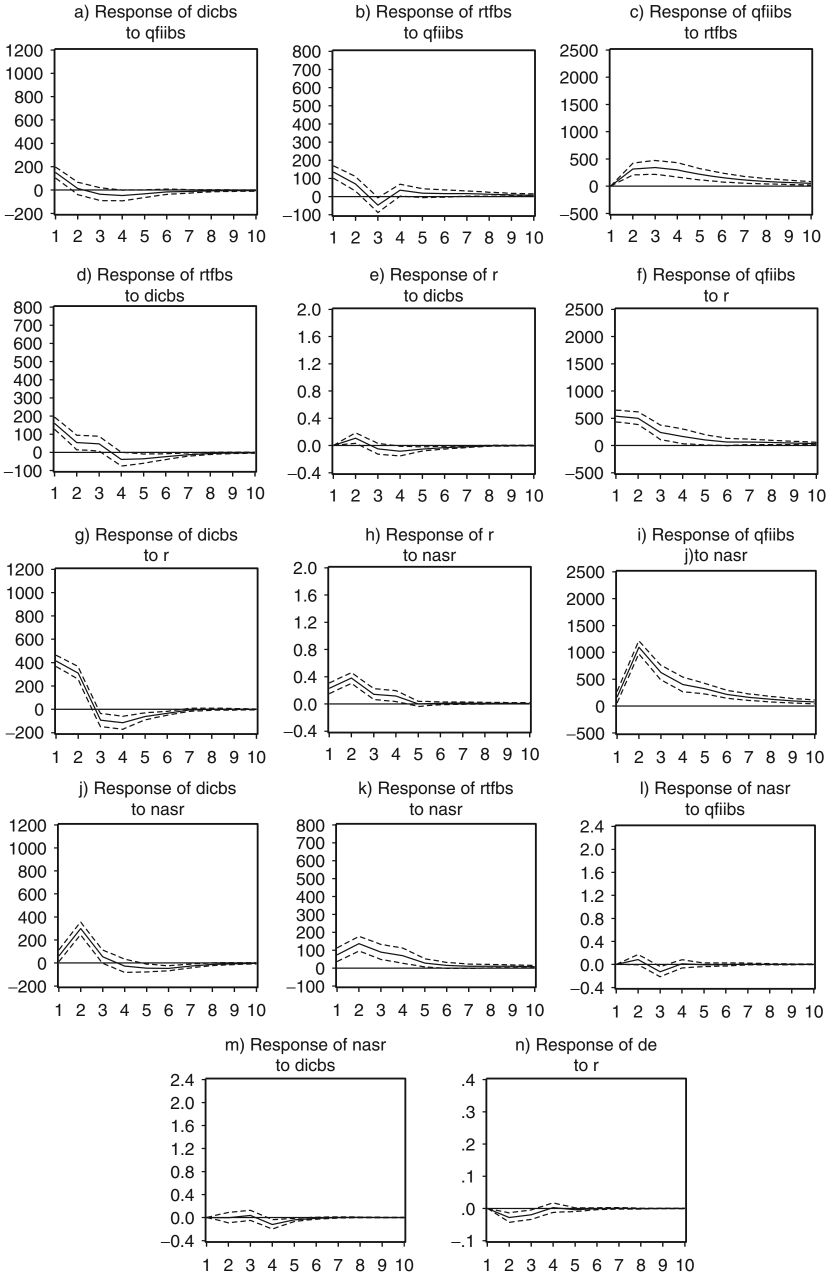


Fig. 17.2 Impulse responses to innovations in the unrestricted VAR models (up to ten periods). Notes: Solid lines represent response path and dotted lines are bands for the 95 % confidence interval around response coefficients. r, TSE returns; nasr, NASDAQ returns; de, exchange rate changes; qfiibs, net *qfii* purchases; dicbs, net *dic* purchases; rtfbs, net *rtf* purchases

returns. Moreover, the IRFs in Fig. 17.2f, g also reveal a significant positive relation between TSE returns and net foreign purchases up to four periods and a significant positive relation between TSE returns and net *dic* purchases for two periods, which becomes negative after Period 3. In other words, foreign investors in the TSE engage in positive-feedback trading, while those of the *dic* tend to change their strategy and adopt negative-feedback trading after Period 3.

On the other hand, Table 17.5 indicates that previous day's NASDAQ returns significantly affect both current TSE returns and net purchases by the three types of institutional investors. The impulse response in Fig. 17.2h–k also confirms that previous day's NASDAQ returns are positively related to both current returns and net purchases by these institutional investors, with the exception that a negative relation between net *dic* purchases and previous day's NASDAQ returns is found after Period 4. Such results are much in sync with the expectation since the largest sector that comprises the TSE-weighted stock index is the electronics industry to which many listed companies on NASDAQ have a strong connection. In addition, although previous day's net *qfii* and *dic* purchases also lead the NASDAQ returns, we find that no significant relation exists except for Period 4 with a significantly negative relation between them.

The liberalization of Taiwan's stock market has ushered in significant amount of short-term inflows and outflows of foreign capital, which have induced fluctuations in the exchange rate. As is seen from Table 17.5, the TSE returns lead currency returns and it appears that the initial significant effect of stock returns on currency returns is negative for the first three periods and then turns to be significantly positive thereafter (Fig. 17.2n). Given that foreign investors are positive-feedback traders, the capital inflows is expected to grow in order to increase their stakes in TSE securities when stock prices rise. Consequently, NT/USD is expected to appreciate.

17.3.2 The Structural VAR Model

The unrestricted VAR model does not consider the effect of current returns on net purchases by institutions. The prior study by Griffin et al. (2003) includes current returns in the institutional imbalance equation and finds a strong contemporaneous positive relation between institutional trading and stock returns.[6] Therefore, to further examine the relationship between institutional trading and stock returns, we introduce the current TSE returns (r_t) in the net purchases equations of the three types of institutional investors and reestimate the VAR model before conducting the corresponding block exogeneity tests. Table 17.6 presents the estimation results.

As indicated in the last row of Table 17.6, we find evidence of a strong contemporaneous correlation between current returns and net institutional

[6]The main purpose of this paper is to improve our understanding on the interaction among institutional investors and the relationship between institutional trading and stock returns. The following discussion will therefore focus on these two issues.

Table 17.6 Results of Granger causality tests using the structural VAR models

SVAR	$nasr_t$	Δe_t	r_t	$qfiibs_t$	$dicbs_t$	$rtfbs_t$
$\sum_{i=1}^{3} nasr_{t-i}$		5.23 (0.16)	103.73* (0.00)	265.94* (0.00)	54.67* (0.00)	6.75*** (0.08)
$\sum_{i=1}^{3} \Delta e_{t-i}$	2.77 (0.43)		2.66 (0.45)	1.85 (0.60)	7.06*** (0.07)	4.45 (0.22)
$\sum_{i=1}^{3} r_{t-i}$	1.26 (0.74)	20.63* (0.00)		98.84* (0.00)	264.71* (0.00)	75.05* (0.00)
$\sum_{i=1}^{3} qfiibs_{t-i}$	12.67* (0.01)	5.81 (0.12)	4.38 (0.22)		14.05* (0.00)	36.64* (0.00)
$\sum_{i=1}^{3} dicbs_{t-i}$	9.38** (0.02)	0.85 (0.84)	15.97* (0.00)	5.09 (0.17)		11.09* (0.01)
$\sum_{i=1}^{3} rtfbs_{t-i}$	0.66 (0.88)	0.81 (0.85)	2.33 (0.51)	50.56* (0.00)	0.83 (0.84)	
r_t				332.55* [10.91]	247.47* [18.10]	213.66* [21.43]

The structural VAR model includes current TSE returns in the institutional trading equations to take the contemporaneous correlations into consideration. See also Table 17.5 for definitions

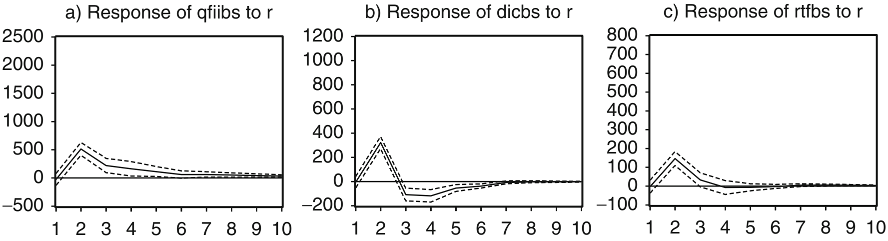


Fig. 17.3 Selected impulse responses to innovations up to ten periods in the structural VAR models. Notes: See also Fig. 17.2

purchases, which confirms the finding by the previous research. Moreover, as shown in Table 17.5, we find no evidence that past returns lead net *rtf* purchases when the unrestricted VAR model is used. In contrast to it, when the contemporaneous impact of stock returns on net institutional purchases is considered, we find that past returns also lead net *rtf* purchases as well as net purchases by *qfii* and *dic* when the structural VAR model is used (Table 17.6). In other words, net purchases by the three types of institutions are affected by past stock returns as was evidenced by previous studies. The corresponding impulse response relations are presented in Fig. 17.3.

Comparing the impulse response relations in Figs. 17.2 and 17.3, it is clear that when the impact of current returns on net institutional purchases is considered, a one-unit standard error shock from r_t does not produce a positive impulse response in institutional trading until Period 2.[7] The responses of foreign investors are rather distinct from those of domestic institutional investors after Period 3. In general, a sustained positive response from foreign investors is observed, while a negative response is witnessed for *dic* and sometimes, an insignificant response for *rtf* manifests itself after Period 3.

17.3.3 The Threshold VAR Analysis

We pool all the data together when estimating either the unrestricted or restricted VAR model; however, the trading activity of institutional investors may depend on whether stock prices rise or fall.[8] A small economy like Taiwan also depends to a large degree on the sign of NASDAQ index returns. Consequently, to investigate institutional trading under distinct regimes based on market returns, we use the multivariate threshold autoregression (MVTAR) model proposed by Tsay (1998)

[7]Figure 17.2f, g show significantly positive responses of $qfiibs_t$ and $dicbs_t$ to r_t in Period 1 if the impact of current returns is not considered.

[8]Recall that both the positive-feedback and negative-feedback trading are associated with the sign of market returns on the previous trading day.

Table 17.7 The $C(d)$ statistic

Threshold variable	Statistic
$nasr_{t-1}$	195.08 (0.00)
r_{t-1}	136.22 (0.08)

Values in parentheses are p values. The delay (d) is assumed to be one

to test the relevant hypotheses. Let $\mathbf{y}_t = [nasr_t,\ \Delta e_t,\ r_t,\ qfiibs_t,\ dicbs_t,\ rtfbs_t]'$ be a 6×1 vector and the MVTAR model can be described as

$$\mathbf{y}_t = \left(\phi_{0,1} + \sum_{i=1}^{p} \phi_{i,1} y_{t-i}\right) \cdot [1 - I(z_{t-d} > c)] + \left(\phi_{0,2} + \sum_{i=1}^{p} \phi_{i,2} y_{t-i}\right) \cdot I(z_{t-d} > c) + \varepsilon_t \tag{2}$$

where $E(\varepsilon) = 0$, $E(\varepsilon\varepsilon') = \mathbf{\Sigma}$, and $I(\cdot)$ is an index function, which equals 1 if the relation in the bracket holds. It equals zero otherwise. z_{t-d} is the threshold variable with a delay (lag) d.

In order to explore whether institutional trading activity would change during different domestic and foreign market return scenarios, the potential threshold variables used are r_{t-1} and $nasr_{t-1}$.[9] Before estimating Eq. 2, we need to test for possible potential nonlinearity (threshold effect) in this equation. Tsay (1998) suggests using the arranged regression concept to construct the $C(d)$ statistic to test the hypothesis $H_o{:}\phi_{i,1} = \phi_{i,2}$, $i = 0, \ldots p$. If H_0 can be rejected, it implies that there exists the nonlinearity in data with z_{t-d} as the threshold variable. Tsay (1998) proves that $C(d)$ is asymptotically a chi-square random variable with $k(pk + 1)$ degrees of freedom, where p is the lag length of the VAR model and k is the number of endogenous variables y_t.[10] Table 17.7 presents the estimation results of the $C(d)$ statistic.

As shown in Table 17.7, the null hypothesis H_0 is rejected using either past returns on the TSE or NASDAQ, suggesting that our data exhibit nonlinear threshold effect. Theoretically, one needs to rearrange the regression based on the size of the threshold variable z_{t-d} before applying a grid search method to find the optimal threshold value c^*. Nonetheless, our goal is to know whether the institutional trading behavior depends on the sign of market returns, as such the threshold is set to zero in a rather arbitrary way.[11] Table 17.8 lists the results of block

[9] Here, we assume that net purchases by institutions are only affected by market returns on the previous trading day.

[10] For more details see Tsay (1998).

[11] A previous study by Sadorsky (1999) also splits data into two regimes based on the sign of the variable to discuss whether variables used would change their behaviors under different regimes.

Table 17.8 Results of Granger causality tests using the MVTAR models with threshold variable r_{t-1}

$r_{t-1} < 0$	$nasr_t$	Δe_t	r_t	$qfiibs_t$	$dicbs_t$	$rtfbs_t$
$\sum_{i=1}^{3} nasr_{t-i}$		3.07 (0.38)	84.21^{*} (0.00)	167.80^{*} (0.00)	53.71^{*} (0.00)	30.12^{*} (0.00)
$\sum_{i=1}^{3} \Delta e_{t-i}$	0.51 (0.92)		2.01 (0.57)	1.00 (0.80)	3.11 (0.37)	0.65 (0.89)
$\sum_{i=1}^{3} r_{t-i}$	2.28 (0.52)	12.41^{*} (0.01)		8.57^{**} (0.04)	22.40^{*} (0.00)	3.14 (0.37)
$\sum_{i=1}^{3} qfiibs_{t-i}$	7.75^{**} (0.05)	3.04 (0.39)	4.33 (0.23)		9.79^{**} (0.02)	32.73^{*} (0.00)
$\sum_{i=1}^{3} dicbs_{t-i}$	3.83 (0.28)	0.46 (0.93)	4.92 (0.18)	2.38 (0.50)		4.62 (0.20)
$\sum_{i=1}^{3} rtfbs_{t-i}$	1.23 (0.75)	1.62 (0.65)	1.62 (0.66)	19.99^{*} (0.00)	2.83 (0.42)	
$\sum_{i=1}^{3} nasr_{t-i}$		3.72 (0.29)	27.57^{*} (0.00)	168.61^{*} (0.00)	49.94^{*} (0.00)	12.25^{*} (0.01)

$\sum_{i=1}^{3} \Delta e_{t-i}$	5.97 (0.11)		1.78 (0.62)	2.73 (0.44)	5.16 (0.16)	7.01^{***} (0.07)
$\sum_{i=1}^{3} r_{t-i}$	0.88 (0.83)	13.00^{*} (0.00)		3.02 (0.39)	43.44^{*} (0.00)	1.68 (0.64)
$\sum_{i=1}^{3} qfiibs_{t-i}$	4.83 (0.18)	5.87 (0.12)	1.85 (0.60)		5.72 (0.13)	5.16 (0.16)
$\sum_{i=1}^{3} dicbs_{t-i}$	8.99^{**} (0.03)	2.04 (0.56)	13.17^{*} (0.00)	3.90 (0.27)		12.54^{*} (0.01)
$\sum_{i=1}^{3} rtfbs_{t-i}$	0.22 (0.97)	0.58 (0.90)	2.53 (0.47)	39.46^{*} (0.00)	2.15 (0.54)	

The results are premised on the condition that when previous day's TSE returns are positive. See also Table 17.5

exogeneity tests for the lead-lag relation in the $r_{t-1} < 0$ and $r_{t-1} \geq 0$ regimes, respectively.

The interaction among institutional investors is depicted in Table 17.8: Current net purchases by foreign investors affect that by domestic institutions when previous day's TSE returns are negative. Note that no such relation is evidenced when previous day's TSE returns are positive. A feedback relation between *rtf* and *qfii* is observed when r_{t-1} is positive or negative. However, *dic* is found to lead *rtf* only when r_{t-1} is positive. Such results reveal different institutional trading strategies under distinct return regimes. The demonstration effect – previous day's net foreign purchases have on domestic institutions using the unrestricted VAR model – seems to surface only when previous day's market returns are negative. Therefore, it may produce misleading results if we fail to consider the sign of previous returns.

The impulse responses in Fig. 17.4 illustrate that the responses of *dic* and *rtf* from the *qfii* shock are quite similar to the ones in Fig. 17.2f, g.[12] As for the impact of previous day's market returns on current net purchases by institutions, it can be shown via the MVTAR model that market returns lead net purchases by *qfii* and *dic* when previous day's market returns are negative, which is consistent with the finding using the one-regime VAR model. When previous day's market returns are positive, market returns lead net purchases by the *dic* only. Obviously, returns have more influence on net institutional purchases when previous day's returns were negative. In addition, we find that net *dic* purchases on the previous day may affect current returns when the one-regime VAR model is used. Actually, the MVTAR analysis reveals that such a relation emerges only when previous day's market returns are positive. The impulse responses depicted in Fig. 17.4 (Panel B) demonstrate that a one-unit standard error shock to net *dic* purchases produces an increase in market returns in Period 2, and then they turns to be negative after Period 4, a result similar to those using the one-regime VAR model.

Among the listed companies on the TSE, the electronics sector has the largest market share, which accounts for more than 60 % of all trades. This being the case, Taiwan's stock market is closely related to the NASDAQ index as is evidenced using the conventional VAR model. To further investigate whether the interaction among institutions and the relationship between institutional trading and stock returns are affected by the sign of previous day's NASDAQ index return, $nasr_{t-1}$ is used as the threshold variable. That is, block Granger causality tests are performed by splitting our data into two regimes based on the sign of the variable $nasr_{t-1}$. Table 17.9 reports the results.

The results indicate that net *qfii* purchases lead to that of two domestic institutional investors regardless of the sign of previous day's NASDAQ returns. Moreover, when $nasr_{t-1} < 0$, net *rtf* purchases lead net *qfii* purchases, which is in line with that using the one-regime VAR model. However, when $nasr_{t-1} \geq 0$, the net purchases by either *dic* or *rtf* lead net *qfii* purchases, and net *qfii* purchases lead net purchases by either *dic* or *rtf*. In other words, we find strong evidence of a feedback

[12]To economize space, only relevant impulse responses are presented here; the remaining are available upon request.

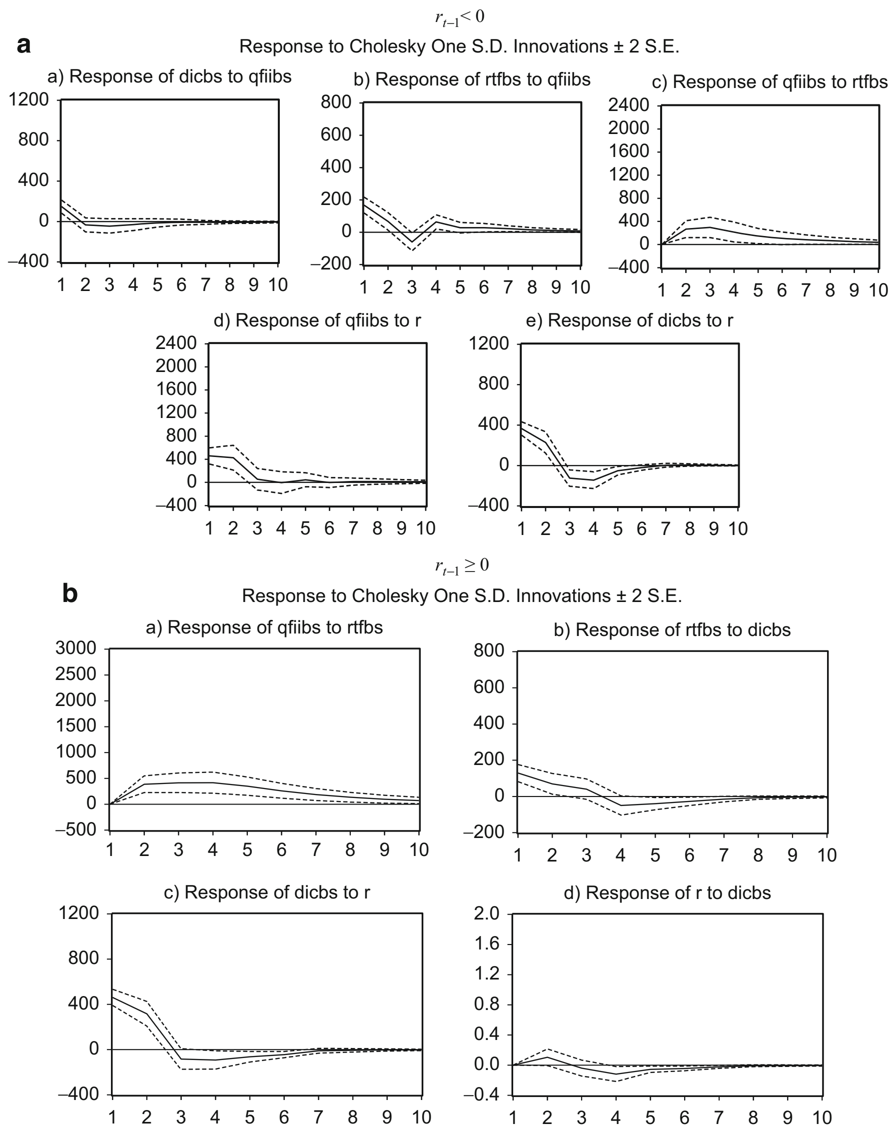


Fig. 17.4 Selected impulse responses to innovations up to ten periods in the MVTAR models. (**a**) $r_{t-1} < 0$, (**b**) $r_{t-1} \geq 0$. Notes: See also Fig. 17.2

relation between net foreign purchases and two domestic net purchases when previous day's NASDAQ returns are positive. The results pertaining to the impact of previous day's returns on institutional trading parallel those using the unrestricted VAR model: Previous day's returns have an impact on the net purchases by *qfii* and *dic*, but not on net purchases by *rtf* regardless of the sign of previous day's NASDAQ returns. As for the impact of net institutional purchases

Table 17.9 Results of Granger causality tests using the MVTAR models with threshold variable $nasr_{t-1}$

$nasr_{t-1} \geq 0$	$nasr_t$	Δe_t	r_t	$qfiibs_t$	$dicbs_t$	$rtfbs_t$
$\sum_{i=1}^{3} nasr_{t-i}$		0.92 (0.82)	36.02^{*} (0.00)	50.15^{*} (0.00)	15.42^{*} (0.00)	12.53^{*} (0.01)
$\sum_{i=1}^{3} \Delta e_{t-i}$	0.50 (0.92)		4.95 (0.18)	0.29 (0.96)	3.81 (0.28)	3.08 (0.38)
$\sum_{i=1}^{3} r_{t-i}$	2.13 (0.55)	8.09^{**} (0.04)		8.54^{**} (0.04)	48.61^{*} (0.00)	1.80 (0.62)
$\sum_{i=1}^{3} qfiibs_{t-i}$	4.62 (0.20)	9.55^{**} (0.02)	3.25 (0.35)		7.43^{***} (0.06)	14.74^{*} (0.00)
$\sum_{i=1}^{3} dicbs_{t-i}$	6.94^{***} (0.07)	1.05 (0.79)	7.97^{**} (0.05)	2.77 (0.43)		9.64^{*} (0.02)
$\sum_{i=1}^{3} rtfbs_{t-i}$	0.12 (0.99)	1.48 (0.69)	1.33 (0.72)	28.94^{*} (0.00)	1.93 (0.59)	
$\sum_{i=1}^{3} nasr_{t-i}$		2.64 (0.45)	29.58^{*} (0.00)	115.60^{*} (0.00)	59.06^{*} (0.00)	16.76^{*} (0.00)

$\sum_{i=1}^{3}\Delta e_{t-i}$	6.72^{***} (0.08)		5.21 (0.16)	3.85 (0.28)	5.63 (0.13)	5.73 (0.13)
$\sum_{i=1}^{3} r_{t-i}$	1.49 (0.68)	27.44^{*} (0.00)		14.75^{*} (0.00)	55.66^{*} (0.00)	0.27 (0.96)
$\sum_{i=1}^{3} qfiibs_{t-i}$	11.39^{*} (0.01)	0.60 (0.90)	7.76^{**} (0.05)		7.00^{***} (0.07)	33.48^{*} (0.00)
$\sum_{i=1}^{3} dicbs_{t-i}$	3.48 (0.32)	3.25 (0.35)	11.08^{*} (0.01)	6.74^{***} (0.08)		3.47 (0.32)
$\sum_{i=1}^{3} rtfbs_{t-i}$	1.82 (0.61)	2.59 (0.46)	4.12 (0.25)	25.09^{*} (0.00)	4.65 (0.20)	

The results are premised on the condition that when previous day's NASDAQ returns are positive. See also Table 17.5

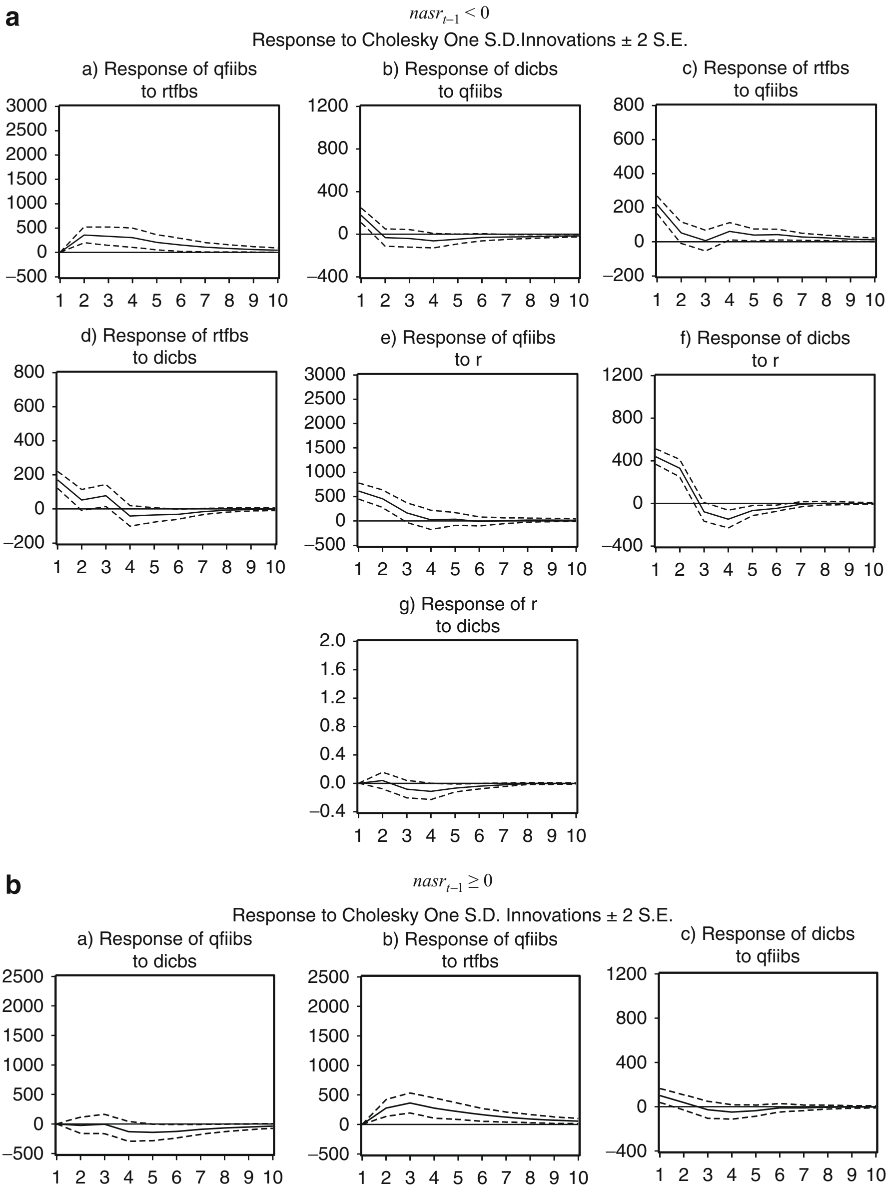


Fig. 17.5 (continued)

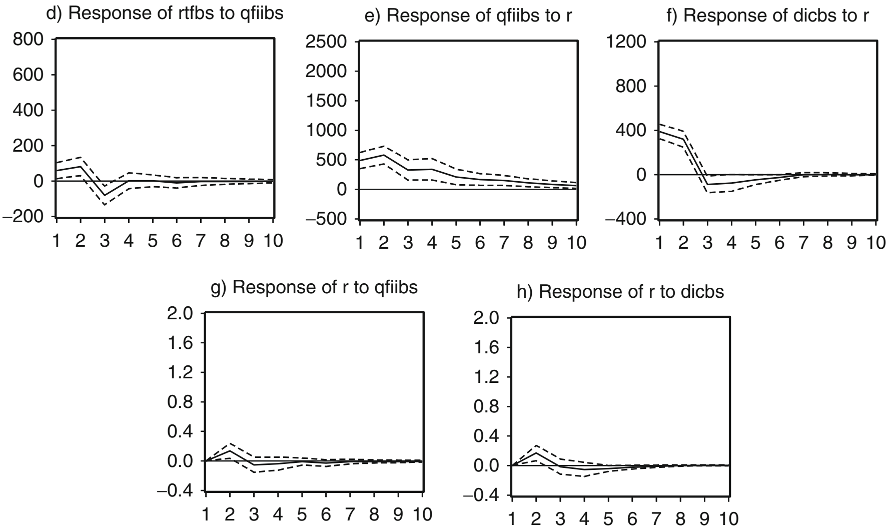


Fig. 17.5 Impulse responses to innovations up to ten periods in the MVTAR models (portion). (**a**) $nasr_{t-1} < 0$, (**b**) $nasr_{t-1} \geq 0$. Notes: See also Fig. 17.2

on previous day's stock returns, only net *dic* purchases still lead stock returns when $nasr_{t-1} < 0$, as was the case in the one-regime model. However, we find that net *qfii* purchases also lead stock returns when $nasr_{t-1} \geq 0$.

The results from Panel B of Fig. 17.5g indicate that a one-unit standard error shock from net *qfii* purchases made on previous days produces a positive response to stock returns in Period 2, but no significantly negative responses are found during other periods. The results of the MVTAR model also capture the phenomenon in the one-regime model: Net *dic* purchases exert a negative impact on market returns. However, such an effect is witnessed when previous day's NASDAQ returns are negative.

17.4 Conclusion

In this paper we investigate whether the trading behavior of foreign investors leads that of Taiwanese institutional investors (i.e., the demonstration effect) and whether institutional trading has a destabilizing effect on the stock market. The reason we select Taiwan in our study is due to her unique role of being gradually opened up to foreign investment and her high stock returns volatility. To provide more information to these issues, this paper has constructed a six-variable VAR model including trading activities of three types of institutional investors, the TSE returns, NASDAQ returns, and currency returns so as to examine the interaction and the dynamic relationship between institutional trading and stock returns using daily data from December 13, 1995 to May 13, 2004.

The results from the conventional unrestricted VAR model indicate that net purchases by foreign investors lead those by domestic institutions (and thus the demonstration effect), while net purchases by *rtf* also lead those by foreign investors. That is, there exists a feedback relation between them. As for the relationship between institutional trading and stock returns, we find that except for *rtf*, both foreign and *dic* net purchases are positively affected by previous day's TSE returns. That is, both the *qfii* and *dic* engage in positive-feedback trading. We also find that previous day's net *dic* purchases first produce a positive and then a negative impact on stock returns. Furthermore, we employ a structural VAR model with the contemporaneous returns included in the three net institutional purchase equations. A comparison of the structural and unrestricted VAR models suggests that the TSE returns positively lead net *rtf* purchases using the structural VAR model, which cannot be observed when the unrestricted VAR model is used. In other words, if the contemporaneous relation between returns and net institutional purchases is taken into account, we find that *rtf* are also positive-feedback traders.

On the other hand, the sign of market returns does affect trading activities of the institutions. As a result, this paper makes use of the MVTAR model introduced by Tsay (1998). By splitting data into two regimes based on the sign of both TSE and NASDAQ returns on the previous trading day, we find that the demonstration effect that foreign investors have on domestic institutions arises only when previous day's TSE returns are negative. In addition, when previous day's TSE returns are negative, stock returns lead both net purchases by *qfii* and *dic*. However, stock returns lead only net *dic* purchases when previous day's TSE returns are positive. Finally, we find the relation that net *dic* purchases lead market returns using the unrestricted VAR model tends to emerge only when previous day's TSE returns are positive.

As for the effect of NASDAQ returns on institutional trading, the results from this paper suggest that when previous day's NASDAQ returns are positive, a feedback relation between net foreign purchases and net domestic purchases prevails. Moreover, it is found that the net *dic* purchases lead the TSE returns, as do the net foreign purchases. The latter, however, exert a positive influence on the TSE returns.

In summary, our results suggest that net foreign purchases do lead net domestic purchases, but more details manifest when the threshold model is applied. When previous day's TSE returns are negative (or previous day's NASDAQ returns are positive), a feedback relation between net foreign purchases and net domestic purchases is observed. It implies that the widespread argument that foreign investors have a demonstration effect on domestic institutions in Taiwan is not entirely correct. In examining the relation between market returns and institutional trading, we find that market returns at least lead net purchases by both *qfii* and *dic* in most cases. Market returns also lead net *rtf* purchases if the relationship of contemporaneous returns and institutional trading is considered. Our analysis also indicates that net *dic* purchases lead market returns and are negatively associated with market returns in Period 4. The results of the MVTAR model suggest that when previous day's NASDAQ returns are positive, net foreign purchases positively lead stock returns and thus will not exert a destabilizing influence on the market.

Appendix 1

The Unrestricted VAR Model

The prices of many Taiwanese electronics securities are affected by the NASDAQ returns and hence foreign portfolio inflows may induce fluctuations of exchange rate. To investigate interactions emanated from the three types of institutional investors and the relationship between institutional trading activity and stock returns in Taiwan, we employ a six-variable VAR model using the NASDAQ returns ($nasr_t$), currency returns (Δe_t), TSE returns (r_t), net foreign purchases ($qfiibs_t$), net *dic* purchases ($dicbs_t$), and net *rtf* purchases ($rtfbs_t$) as the underlying variables. We attempt to answer the issues pertaining to (i) the interaction among trading activities of the three types of institutions and (ii) the relationship between stock returns and institutional trading. First, we propose a six-variable unrestricted VAR model shown below:

$$\begin{bmatrix} nasr_t \\ \Delta e_t \\ r_t \\ qfiibs_t \\ dicbs_t \\ rtfbs_t \end{bmatrix} = \begin{bmatrix} \phi_{11}(L) & \phi_{12}(L) & \phi_{13}(L) & \phi_{14}(L) & \phi_{15}(L) & \phi_{16}(L) \\ \phi_{21}(L) & \phi_{22}(L) & \phi_{23}(L) & \phi_{24}(L) & \phi_{25}(L) & \phi_{26}(L) \\ \phi_{31}(L) & \phi_{32}(L) & \phi_{33}(L) & \phi_{34}(L) & \phi_{35}(L) & \phi_{36}(L) \\ \phi_{41}(L) & \phi_{42}(L) & \phi_{43}(L) & \phi_{44}(L) & \phi_{45}(L) & \phi_{46}(L) \\ \phi_{51}(L) & \phi_{52}(L) & \phi_{53}(L) & \phi_{54}(L) & \phi_{55}(L) & \phi_{56}(L) \\ \phi_{61}(L) & \phi_{62}(L) & \phi_{63}(L) & \phi_{64}(L) & \phi_{65}(L) & \phi_{66}(L) \end{bmatrix}$$

$$\times \begin{bmatrix} nasr_{t-1} \\ \Delta e_{t-1} \\ r_{t-1} \\ qfiibs_{t-1} \\ dicbs_{t-1} \\ rtfbs_{t-1} \end{bmatrix} + \begin{bmatrix} \varepsilon_{1t} \\ \varepsilon_{2t} \\ \varepsilon_{3t} \\ \varepsilon_{4t} \\ \varepsilon_{5t} \\ \varepsilon_{6t} \end{bmatrix}$$

Where $\phi_{ij}(L)$ is the polynomial lag of the *j*th variable in the *i*th equation. To investigate the lead-lag relation among three types of institutional investors, we need to test the hypothesis that each off-diagonal element in the sub-matrix $\begin{bmatrix} \phi_{44}(L) & \phi_{45}(L) & \phi_{46}(L) \\ \phi_{54}(L) & \phi_{55}(L) & \phi_{56}(L) \\ \phi_{64}(L) & \phi_{65}(L) & \phi_{66}(L) \end{bmatrix}$ is zero.

On the other hand, to determine whether the TSE returns of the previous day lead net purchases by the three types of institutional investors, we test the hypothesis that each polynomial lag in the vector $[\phi_{43}(L)\phi_{53}(L)\phi_{63}(L)]'$ is zero. Conversely, if we want to determine whether previous day's net purchases by institutional investors lead current market returns, we test the hypothesis that each element in the vector $[\phi_{34}(L)\phi_{35}(L)\phi_{36}(L)]$ is zero. Before applying the VAR model, an appropriate lag structure needs to be specified. A 3-day lag is selected based on the Akaike information criterion (AIC). Table 17.5 presents the lead-lag relation among the six time series using block exogeneity tests.

The $\begin{bmatrix} \phi_{44}(L) & \phi_{45}(L) & \phi_{46}(L) \\ \phi_{54}(L) & \phi_{55}(L) & \phi_{56}(L) \\ \phi_{64}(L) & \phi_{65}(L) & \phi_{66}(L) \end{bmatrix}$ block represents the potential interaction among three types of institutional investors. The results indicate that net foreign purchases lead net *dic* purchases and the dynamic relationship between these two variables can be provided by the impulse response function (IRF) in Fig. 17.2a. Clearly, a one-unit standard error shock to net foreign purchases leads to an increase in net *dic* purchases, but this effect dissipates quickly by Period 2. Figure 17.2b, c indicate a feedback relation between net purchases by *qfii* and *rtf*. A one-unit standard error shock to net foreign purchases results in a positive response to net *rtf* purchases over the next two periods, which become negative in Period 3, followed by a positive response again after Period 4. Furthermore, a one-unit standard error shock to net *rtf* purchases also gives rise to an increase in net foreign purchases, which decays slowly over ten-period horizon.

Figure 17.2d shows that net purchases by *dic* lead net *rtf* purchases. A one-unit standard error shock to net *dic* purchases produces an increase in net *rtf* purchases in the first three periods and then declines thereafter. Overall, these impulse responses suggest that previous day's net foreign purchases exert a noticeable impact on net *rtf* purchases, while previous day's net *rtf* purchases also has an impact on net foreign purchases. It implies that not only do foreign capital flows affect the trading activity of domestic institutional investors but also the relation is not unidirectional. To be specific, there is a feedback relation between net *rtf* purchases and net foreign purchases.

As for the effect of the three types of institutional trading activity on stock returns in the TSE, Table 17.5 reveals that net *dic* purchases on previous day lead the TSE returns. We can also see in Fig. 17.2e that after Period 3, net *dic* purchases exert a negative (and thus destabilizing) effect on market returns, while the other two institutional investors do not have such an effect over the sample period. Examining the relationship between market returns and trading activity of the three types of institutional investors, we find that either the net foreign purchases or the net *dic* purchases on previous trading day are affected by the previous day's TSE returns. Moreover, the IRFs in Fig. 17.2f, g also reveal a significant positive relation between TSE returns and net foreign purchases up to four periods and a significant positive relation between TSE returns and net *dic* purchases for two periods, which becomes negative after Period 3. In other words, foreign investors in the TSE engage in positive-feedback trading, while those of the *dic* tend to change their strategy and adopt negative-feedback trading after Period 3.

Table 17.5 indicates that previous day's NASDAQ returns significantly affect both current TSE returns and net purchases by the three types of institutional investors. The impulse response in Fig. 17.2h–k also confirms that previous day's NASDAQ returns are positively related to both current returns and net purchases by these institutional investors, with the exception that a negative relation between net *dic* purchases and previous day's NASDAQ returns is found after Period 4.

Such results are much in sync with the expectation since the largest sector that comprises the TSE-weighted stock index is the electronics industry to which many listed companies on NASDAQ have a strong connection. In addition, although previous day's net *qfii* and *dic* purchases also lead the NASDAQ returns, we find that no significant relation exists except for Period 4 with a significantly negative relation between them (Figs. 17.1 and 17.2).

The liberalization of Taiwan's stock market has ushered in significant amount of short-term inflows and outflows of foreign capital, which have induced fluctuations in the exchange rate. As is seen from Table 17.5, the TSE returns lead currency returns and it appears that the initial significant effect of stock returns on currency returns is negative for the first three periods and then turns to be significantly positive thereafter (Fig. 17.2n). Given that foreign investors are positive-feedback traders, the capital inflows is expected to grow in order to increase their stakes in TSE securities when stock prices rise. Consequently, NT/USD is expected to appreciate.

The Structural VAR Model

The unrestricted VAR model does not consider the effect of current returns on net purchases by institutions. The prior study by Griffin et al. (2003) includes current returns in the institutional imbalance equation and finds a strong contemporaneous positive relation between institutional trading and stock returns. Therefore, to further examine the relationship between institutional trading and stock returns, we introduce the current TSE returns (r_t) in the net purchases equations of the three types of institutional investors and reestimate the VAR model before conducting the corresponding block exogeneity tests. Table 17.6 presents the estimation results.

As indicated in the last row of Table 17.6, we find evidence of a strong contemporaneous correlation between current returns and net institutional purchases, which confirms the finding by the previous research. As shown in Table 17.5, we find no evidence that past returns lead net *rtf* purchases when the unrestricted VAR model is used. When the contemporaneous impact of stock returns on net institutional purchases is considered, we find that past returns also lead net *rtf* purchases as well as net purchases by *qfii* and *dic* when the structural VAR model is used (Table 17.6). In other words, net purchases by the three types of institutions are affected by past stock returns as was evidenced by previous studies. The corresponding impulse response relations are presented in Fig. 17.3.

Comparing the impulse response relations in Figs. 17.2 and 17.3, it is clear that when the impact of current returns on net institutional purchases is considered, a one-unit standard error shock from r_t does not produce a positive impulse response in institutional trading until Period 2. The responses of foreign investors are rather distinct from those of domestic institutional investors after Period 3. In general, a sustained positive response from foreign investors is observed, while a negative response is witnessed for *dic* and sometimes, an insignificant response for *rtf* manifests itself after Period 3.

The Threshold VAR Analysis

We pool all the data together when estimating either the unrestricted or restricted VAR model; however, the trading activity of institutional investors may depend on whether stock prices rise or fall. A small economy like Taiwan also depends to a large degree on the sign of NASDAQ index returns. Consequently, to investigate institutional trading under distinct regimes based on market returns, we use the multivariate threshold autoregression (MVTAR) model proposed by Tsay (1998) to test the relevant hypotheses. Let $\mathbf{y}_t = [nasr_t, \Delta e_t, r_t, qfiibs_t, dicbs_t, rtfbs_t]'$ be a 6×1 vector and the MVTAR model can be described as

$$\begin{aligned}\mathbf{y}_t = &\left(\phi_{0,1} + \sum_{i=1}^{p} \phi_{i,1} y_{t-i}\right) \cdot [1 - I(z_{t-d} > c)] \\ &+ \left(\phi_{0,2} + \sum_{i=1}^{p} \phi_{i,2} y_{t-i}\right) \cdot I(z_{t-d} > c) + \varepsilon_t\end{aligned}$$

where $E(\varepsilon) = 0$, $E(\varepsilon\varepsilon') = \mathbf{\Sigma}$, and $I(\cdot)$ is an index function, which equals 1 if the relation in the bracket holds. It equals zero otherwise. z_{t-d} is the threshold variable with a delay (lag) d.

In order to explore whether institutional trading activity would change during different domestic and foreign market return scenarios, the potential threshold variables used are r_{t-1} and $nasr_{t-1}$. Before estimating Eq. 2, we need to test for possible potential nonlinearity (threshold effect) in this equation. Tsay (1998) suggests using the arranged regression concept to construct the $C(d)$ statistic to test the hypothesis H_o:$\phi_{i,1} = \phi_{i,2}$, $i = 0, \ldots p$. If H_0 can be rejected, it implies that there exists the nonlinearity in data with z_{t-d} as the threshold variable. Tsay (1998) proves that $C(d)$ is asymptotically a chi-square random variable with $k(pk + 1)$ degrees of freedom, where p is the lag length of the VAR model and k is the number of endogenous variables $\mathbf{y}_t$. Table 17.7 presents the estimation results of the $C(d)$ statistic.

As shown in Table 17.7, the null hypothesis H_0 is rejected using either past returns on the TSE or NASDAQ, suggesting that our data exhibit nonlinear threshold effect. Theoretically, one needs to rearrange the regression based on the size of the threshold variable z_{t-d} before applying a grid search method to find the optimal threshold value c^*. Nonetheless, our goal is to know whether the institutional trading behavior depends on the sign of market returns, as such the threshold is set to zero in a rather arbitrary way. Table 17.8 lists the results of block exogeneity tests for the lead-lag relation in the $r_{t-1} < 0$ and $r_{t-1} \geq 0$ regimes, respectively.

The interaction among institutional investors is depicted in Table 17.8: Current net purchases by foreign investors affect that by domestic institutions when previous day's TSE returns are negative. Note that no such relation is evidenced when previous day's TSE returns are positive. A feedback relation between *rtf* and *qfii* is observed when r_{t-1} is positive or negative. However, *dic* is found to lead *rtf* only when r_{t-1} is positive. Such results reveal different institutional trading strategies under distinct return regimes. The demonstration effect – previous day's net foreign purchases have on domestic institutions using the unrestricted VAR model – seems to surface only when previous day's market returns are negative. Therefore, it may produce misleading results if we fail to consider the sign of previous returns.

The impulse responses in Fig. 17.4 illustrate that the responses of *dic* and *rtf* from the *qfii* shock are quite similar to the ones in Fig. 17.2f, g. As for the impact of previous day's market returns on current net purchases by institutions, it can be shown via the MVTAR model that market returns lead net purchases by *qfii* and *dic* when previous day's market returns are negative, which is consistent with the finding using the one-regime VAR model. When previous day's market returns are positive, market returns lead net purchases by the *dic* only. Obviously, returns have more influence on net institutional purchases when previous day's returns were negative. In addition, we find that net *dic* purchases on the previous day may affect current returns when the one-regime VAR model is used. Actually, the MVTAR analysis reveals that such a relation emerges only when previous day's market returns are positive. The impulse responses depicted in Fig. 17.4 (Panel B) demonstrate that a one-unit standard error shock to net *dic* purchases produces an increase in market returns in Period 2, and then they turns to be negative after Period 4, a result similar to those using the one-regime VAR model.

To further investigate whether the interaction among institutions and the relationship between institutional trading and stock returns are affected by the sign of previous day's NASDAQ index return, $nasr_{t-1}$ is used as the threshold variable. That is, block Granger causality tests are performed by splitting our data into two regimes based on the sign of the variable $nasr_{t-1}$. Table 17.9 reports the results.

The results indicate that net *qfii* purchases lead that of two domestic institutional investors regardless of the sign of previous day's NASDAQ returns. Moreover, when $nasr_{t-1} < 0$, net *rtf* purchases lead net *qfii* purchases, which is in line with that using the one-regime VAR model. However, when $nasr_{t-1} \geq 0$, the net purchases by either *dic* or *rtf* lead net *qfii* purchases, and net *qfii* purchases lead net purchases by either *dic* or *rtf*. In other words, we find strong evidence of a feedback relation between net foreign purchases and two domestic net purchases when previous day's NASDAQ returns are positive. The results pertaining to the impact of previous day's returns on institutional trading parallel those using the unrestricted VAR model: Previous day's returns have an impact on the net purchases by *qfii* and *dic*, but not on net purchases by *rtf* regardless of the sign of previous day's NASDAQ returns. As for the impact of net institutional purchases on previous day's stock returns, only net *dic* purchases still lead stock returns when $nasr_{t-1} < 0$, as was the case in the one-regime model. However, we find that net *qfii* purchases also lead stock returns when $nasr_{t-1} \geq 0$.

References

Brennan, M., & Cao, H. (1997). International portfolio investment flows. *Journal of Finance, 52*, 1851–1880.

Cai, F., & Zheng, L. (2002). *Institutional trading and stock returns* (Working paper). Ann Arbor: University of Michigan Business School.

Choe, H., Kho, B. C., & Stulz, R. (1999). Do foreign investors destabilize stock markets? The Korean experience in 1997. *Journal of Financial Economics, 54*, 227–264.

Froot, K. A., O'Connell, P. G. J., & Seasholes, M. S. (2001). The portfolio flows of international investors. *Journal of Financial Economics, 59*, 151–193.

Griffin, J. M., Harris, J. H., & Topaloglu, S. (2003). The dynamics of institutional and individual trading. *Journal of Finance, 58*, 2285–2320.

Grinblatt, M., & Keloharju, M. (2000). The investment behavior and performance of various investor-types: A study of Finland's unique data set. *Journal of Financial Economics, 55*, 43–67.

Hamao, Y., & Mei, J. (2001). Living with the "enemy": An analysis of foreign investment in the Japanese equity market. *Journal of International Money and Finance, 20*, 715–735.

Huang, C.-H., & Hsu, Y. Y. (1999). The impact of foreign investors on the Taiwan stock exchange. *Taipei Bank Monthly, 29*(4), 58–77.

Kamesaka, A., Nofsinger, J. R., & Kawakita, H. (2003). Investment patterns and performance of investor groups in Japan. *Pacific-Basin Finance Journal, 11*, 1–22.

Karolyi, G. A. (2002). Did the Asian financial crisis scare foreign investors out of Japan? *Pacific-Basin Finance Journal, 10*, 411–442.

Lakonishok, J., Shleifer, A., & Vishny, R. (1992). The impact of institutional trading on stock price. *Journal of Financial Economics, 32*, 23–44.

Lee, T. S., & Oh, Y. L. (1995). Foreign investment, stock market volatility and macro variables. *Fundamental Financial Institutions, 31*, 45–47 (in Chinese).

Lee, Y.-T., Lin, J. C., & Liu, Y. J. (1999). Trading patterns of big versus small players in an emerging market: An empirical analysis. *Journal of Banking and Finance, 23*, 701–725.

Nofsinger, J. R., & Sias, W. (1999). Herding and feedback trading by institutional and individual investors. *Journal of Finance, 54*, 2263–2295.

Sadorsky, P. (1999). Oil price shocks and stock market activity. *Energy Economics, 21*, 449–469.

Sias, R. W., & Starks, L. T. (1997). Return autocorrelation and institutional investors. *Journal of Financial Economics, 46*, 103–131.

Tsay, R. S. (1998). Testing and modeling multivariate threshold models. *Journal of American Statistical Association, 93*(443), 1188–1202.

Wang, L.-R., & Shen, C. H. (1999). Do foreign investments affect foreign exchange and stock markets? – The case of Taiwan. *Applied Economics, 31*(11), 1303–1314.

Methods of Denoising Financial Data 18

Thomas Meinl and Edward W. Sun

Contents

Abstract

Denoising analysis imposes new challenges for financial data mining due to the irregularities and roughness observed in financial data, particularly, for instantaneously collected massive amounts of tick-by-tick data from financial markets for information analysis and knowledge extraction. Inefficient decomposition of the systematic pattern (the trend) and noises of financial data will lead to erroneous conclusions since irregularities and roughness of the financial data make the application of traditional methods difficult.

In this chapter, we provide a review to discuss some methods applied for denoising analysis of financial data.

T. Meinl (✉)
Karlsruhe Institute of Technology (KIT), Karlsruhe, Germany
e-mail: thomas.meinl@kit.edu

E.W. Sun
KEDGE Business School and BEM Management School, Bordeaux, France
e-mail: edward.sun@bem.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_18,

Keywords
Jump detection • Linear filters • Nonlinear filters • Time series analysis • Trend extraction • Wavelet

18.1 Introduction

Technical developments allow corporations and organizations to instantaneously collect massive amounts of data, particularly the tick-by-tick data from financial markets. Mining financial data turns to be the foundation for financial informatics and stimulates the research interest in analyzing information conveyed at different frequencies in this data. Financial data have the complex structure of irregularities and roughness that are caused by a large number of instantaneous changes of the markets and trading noises. The noises conveyed in the financial data usually illustrate heavy tailedness, that is, the underlying time series data exhibits a large number of occasional jumps. Ignoring these irregularities can easily lead to erroneous conclusions for data mining and statistical modeling. As consequence, the statistical data mining methods (or models) require a denoising algorithm to clean the data in order to obtain more significant results (see Sun and Meinl (2012)).

Most data cleaning methods focus only on a known type of irregularity. In financial data, the irregularity is manifold, that is, the irregularity varies along with time and changes with different measuring scales (see Fan and Yao (2003) and Sun et al. (2008)). Finding an effective denoising algorithm turns out to be the initial task of financial data mining (see Au et al. (2010) and Meinl and Sun (2012)).

In this chapter we will outline the classic and newly established methods for the denoising analysis (trend extraction) of financial data analysis. We talk about different approaches and a respective classification of them. Based on this classification we focus on nonparametric methods (i.e., linear and nonlinear filters) and examine their pros and cons.

18.2 Denoising (Trend Extraction) of Financial Data

There is no universal definition of *trend* which applies to all application in different fields, and it is generally accepted that a trend is a slowly evolving component that is the main driving force for long-term development beneath the system. Pollock (2001) characterizes trend as being limited to certain low frequencies of the data. This notion excludes any noisy influences and fluctuations from higher frequency levels. However, this notion of trends is not satisfactory for many financial data we encounter in practice.

Some theoretical models (like the Black-Scholes model) do not incorporate aspects like seasonalities or even jumps; they are still widely used today in practice, assuming perfect division between the trend and stochastic fluctuations. It is insufficient for many financial data we measure today, especially, when considering trends over longer time horizon, we perceive significant jumps or steep slopes which cannot be attributed to be part of the persistent stochastic noise. Another example we can observe in electricity markets, where it has been found that jumps occur on such a regular basis that is has been found reasonable to model these by specific stochastic processes; see Seifert and Uhrig-Homburg (2007). These patterns contradict the slow evolving characteristic and the low-frequency notion generally associated with trends and have been considered as an inherent part of them.

A basic model (among others proposed by Fan and Yao (2003)) is stated as follows. Given a time series X_t, $t \geq 0$, this series can be decomposed into

$$X_t = \vartheta_t + s_t + Y_t, \tag{18.1}$$

where ϑ_t represents a *slowly varying function known as trend component*, s_t a single of combination of periodic functions (i.e., seasonal components), and Y_t the stochastic component, assumed to be stationary.

A trend is a mostly slow (in respect to the noise) evolving pattern in a time series, with its driving force not being attributed to any noise present in the signal. Trends may also exhibit edged frontiers (i.e., jumps and sudden regime changes) as well as steep slopes, roofs, and valleys (see Joo and Qiu (2009)), as long as these patterns can be contributed to the long-term dynamics of the time series and do not stem from the noise component responsible for short-term variations.

We note that trend must always be interpreted in respect to the time series at hand (i.e., the period coverage of the financial data) and the goal of the data analysis, that is, on which scale the trend and the noise are relevant. We ignore any distinction between long-, intermediate-, and short-term trends and/or seasonalities, as these usually depend on the context of the financial data.

The question that may now arise is if we consider a trend ϑ to be specified according to Eq. 18.1, how does this agree with the notion that the trend may also exhibit jumps, which clearly contradict this definition? Without going further into this, we only note that this kind of trend would require a different (and more accurate) definition of how trend is to be defined in a time series, than the generally accepted notion in today's literature. While we will not propose a new definition either and leave this task to others, we can point out how jumps enter the trend.

As pointed out by, for example, Tsay (2005), jumps in financial time series data, and particularly in high-frequency data, are attributed to external events, like the increase or drop in interest rates by some governmental financial institution. These events can be considered to happen only occasionally and are very sparse in relation to the frequency the data is measured, that is, the amount of data exhibiting no jumps at all. In the field of high-frequency financial data analysis, jumps are thus assumed to be extreme events that happen with low probability but form nevertheless part of the stochastic distribution and must be considered to be modeled there. Thus, Y will be modeled either by stochastic processes including jump components (see, e.g., Seifert and Uhrig-Homburg (2007)) or by a distribution itself, depending on the model. Such a distribution for high-frequency financial data has then found to be heavy tailed, that is, jumps happen with enough regularity that they cannot simply be discarded as nonrecurring events; see, for example, Rachev (2003).

However, as these extreme events can have an enormous impact on the stochastic variance analysis and its succeeding usage, and furthermore could lead to misleading results in the regions of the signal without any jumps, it is generally preferred to include them into the trend component ϑ rather than the stochastic component Y.

In, for example, Wang (1995), a definition of an α-cusp in a continuous function f at x_0 is given, that is, if there exists an $\varepsilon > 0$ so that

$$|f(x_0 + h) - f(x_0)| \geq K|h|^{\alpha} \tag{18.2}$$

holds for all $h \in [x_0 - \varepsilon, x_0 + \varepsilon]$. For $\alpha = 0$, f is said to have a jump at x_0. It is commonly agreed that the jump should significantly differ from the other fluctuations (i.e., the noise) in the signal. As said above, jumps are just one particular pattern of extreme events we are interested in. Others are steep slopes, roofs, and valleys, which in Joo and Qiu (2009) are defined by at this point having a jump in the first-order derivative of the regression curve.

Other extreme events frequently occurring in many practical applications are spikes and outliers. However, these are usually undesirable features that should not be included in the trend of affect it by any means. This is due to the following reasons. First, in many cases, these outliers or spikes consist only of one or very few points often caused by measurement errors, and it is obvious that they were caused by some factor that plays no vital role in the ongoing time series analysis (unless the focus is on what caused these outliers). Second, while jumps imply a permanent change in the whole time series, outliers do not contribute to this. While we are aware that the distinction between a few (adjacent) outliers and roofs/valleys is not precise, from the context of the time series in most cases, it is evident whether an occurrence should be considered as an outlier that is to be neglected or a significant feature to be included in the trend.

In this work we only consider homogeneous financial time series data. However, in many applications and particularly for high-frequency financial time series data, this data is initially irregularly spaced. Homogeneity in this context means that for a given series X_t, $t \in \mathrm{N}$, holds $t + 1 - t = c$, with a constant $c > 0$, that is, all time steps are equally spaced. As we will see, this is not always the case for empirical time series, especially in the area of financial high-frequency data. In this case, it is necessary to preprocess the inhomogeneous (i.e., irregularly spaced) time series by interpolation methods in order to regularize the raw data. Though there exist models which can handle inhomogeneous time series directly (see Dacorogna (2001), but they also remark that most today's models are suited for regularly spaced time series only), regarding the methods we discuss in the following sections, we restrict ourselves to homogeneous ones.

In this chapter we focus on nonparametric methods for trend extraction. This is due to the reason that in most time series we analyze in this work, we cannot reasonably assume any model for the underlying trend. Yet, as noted in the framework above, in case such assumptions hold, we can expect those models to perform better than nonparametric models, since they exploit information that nonparametric approaches cannot. Furthermore, the commitment to certain parametric time series or trend models can be seen as a restriction when considering the general case, which may lead even to misleading results in case the trend does not match model, as certain patterns might not be captured or considered by the model itself. This can easily be seen at a most basic example, in case a linear trend is expected, which in most cases will only be a poor estimator for any nonlinear trend curve. In this case nonparametric approaches are less restrictive and can more generally be applied, while of course not delivering the same accuracy as parametric models which exactly match the underlying trend, with only their parameters to be calibrated. If, e.g., the trend follows sinusoidal curve, a sinusoidal curve with its parameters being estimated by, e.g., the least-squares method will almost surely provide a better accuracy than any other nonparametric approach. On the other hand, if the underlying trend is linear or even contains only little deviations from a perfect sinusoidal curve, a sinusoidal fit to this trend, it has been shown at simple examples that the parametric sinusoidal approach can lead to confusing results and conclusions.

We further require that the method used for trend extraction be robust, that is, the results are reliable and the error can be estimated or is at least bounded in some way. In many cases (see, e.g., Donoho and Johnstone (1994)), the robustness of a method is shown by proving its asymptotic consistency, that is, its convergence towards a certain value for certain parameters tending towards infinity. It should be remarked that the robustness should be independent of the time series itself and/or any specific algorithm parameter sets, in order to be applicable in practice. Of course this does not exclude specific assumptions on the time series that must be met or parameter ranges for which the algorithm is defined.

Therefore, as we cannot reasonably assume any model for the any time series in general, in this work, we focus on nonparametric approaches only. Within these approaches we focus on two main branches for trend extraction: linear and nonlinear filters. This is for the reason because linear filters are known and have proven to deliver a very smooth trend (given the filtering window size is large enough), while nonlinear filters excel at preserving characteristic patterns in a time series, i.e., especially jumps. Both methods in general require only very few (or none at all) information about the underlying data, besides their configuration of weights and calibration parameters, and are thus applicable to a wide range of time series, independent of the field the data was measured in.

Although there exists a variety of other nonparametric methods, most of these already rely on specific assumptions or choices of parameters which in general cannot easily be derived for any time series data or different analysis goals. Nevertheless, for the sake of completeness, later in this chapter we list some alternative methods, also including parametric approaches, which have been applied in economic, financial, or related time series data.

18.3 Linear Filters

Linear filters are probably the most common and well-known filters used for trend extraction and additive noise removal. We first provide the most general notion of this filter class and then depict the two viewpoints of linear filters and how they can be characterized. While this characterization on the one hand is one of the most distinguishable advantages of linear filters, on the other hand at the same time, it leads to the exact problem we are facing in this work, that is, the representation of sharp edges in otherwise smooth trends.

18.3.1 General Formulation

The filtered output depends linearly from the time series input. Using the notation of Fan and Yao (2003), a linear filter of length $2h + 1$ can be defined as

$$\hat{X}_t = \sum_{i=-h}^{h} w_i X_{t+i}, \tag{18.3}$$

with $\hat{X}_t$ the filtered output and w_i the filter weights. These kinds of filters are also known as (discrete) convolution filters, as the outcome is the convolution of the input signal with a discrete weight function, where often the following notation is used:

$$w^* X_t := \sum_{i=-h}^{h} w_i X_{t-i} = \sum_{i=-h}^{h} w_{t-i} X_i. \tag{18.4}$$

Thus, for every data point X_t, the filtered output $\hat{X}_t$ is the result of weighted summation of data points around t. Applied for the whole time series, this results in weighted average window of size $L = 2h + 1$ which is moved throughout the whole series. The size of this window is also called the *bandwidth* of the filter. It is also common notation to set $h \equiv \infty$ even if the window filter size should be finite, with $w_i = 0$ for $|i| > h$.

The probably best known linear filter is the mean filter, with $w_i = 2h + 1$, that is, all filter weights are uniformly distributed. A more general viewpoint is given by the notion of *kernel filters*. Given a kernel function w.l.o.g. with support $[-1, 1]$, this function assigns the weights according to

$$w_i = \frac{K(i/h)}{\sum_{j=-h}^{h} K(j/h)}. \tag{18.5}$$

Commonly used examples are the *Epanechnikov kernel*

$$K^{\mathrm{E}}(u) = \frac{3}{4}\left(1 - u^2\right) + , \tag{18.6}$$

the *Gaussian kernel*

$$K^{\mathrm{G}}(u) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-u^2}{2}\right), \tag{18.7}$$

and the symmetric beta family

$$K_\gamma^{\mathrm{B}}(u) = \frac{1}{B(1/2,\ \gamma+1)} \left(1-u^2\right)^\gamma I_{|u|\leq 1}. \tag{18.8}$$

For the values $\gamma \in \{0, 1, 2, 3\}$, the kernel $K_\gamma^{\mathrm{B}}(u)$ corresponds to the uniform, Epanechnikov, biweight, and triweight kernel functions, respectively.

18.3.2 Transfer Functions: Time Versus Frequency Domain

The previous section depicts the linear filtering method in the time domain, i.e., we look at the time series X_t and its respective filtered output $\hat{X}_t$ and how they evolve over time t. Another perception can be given by taking the frequency domain into account. For all linear filters, we cannot only give its definition as depicted earlier, but also in respect to the frequencies, the filters let pass. This notion can be derived as follows.

While the sequence of filter weights w_i, also called *impulse response sequence*, determines the filtered output in the time domain (or equivalent are the linear filter's time domain representation), via the discrete Fourier transform (DFT), we can derive the *transfer function*

$$\mathcal{W}(f) = \sum_{j=-\infty}^{\infty} w_j e^{-\mathrm{i}2\pi f j}, \tag{18.9}$$

its counterpart in the frequency domain, also called *frequency response function*. Alternatively, if this formulation is given in the beginning, we can also derive the weights via the inverse transform:

$$w_j = \int_{-1/2}^{1/2} \mathcal{W}(f) e^{\mathrm{i}2\pi f j} \mathrm{d}f. \tag{18.10}$$

Obviously, these two formulations are equivalent, as one can be derived from the other, and vice versa. By considering the transfer function's polar representation,

$$\mathcal{W}(f) = |\mathcal{W}(f)| e^{\mathrm{i}\theta(f)}, \tag{18.11}$$

with $|\mathcal{W}(f)|$ the gain function. The magnitude in gain $|\mathcal{W}(f)|$ describes the linear filters behavior in the frequency domain, that is, what kind of frequencies and their respective proportions will let passed or be blocked. For our needs, it satisfies to distinguish between high – and low-pass filters, that is, filters that let pass either the high frequencies and block the lower ones, or vice versa. Of course, other filter types exist; for example, by combining high – and low-pass filters (e.g., in a filter cascade), one can derive band-pass and stop filters, so that the frequency domain output will be located only in a certain frequency range. In this work we are specifically interested in low-pass filters, as they block the high-frequency noise and the output consists of the generally low-frequency trend.

As we assume that the weights w_j are real valued, one can show (see, e.g., Percival and Walden (2006)) that $\mathcal{W}(-f) = \mathcal{W}^*(f)$, and, with $|\mathcal{W}^*(f)| = |\mathcal{W}(f)|$, it follows that $|\mathcal{W}(-f)| = |\mathcal{W}(f)|$. Therefore, the transfer functions are symmetric around zero. Because of its periodicity, we need to consider $\mathcal{W}(f)$ only an interval of unit length. For convenience, this interval is often taken to be $[-1/2, 1/2]$, i.e., $|f| \leq 1/2$. Therefore, with above-depicted symmetry, it suffices to consider $f \in [0, 1/2]$ in order to fully specify the transfer function.

While saying that certain frequencies are blocked and others are passed, this holds only approximately true, since the design of such exact frequency filters is not possible, but always a transition between the blocked and passing frequencies. The goal of many linear filters is either to minimize these transitions (i.e., the range of by this affected frequencies), which, on the other hand, inevitably causes ripples in the other frequencies, that is, they are not any longer blocked, or to let them pass completely (see Bianchi and Sorrentino (2007) and the references therein for further details about this topic).

As we have seen at above examples, linear filters can be designed either from a time or from a frequency perspective. The time domain usually focuses on putting more weights on the surrounding events (i.e., events that recently before or occurred shortly after) and, thus, gives an (economic) interpretation similar to the, for example, ARMA and GARCH models. On the contrary, the frequency domain is based on the point of view that certain disturbances (almost) exclusively are located in a certain frequency range and are thus isolated from the rest of the signal. Also, the slowly evolving trend can be seen to occupy only the lower frequency ranges. Thus, the (economic) meaning lies here in the frequency of events; see, for example, Dacorogna (2001).

Although linear filters can be designed to block or let pass certain frequencies nearly optimal, at the same time this poses a serious problem when facing trends that exhibit jumps or slopes. As these events are also located in the same (or in case of slopes the adjacent) frequency range as the high-frequency noise, this has the effect that jumps and edged frontiers are blurred out, while steep slopes mostly are captured with poor precision only. Hence, from a frequency perspective, a smooth trend and edge preservation are two conflicting goals. This is as the linear filters are not capable to distinguish between the persistent noise and a single events that, while located in the same frequency range, usually have a higher energy and, thus, significantly larger amplitude. Thus, the same filtering rule is applied throughout the whole signal, without any adaption. Note, however, that linear filters still give some weight to undesirable events like outliers, due to their moving average nature. Thus, while some significant features like jumps are not displayed in enough detail, other unwanted patterns, like spikes, still partially carry over to the filtered output. To overcome all these drawbacks for that kind of trends or signals, besides the class of linear filters, the class of nonlinear filters has been developed.

18.4 Nonlinear Filters

As we have seen, linear filters tend to blur out edges and other details though they form an elementary part of the time series' trend. In order to avoid this, a wide range of nonlinear filters has been developed which on one hand preserve those details while on the other try to smooth out as much of the noise as possible. We do not find nonlinear filters only in time series, but they were in many cases developed specifically for two-dimensional signals, specifically images, where the original image, probably corrupted by noise during data transmission, consists mainly of edges, which form the image.

18.4.1 General Perception

While linear filters generally provide a very smooth trend achieved through averaging, two characteristics pose a problem for this class of filters:

- Outliers and spikes
  Single, extreme outliers and spikes can cause the whole long-term trend to deviate in the same direction, though they obviously do not play a part in it.
- Jumps, slopes, and regime changes
  Whenever there occurs a sudden external event in the underlying main driving force, it causes the trend to jump, that is, contrary to spikes it changes permanently onto another plane. While slopes are not that extreme, they also show a similar behavior as they decay or rise with the trend's unusual degree.

The reasons for the deviation sensitivity to these events are given by one of the most favorable linear filters' characteristics themselves: It follows directly from them being characterizable in terms of frequency passbands we explained earlier that all frequencies are treated the same (i.e., filtered according to the same rule) throughout the whole signal. This means that no distinction is made (and even cannot be made) between noise and those patterns, as they are located in approximately the same frequency range. Technically, as long as an outlier or a jump is contained in the weighted moving average filtering window, also a weight is assigned to these outlier data points or the points on the other plane, i.e., before and after the jump. Nonlinear filtering procedures try to avoid this by using a different approach, for example, by considering a single value only (instead of multiple weighted ones) that was selected from an ordered or ranked permutation of the original values in the filtering window.

Although we cannot characterize nonlinear filters in the same way as we do with linear filters (i.e., by transfer functions) besides their characteristics not being classifiable in that way, according to Peltonen et al. (2001), we can divide the whole class of these filters into several subclasses that share the same or similar approaches. Among them, there are stack filters, weighted median filters, polynomial filters, and order statistic filters. Astola and Kuosmanen (1997) provide two different taxonomies for further classification, though they remark that these divisions are not unique. In their work, they extensively show how those different filters behave (i.e., their characteristics) when applied onto different benchmark signals in respect to the mean absolute error (MAE) and mean squared error (MSE) measures.

The behavior of nonlinear filters is generally characterized by their impulse and step response, i.e., the filtered output when the input consists of a single impulse or step only. These impulses generally are given by the sequence $[\ldots, 0, 0, 0, a, 0, 0, 0, \ldots]$ and $[\ldots, 0, 0, 0, a, a, a, \ldots]$, respectively, with $a \neq 0$. Though in most cases no analytical result can be given, these characteristics help us to understand how the nonlinear filter behaves in respect to those patterns, for which the linear filters generally fail to deliver adequate results.

Despite their undoubtedly good ability to preserve outstanding features in a time series while extracting the trend, nonlinear filters also suffer some drawbacks:

1. Insufficient smoothness

   Though most nonlinear filters try to deliver a smooth trend and a good resolution of edged frontiers, beyond the jumps' surrounding regions, they usually fail to deliver a smooth trend as accurate as even simple linear filter (e.g., the mean filter) provides. Yet, by applying further smoothing procedures (e.g., by recursive filters or some kind of linear filtering on the nonlinear output, with a smaller bandwidth, as most of the noise is already smoothed out) comes at the price that the prior preserved details of jumps or slopes tend to get lost.

   This effect is even aggravated when the high-frequency noise present in the signal is extremely high, i.e., the trend which evolves, besides the jumps, quite slowly is dominated by the noise with extremely high amplitudes. Some filters try to counter this effect, but they either:

   - Provide only a tradeoff between an overall smooth signal and poor jump resolution, or a trend still exhibiting ripples but preserved edges, or
   - Rely on further information about the time series itself, that is, the noise component and its structure, the jumps or the trend itself

   This is due to the problem that most filtering rules are applied throughout the whole signal, that is, they do not adapt themselves sufficiently when the filtering window approaches a jump. An overview of these and other nonlinear filters' performance is given by Astola and Kuosmanen (1997), proving again the well-known insight that there cannot exist one solution that performs optimal for all cases. Though the authors also report some approaches that try to incorporate that behavior, these filters provide only dissatisfactory results (see the examples below).

2. Lack of frequency control
Another feature nonlinear filters are lacking is the ability to regulate the filtered output in terms of frequency passbands, as linear filters do. Since Pollock (2001) defines the trend in terms of frequency bands and Ramsey (1999, 2002) points out that frequency analysis is an important aspect in financial time series, so is frequency control. Though not necessary for all applications, the ability to a priori control and regulate the filters output (in contrast to an only a posteriori frequency analysis of the filtered result) may come handy when one wants to ensure that certain frequencies are not contained in the output, that is, when certain information about the noise frequencies is at hand, the analyst can decide before the actual filtering process (and thus, without any try and error procedures) what frequency parts should be filtered out. Although a nonlinear filter can also provide the same or a similar result, no theoretical results or statements are available before the filtering procedure has been carried out completely. This incapacity of the nonlinear filter follows directly from the fact that nonlinear filters do not rely on frequency passbands, as they must be able to filter events, even though they occupy nearly the same frequency range (i.e., noise and jumps), due to different rules.

As Astola and Kuosmanen (1997) classify linear filters to be a subclass of the class of nonlinear filters, above statement does not exactly hold true, i.e., it would mean that a subclass of nonlinear filters can be characterized by their transfer function and frequency output. We, however, separate the classes of linear and nonlinear filters by whether or not a filter can be described by a transfer function or not, which marks a strict division of these two classes.

18.4.1.1 Examples

To illustrate the different courses of action of nonlinear filters and give the reader an idea of their general procedure, in this section we outline several examples of above named subclasses. As this list cannot be exhaustive by any means, of course, we note that we do not take filters into account that already rely on specific assumptions of the systems beneath the time series themselves.

Trimmed Mean Filter

This filter works essentially as the mean filter, with the difference that the extreme values of the *ordered* series $X_{(i)}$ are trimmed. Index t is omitted here as the order is no longer in concordance with time. Therefore, an (r, s)-fold trimmed mean filter is given by

$$\frac{1}{N-r-s}\sum_{i=r+1}^{N-s} X_{(i)}. \tag{18.12}$$

A special case is the choice of $r = s$. A further modification of the trimmed mean filter is not to discard the ordered values beyond $X_{(r)}$ and $X_{(s)}$, but instead replace them by $X_{(r+1)}$ and $X_{(s+1)}$ themselves. This is the Winsorized mean filter:

$$\frac{1}{N}\left(r \cdot X_{(r+1)} \sum_{i=r+1}^{N-s} X_{(i)} + s \cdot X_{(N-s)}\right). \tag{18.13}$$

In these methods, the (r, s) tuple is dependent of the data itself. Other filters consider to make these values independent from the data or dependent from the central sample itself, i.e., nearest neighbor techniques. All those filters have in common that they discard all samples from the ordered series being too far away (respectively) according to some measure.

L-Filters and Weighted Median

L-filters (also called order statistics filters) make a compromise between the weighted moving averages of linear filters and the nonlinear ordering operation. The idea is that the filtered output is generated by weighted averages over the ordered samples, that is,

$$\sum_{i=1}^{N} w_i X_{(i)}, \tag{18.14}$$

with w_i the filter weights. A similar notion is given by weighted median filters, where the weights are assigned to the time ordered sample X_t and where the weights denote a duplication operation, i.e., $w_i \circ X_t = X_{t,1}, \ldots, X_{t,w_i}$. The output is then given by

$$\text{median}\,\{w_1 \circ X_1, \ldots, w_N \circ X_N\}. \tag{18.15}$$

Ranked and Weighted Order Statistic Filters

An rth-ranked-order statistic filter is simply given by taking $X_{(r)}$ as the filter output. Examples are the median, the maximum ($r = N$), and the minimum ($r = 1$) operation. This can also be combined with weights as depicted above, that is,

$$r\text{th order statistic}\{w_1 \circ X_1, \ldots, w_N \circ X_N\}. \tag{18.16}$$

Hybrid Filters

Another approach is the design of nonlinear filters consisting of filter cascades, that is, the repeated application of different filters on the respective outputs. A general formulation is, for example, given by rth-order statistic $F_1(X_1, \ldots, X_n), \ldots, F_M(X_1, \ldots, X_n)$, where $F_1, \ldots, F_M$ can denote any other filtering procedure. A concrete example is the median hybrid filter that combines prior linear filtering procedure with a succeeding median ordering operation, i.e.,

$$\text{median}\left\{(1/k)\sum_{i=1}^{k} X_i,\ \ X_{k+1},\ \ (1/k)\sum_{i=k+2}^{N} X_i\right\}. \tag{18.17}$$

Selective Filters

An interesting approach is given by the principle of switching between different output rules depending on some selection rule, for example, based on the fact that the mean filter delivers a larger (smaller) output than the median filter when the filtering window approaches an upward (downward) jump. Thus, a selection rule would be given by

$$\text{mean}\{X_1, \ldots, X_N\} \geq \text{median}\{X_1, \ldots, X_N\}. \tag{18.18}$$

A certain drawback of this selection rule is that it is one-sided, that is, it considers only the first half of the region around the jump. This is due to the fact that the mean for the second half, after the jump has occurred, is generally smaller (larger) than the median. Other rules can include thresholds and aim at deciding whether there has actually happened a jump or there was an impulse in the signal, which is not caused by the noise distribution, but due to some other explanatory effect.

Above-depicted examples of nonlinear filters should give the reader an overview over the most common methods applied in practice. For detailed information about each filter's characteristics, advantages, and drawbacks, we refer to Astola and Kuosmanen (1997), where there are also the references to the original works to be found. Yet, we see that basically most of these filters rely on some ordered statistics, with their input or output modified prior or afterwards, respectively. Since this basic principle applies to most filters not directly dependent on some specific characteristic or assuming a certain structure of the original series to be filtered, the different methods pointed out above can be combined in numerous ways. In many cases, however, even though we portrait only the very basic methods, we see that almost all of them already incorporate an implicit or explicit choice of additional parameters besides the filter bandwidth, either by weights, rules, or thresholds. These choices introduce further biases into the filtering process. Though some of these parameters can be chosen to be optimal in some sense (i.e., minimize a certain distance measure, e.g., the MSE for the L-filters), they lack the concrete meaning of weights we have for linear filters.

18.5 Specific Jump Detection Models and Related Methods

In this section we list further related methods that are concerned with the estimation of long-term trends exhibiting edged frontiers, i.e., jumps and/or steep slopes. We first review methods being explicitly developed either only for the detection of jumps in a signal corrupted by noise or approaches that also include capturing (i.e., modeling) those very jumps. We show the advantages and limits of applications of these methods, highlighting in which aspects further research is still necessary. We conclude this chapter by listing some of the most in practice well-known methods.

18.5.1 Algorithms for Jump Detection and Modeling

The issue of detecting and modeling jumps in time series has been recognized as an essential task in time series analysis and therefore has already been considered extensively in the literature. Though we introduce wavelet methods in the next chapter only, we list the works based on these methods here as well without going into details. We only note that wavelets, based on their characteristics, make excellent tools for jump and spike detection, as it is this what they were developed for in the first place (see Morlet (1983)).

Generally, the most appreciated procedures in the recent literature can be seen as two different general approaches. One is via wavelets, and the other uses local (linear) estimators and derivatives.

One of the first approaches using wavelets for jump detection in time series, besides the classical wavelet literature, for example, Mallat and Hwang (1992), was given by Wang (1995, 1998). He uses wavelets together with certain data-dependent thresholds in order to determine where in the signal jumps have happened and whether they are significantly different from short-varying fluctuations, and provides several benchmark signals. Assumptions about the noise structure were made according to Donoho and Johnstone (1994), that is, the approach is applicable for white (i.e., uncorrelated) Gaussian noise only. This work was extended by Raimondo (1998) to include even more general cusp definitions. More recent contributions extend these works to stationary noise (Yuan and Zhongjie (2000)) and other distributions (Raimondo and Tajvidi (2004)) and also provide theoretical results about asymptotic consistency. A further application specifically on high-frequency data is given by Fan and Wang (2007).

The other line is given by Qiu and Yandell (1998) who estimate jumps using local polynomial estimators. This work is continued in Qiu (2003) where jumps are not only detected but also represented in the extracted time series. Gijbels et al. (2007) further refine the results by establishing a compromise between a smooth estimation for the continuous parts of a curve and a good resolution of jumps. Again, this work is limited to pure jumps only and, since it uses local linear estimators as the main method, has no frequency interpretation available. Sun and Qiu (2007) and Joo and Qiu (2009) use derivatives to test the signal for jumps and to represent them.

Finally, we note a completely different approach that is presented in Kim and Marron (2006): a purely graphical tool for the recognition of probable jumps (and also areas where almost certainly no jumps occurred). Yet, the authors confess that their work is only to be seen as a complementary approach, and refer to Carlstein et al. (1994) for further thorough investigation.

We note that there exists already a good body of work about how to detect (i.e., estimate their location), and even how to model jumps (i.e., estimate their height), though in most works bounded to strict requirements on the noise or the trend model. However, although most models will also automatically include the detection of steep slopes, they fail at modeling the slope itself. While a jump can easily be presented (either by indicator functions or any other methods used in the cited works above), matters are different with slopes: Since there can be given no general formulation or model of the exact shape of a slope, any parametric approach will fail or deliver only a poor approximation if the model does not fit the occurred slope. Examples of such different kinds of slopes are innumerous: Sine, exponential, and logarithmic decays are only the most basic forms to approximate such events, which in practical examples rarely follow such idealized curves. Naturally, only nonparametric approaches will adapt themselves to the true shape of the underlying trend but generally suffer the same drawbacks as all linear and nonlinear methods pointed out above, i.e., they always have to bargain a tradeoff between bias and signal fidelity.

18.5.2 Further Related Methods

We conclude this section by outlining the most popular and established filtering methods.

18.5.2.1 General Least-Squares Approaches

The most general approach can be seen by setting up a parameterized model and calibrating the model afterwards, generally using some minimization procedure in respect to some error measure. This can be seen as a straightforward approach, requiring only the initialization of an appropriate model and choice of error measure.

An example from the energy market is given by Seifert and Uhrig-Homburg (2007) and Lucia and Schwartz (2002), who set up a trigonometric model in conjunction with indicator functions and minimize the squared error for each time step in order to estimate the deterministic trend as well as values for different seasons and days. Though they find that their model works well in practice, for general cases it might be difficult always finding an appropriate model, especially when there is no information about the trend, its seasonal cycles, and other (deterministic or stochastic) influences. This is especially the case when the data set covers only a short period of time.

When analyzing financial time series data, one could suggest the trigonometric model and the order of the sinusoidal functions; in general it may be difficult to set up such a model or reason, why this model and its estimated trend are appropriate for the respective time series. This requires either a rigorous a priori analysis of the series itself or further information about the external factors (i.e., the system the time series is derived from) and their interaction. In addition to this, the estimation can never be better than the model and to which accuracy it approximates the true trend.

We note another probable critical issue when using indicator functions in combination with least-squares estimation. First, using indicator functions confines the model to jumps only, that is, slopes or similar phenomena cannot be captured by that approach, as the indicator functions automatically introduce jumps in the trend component. Thus, indicator functions excel at modeling jumps but perform poorly with other types of sudden changes. Second, for this approach, it is extremely important to determine the location of the jump as exact as possible, as otherwise the estimated trend in this area may be highly inaccurate.

It is therefore dubious, whether such parametric approaches are appropriate to deal with economic and financial time series, though they will perform very well, if their requirements are met.

18.5.2.2 Smoothing Splines

Smoothing splines, though also utilizing the least-squares methods, on the contrary do not rely on a specific model assumption. Instead, they penalize the regression spline in respect to its roughness, i.e.,

$$\min_{m} \sum_{t=1}^{N} (X_t - m(t))^2 + \omega \int (m''(x))^2 \, dx. \tag{18.19}$$

It follows directly from this definition that for $\omega = 0$ this yields us an interpolation while for $\omega \to \infty$ m will approximate a linear regression.

In practice, there emerges another difficulty: In many cases, the (optimal) choice of ω remains unclear. Although there exist several works that have established some data-dependent rules for this, in many cases, when the assumptions about the noise do not hold or the time series incorporates additional deterministic (e.g., cycles) or stochastic components (e.g., outliers that are part of the system and not due to measurement or other errors), the choice of the penalizing smoothing parameter is a difficult task that has been and is still undergoing extensive research (see Morton et al. (2009), Lee (2003), Hurvich et al. (1998), Cantoni and Ronchetti (2001), Irizarry (2004)). We only mention particularly the cross-validation method which is used to determine the optimal smoothing parameters (see Green and Silverman (1994)). Furthermore, though ω is eventually responsible for the degree of smoothness (i.e., on which scale or level the trend shall be estimated), one can hardly neither impose nor derive any additional meaning on or from this parameter.

18.5.2.3 The Hodrick-Prescott Filter

The Hodrick-Prescott (HP) filter was first introduced by Leser (1961) and later became popular due to the advanced works of Hodrick and Prescott (1997). In order to extract the trend $\tau = [\tau_0, \ldots, \tau_{N+1}]$ from a given time series X, this trend is derived by solving

$$\min_{\tau} \sum_{t=1}^{N} (X_t - \tau_t)^2 + \omega((\tau_{t+1} - \tau_t) - (\tau_t - \tau_{t-1}))^2. \tag{18.20}$$

We note that (besides from not using a spline basis for approximation) this approach can be seen as a discretized formulation of the smoothing spline.

The smoothing parameter ω plays the same role, while the penalized smoothness measure is the discretized version of the second derivative. Thus, though several authors propose explicit rules (of thumb) for choosing ω (see, e.g., Ravn and Uhlig (2002), Dermoune et al. (2008), Schlicht (2005)), some researchers like Harvey and Trimbur (2008) also recognize that in some way this choice still remains kind of arbitrary or problematic for many time series which cannot be associated in the same time terms.

18.5.2.4 The Kalman Filter

Another sophisticated filter was developed by Kalman (1960). It is a state-space system specifically designed to handle unobserved components models and can be used to either filter past events from noise or forecast. Another good introduction to the Kalman filter can be found in Welch and Bishop (1995) and a thorough discussion in Harvey (1989).

The basic Kalman filter assumes that the state $x \in \mathbb{R}^n$ of underlying process in a time series can be described by a linear stochastic difference equation:

$$x_t = Ax_{t-1} + Bu_{t-1} + w_{t-1}, \tag{18.21}$$

with A the state transition model that relates in conjunction with the (optional) control input B, the respective input u_t, and the noise component wt the previous state to the next. In the above equation, A and B are assumed to be constant but may also change over time. Of this model (i.e., the true state x_t), only

$$z_k = Hx_k + v_k \tag{18.22}$$

can be observed. Both noise components are assumed to be independent of one another and to be distributed according to

$$w \sim N(0, Q) \quad \text{and} \quad v \sim N(0, R). \tag{18.23}$$

With A, B, Q, and R assumed to be known, the filter predicts the next state x_t based on x_{t-1} and also provides an estimate of the accuracy of the actual prediction.

Since its first development, the Kalman filter has become popular in many areas; see, for example, Grewal et al. (2001). However, a serious drawback of this Kalman procedure is that many real-world models do not fit the assumptions of the model, for example, above requirement of a linear underlying system is often not met. Though there exist extensions for nonlinear systems (see e.g., Julier and Uhlmann (1997)), there still exists the problem that one or more of the required parameters are unknown. While for many technical systems (e.g., car or missile tracking systems) based on physical laws the state transition model A is exactly known, this becomes a difficult issue in many other application areas, including finance. Additionally, as, e.g., Mohr (2005) notes, the performance of the Kalman filter can be very sensitive to initial conditions of the unobserved components and their variances, while at the same time it requires an elaborate procedure of model selection. He also notes that in macroeconomic time series, the Kalman filter does not work with annual data. Therefore, we see that while the Kalman filter unquestionably delivers excellent results in many areas (and has also applied for financial time series as well, though not without critique), its usage is not convenient for general cases we treat in this work, requiring more assumptions and knowledge about the underlying model.

18.6 Summary

In this chapter we provided an overview of the different tasks of time series analysis in general and the specific challenges of time series trend extraction. We pointed out several methods and outlined their advantages and disadvantages, together with their requirements. The reader should keep in mind the following key points:

- Linear filters provide a very smooth trend but fail to capture sudden changes.
- Nonlinear filters capture jumps extremely well but are very sensitive to higher levels of noise.
- Advanced methods may provide good results depending on the scenario but often rely on very specific assumptions that can yield misleading results, in case they are not completely met.
- The specific task of jump detection and modeling is well understood but lacks appropriate methods for more general changes like steep slopes.
- And the most important: No methods universally perform best for all tasks!

References

Astola, J., & Kuosmanen, P. (1997). *Fundamentals of nonlinear filtering*. Boca Raton: CRC Press.

Au, S., Duan, R., Hesar, S., & Jiang, W. (2010). A framework of irregularity enlightenment for data pre-processing in data mining. *Annals of Operations Research, 174*(1), 47–66.

Bianchi, G., & Sorrentino, R. (2007). *Electronic filter simulation & design*. New York: McGraw-Hill. ISBN 0071494677.

Cantoni, E., & Ronchetti, E. (2001). Resistant selection of the smoothing parameter for smoothing splines. *Statistics and Computing, 11*(2), 141–146.

Carlstein, E., Mueller, H. G., & Siegmund, D. (1994). *Change-point problems*. Hayward, CA: Institute of Mathematical Statistics.

Dacorogna, M. M. (2001). *An introduction to high-frequency finance*. San Diego, CA: Academic Press.

Dermoune, A., Djehiche, B., & Rahmania, N. (2008). Consistent estimators of the smoothing parameter in the hodrick-prescott filter. *Journal of the Japan Statistical Society, 38*, 225–241.

Donoho, D. L., & Johnstone, I. M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika, 81*(3), 425–455.

Fan, J., & Wang, Y. (2007). Multi-scale jump and volatility analysis for high-frequency financial data. *Journal of the American Statistical Association, 102*(480), 1349–1362.

Fan, J., & Yao, Q. (2003). *Nonlinear time series: Nonparametric and parametric methods*. New York: Springer.

Gijbels, I., Lambert, A., & Qiu, P. (2007). Jump-preserving regression and smoothing using local linear fitting: A compromise. *Annals of the Institute of Statistical Mathematics, 59*(2), 235–272.

Green, P. J., & Silverman, B. W. (1994). *Nonparametric regression and generalized linear models: A roughness penalty approach*. New York: Chapman & Hall/CRC. ISBN 0412300400.

Grewal, M. S., Andrews, A. P., & Ebooks Corporation. (2001). *Kalman filtering theory and practice using MATLAB*. New York: Wiley Online Library. ISBN 0471392545.

Harvey, A. C. (1989). *Structural time series models and the Kalman filter forecasting*. Cambridge: Cambridge University Press.

Harvey, A., & Trimbur, T. (2008). Trend estimation and the Hodrick-Prescott filter. *Journal of the Japan Statistical Society, 38*(1), 41–49.

Hodrick, R. J., & Prescott, E. C. (1997). Postwar US business cycles: An empirical investigation. *Journal of Money, Credit and Banking, 29*(1), 1.

Hurvich, C. M., Simonoff, J. S., & Tsai, C. L. (1998). Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. *Journal of the Royal Statistical Society, Series B: Statistical Methodology, 60*(2), 271–293.

Irizarry, R. A. (2004). *Choosing smoothness parameters for smoothing splines by minimizing an estimate of risk* (p. 30). Johns Hopkins University, Department of Biostatistics Working Papers.

Joo, J. H., & Qiu, P. (2009). Jump detection in a regression curve and its derivative. *Technometrics, 51*(3), 289–305.

Julier, S. J., & Uhlmann, J. K. (1997). A new extension of the Kalman filter to nonlinear systems. In *International symposium on aerospace/defense sensing, simulation and controls* (Vol. 3, p. 26). Citeseer.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of Basic Engineering, 82*(1), 35–45.

Kim, C. S., & Marron, J. S. (2006). Sizer for jump detection. *Journal of Nonparametric Statistics, 18*(1), 13–20.

Lee, T. (2003). Smoothing parameter selection for smoothing splines: A simulation study. *Computational Statistics & Data Analysis, 42*(1–2), 139–148.

Leser, C. E. V. (1961). A simple method of trend construction. *Journal of the Royal Statistical Society: Series B: Methodological, 23*(1), 91–107.

Lucia, J. J., & Schwartz, E. S. (2002). Electricity prices and power derivatives: Evidence from the nordic power exchange. *Review of Derivatives Research, 5*(1), 5–50.

Mallat, S. G., & Hwang, W. L. (1992). Singularity detection and processing with wavelets. *IEEE Transactions on Information Theory, 38*(2), 617–643.

Meinl, T., & Sun, E. W. (2012). A nonlinear filtering algorithm based on wavelet transforms for high-frequency financial data analysis. *Studies in Nonlinear Dynamics and Econometrics, 16*(3), 1–24.

Mohr, M. F. (2005). *A trend-cycle (-season) filter*. Working Paper Series European Central Bank, (499).

Morlet, J. (1983). Sampling theory and wave propagation. *Issues in Acoustic Signal/Image Processing and Recognition, 1*, 233–261.

Morton, R., Kang, E. L., & Henderson, B. L. (2009). Smoothing splines for trend estimation and prediction in time series. *Environmetrics, 20*(3), 249–259.

Peltonen, S., Gabbouj, M., & Astola, J. (2001). Nonlinear filter design: Methodologies and challenges. In *Proceedings of the 2nd International Symposium on Image and Signal Processing and Analysis, 2001*. ISPA 2001, (pp. 102–107).

Percival, D. B., & Walden, A. T. (2006). *Wavelet methods for time series analysis*. Cambridge: Cambridge University Press.

Pollock, D. S. G. (2001). Methodology for trend estimation. *Economic Modelling, 18*(1), 75–96.

Qiu, P. (2003). A jump-preserving curve fitting procedure based on local piecewise-linear kernel estimation. *Journal of Nonparametric Statistics, 15*(4), 437–453.

Qiu, P., & Yandell, B. (1998). A local polynomial jump-detection algorithm in nonparametric regression. *Technometrics, 40*(2), 141–152.

Rachev, S. T. (2003). *Handbook of heavy tailed distributions in finance*. North-Holland.

Raimondo, M. (1998). Minimax estimation of sharp change points. *Annals of Statistics, 26*(4), 1379–1397.

Raimondo, M., & Tajvidi, N. (2004). A peaks over threshold model for change-point detection by wavelets. *Statistica Sinica, 14*(2), 395–412.

Ramsey, J. B. (1999). The contribution of wavelets to the analysis of economic and financial data. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences, 357*(1760), 2593–2606.

Ramsey, J. B. (2002). Wavelets in economics and finance: Past and future. *Studies in Nonlinear Dynamics & Econometrics, 6*(3), 1–27.

Ravn, M. O., & Uhlig, H. (2002). On adjusting the Hodrick-Prescott filter for the frequency of observations. *Review of Economics and Statistics, 84*(2), 371–376.

Schlicht, E. (2005). Estimating the smoothing parameter in the so-called Hodrick-Prescott filter. *Journal of the Japan Statistical Society, 35*(1), 99–119.

Seifert, J., & Uhrig-Homburg, M. (2007). Modelling jumps in electricity prices: Theory and empirical evidence. *Review of Derivatives Research, 10*(1), 59–85.

Sun, E. W., & Meinl, T. (2012). A new wavelet-based denoising algorithm for high-frequency financial data mining. *European Journal of Operational Research, 217*(3), 589–599.

Sun, J., & Qiu, P. (2007). Jump detection in regression surfaces using both first-order and second-order derivatives. *Journal of Computational and Graphical Statistics, 16*(2), 289–311.

Sun, W., Rachev, S. T., Fabozzi, F., & Stoyan S. (2008). Multivariate skewed student's t copula in analysis of nonlinear and asymmetric dependence in German equity market. *Studies in Nonlinear Dynamics and Econometrics*, *12*(2).

Tsay, R. S. (2005). *Analysis of financial time series*. Hoboken: Wiley.

Wang, Y. (1995). Jump and sharp cusp detection by wavelets. *Biometrika, 82*(2), 385–397.

Wang, Y. (1998). Change curve estimation via wavelets. *Journal of the American Statistical Association, 93*(441), 163.

Welch, G., & Bishop, G. (1995). *An introduction to the Kalman filter*. Chapel Hill: University of North Carolina at Chapel Hill.

Yuan, L., & Zhongjie, X. (2000). The wavelet detection of the jump and cusp points of a regression function. *Acta Mathematicae Applicatae Sinica (English Series), 16*(3), 283–291.

Analysis of Financial Time Series Using Wavelet Methods

19

Philippe Masset

Contents

Abstract

This chapter presents a set of tools, which allow gathering information about the frequency components of a time series. In a first step, we discuss spectral analysis and filtering methods. Spectral analysis can be used to identify and to quantify the different frequency components of a data series. Filters permit to capture specific components (e.g., trends, cycles, seasonalities) of the original time series. Both spectral analysis and standard filtering methods have two main drawbacks: (i) they impose strong restrictions regarding the possible processes underlying the dynamics of the series (e.g., stationarity) and (ii) they lead to a pure frequency-domain representation of the data, i.e., all information from the time-domain representation is lost in the operation.

In a second step, we introduce wavelets, which are relatively new tools in economics and finance. They take their roots from filtering methods and Fourier analysis, but overcome most of the limitations of these two methods. Their principal

P. Masset
Ecole Hôtelière de Lausanne, Le-Chalet-à-Gobet, Lausanne 25, Switzerland
e-mail: philippe.masset@ehl.ch

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_19,

advantages derive from (i) combined information from both time domain and frequency domain and (ii) their flexibility as they do not make strong assumptions concerning the data-generating process for the series under investigation.

Keywords
Filtering methods • Spectral analysis • Fourier transform • Wavelet filter • Continuous wavelet transform • Discrete wavelet transform • Multiresolution analysis • Scale-by-scale decomposition • Analysis of variance • Case-Shiller home price indices

19.1 Introduction

The purpose of this chapter is to present a set of methods and tools belonging to the so-called *frequency-domain* analysis and to explain why and how they can be used to enhance the more conventional *time-domain* analysis. In essence, time-domain analysis studies the evolution of an economic variable with respect to time, whereas frequency-domain analysis shows at which frequencies the variable is active. We focus on concepts rather than technicalities and illustrate each method with examples and applications using real datasets.

The usual *time-domain* approach aims at studying the temporal properties of a financial or economic variable, whose realizations are recorded at a predetermined frequency. This approach does not convey any information regarding the frequency components of a variable. Thus, it makes the implicit assumption that the relevant frequency to study the behavior of the variable matches with its sampling frequency. An issue arises, however, if the variable realizations depend (in a possibly complicate manner) on several frequency components rather than just one. In such a case, the time-domain approach will not be able to efficiently process the information contained in the original data series.

In this chapter, we start by discussing methods belonging to the *frequency-domain* analysis. These tools are very appealing to study economic variables that exhibit a cyclical behavior and/or are affected by seasonal effects (e.g., GDP, unemployment). Spectral analysis and Fourier transforms can be used to quantify the importance of the various frequency components of the variable under investigation. In particular, they allow inferring information about the length of a cycle (e.g., business cycle) or a phase (e.g., expansion or recession). The presence of such patterns also imposes the use of appropriate methods when it comes to model the dynamics of the variable. Filtering methods have proven useful in this context. Notably, filters may serve to remove specific frequency components from the original data series.

In a second step, we introduce wavelets. During the last two decades, wavelets have become increasingly popular in scientific applications such as signal processing and functional analysis. More recently, these methods have also started to be applied to financial datasets. They are very attractive as they possess the unique ability to provide a complete representation of a data series from both the time and frequency perspectives simultaneously. Hence, they allow breaking down the activity on the market into different frequency components and to study the dynamics of each of these components separately. They do not suffer from some of the limitations of standard frequency-domain methods and can be employed to study a financial variable, whose evolution through time is dictated by the interaction of a variety of different frequency components. These components may also behave according to nontrivial (noncyclical) dynamics – e.g., regime shifts, jumps, and long-term trends.

For instance, the presence of heterogeneous agents with different trading horizons may generate very complex patterns in the time series of stock prices (see Müller et al. 1995; Lynch and Zumbach 2003). This heterogeneity may in particular induce long memory in stock return volatility. In such a case, studying the properties of a time series and trying to model it from the perspective of a single frequency can be misleading. Much information will be lost because of the naive and implicit aggregation of the different frequency components into a single component. Furthermore, as these components may interact in a complicated manner and may be time varying or even nonstationary, standard methods like Fourier analysis are not appropriate. Therefore, one has to resort to more flexible filtering methods like wavelets.

The remaining of this chapter is structured as follows. In Sect. 19.2, we discuss spectral analysis and filtering methods. Section 19.3 is devoted to the presentation of wavelets, Sect. 19.3.1 explains the relevant theoretical background, Sect. 19.3.2 discusses the implementation of these methods, and Sect. 19.3.3 presents a complete case study. Section 19.4 offers a short conclusion.

19.2 Frequency-Domain Analysis

19.2.1 Spectral Analysis: Some Basics and an Example

Studying the properties of an economic variable in the *time domain* is done through time series analysis. Similarly, the purpose of spectral analysis is to study the properties of an economic variable over the frequency spectrum, i.e., in the *frequency domain*. In particular, the estimation of the population spectrum or the so-called power spectrum (also known as the energy-density spectrum) aims at describing how the variance of the variable under investigation can be split into a variety of frequency components. The subject has been extensively researched during the previous 40 years (see Iacobucci (2003) for a short literature review). Our discussion is based primarily on Hamilton (1994) and Gençay et al. (2002).

19.2.1.1 Fourier Transform

The basic idea of spectral analysis is to reexpress a covariance-stationary process x(t) as a new sequence X(f), which determines the importance of each frequency component f in the dynamics of the original series. This is achieved using the discrete version of the Fourier transform[1]

$$X(\mathrm{f}) = \sum_{t=-\infty}^{\infty} \mathrm{x}(\mathrm{t}) e^{-i2\pi f t}, \tag{19.1}$$

where f denotes the frequency at which X(f) is evaluated. In order to gain a deeper insight into this decomposition, one may think about the De Moivre's (Euler's) theorem, which allows to write $e^{-i2\pi f t}$ as

$$\mathrm{e}^{-\mathrm{i}2\pi f t} = \cos(2\pi f \mathrm{t}) - \mathrm{i}\ \sin(2\pi f t). \tag{19.2}$$

Hence application of formula (19.1) is similar to projecting the original signal x(t) onto a set of sinusoidal functions, each corresponding to a particular frequency component. Furthermore, one can use the inverse Fourier transform to recover the original signal x(t) from X(f):

$$\mathrm{x}(\mathrm{t}) = \frac{1}{2\pi} \int_{-\pi}^{\pi} X(\mathrm{f}) e^{-i2\pi f t} df. \tag{19.3}$$

Equation 19.3 shows that X(f) determines how much of each frequency component is needed to synthesize the original signal x(t).

19.2.1.2 Population Spectrum and Sample Periodogram

Following Hamilton (1994), we define the *population* spectrum of x(t) as

$$s_x(\omega) = \frac{1}{2\pi} \sum_{j=-\infty}^{\infty} \gamma_j e^{-i\omega j}, \tag{19.4}$$

where γ_j is the jth autocovariance of x(t)[2]; $\omega = 2\pi \mathrm{f}$ is a real scalar, which is related to the frequency $\mathrm{f} = 1/\tau$ at which the spectrum is evaluated; and τ is the period length of one cycle at frequency f.[3] One may notice that the right part of Eq. 19.4 is indeed the discrete-time Fourier transform of the autocovariance series. There is also a close link between this expression and the autocovariance generating function, which is defined by

$$g_x(\mathrm{z}) = \sum_{j=-\infty}^{\infty} \gamma_j z^{\mathrm{j}}, \tag{19.5}$$

where z denotes a complex scalar. This implies that one can easily recover the autocovariance generating function from the spectrum. In the same spirit, an

[1]The discrete version of the Fourier transform is used because the time series is recorded at discrete-time intervals.

[2]The jth autocovariance of x(t) is given by $\gamma_j = E[(x(t) - \mu)(x(t - j) - \mu)]$, where μ denotes the expected value of x(t).

[3]As an example, let us consider an economic variable, whose evolution is fully determined by the state of the economy. A complete business cycle lasts on average 36 months and therefore f = 1/36 months.

application of the inverse discrete-time Fourier transform allows a direct estimation of the autocovariances from the population spectrum.

In practice, the sample *periodogram* $\hat{s}_x(\omega)$ can be estimated with either nonparametric or parametric approaches. We briefly describe these two approaches hereafter. The first approach is nonparametric because it infers the spectrum from a sample of realizations of the variable x without trying to assign an explicit structure to the data-generating process underlying its evolution. The estimation of the sample periodogram is straightforward as it is directly related to the squared magnitude of the discrete-time Fourier transform $|X(\mathrm{f})|$ of the time series x(t)

$$\hat{s}_x(\omega) = \frac{1}{2\pi}\frac{1}{T}|X(\mathrm{f})|^2, \tag{19.6}$$

where T is the length of the time series x(t). $|X(\mathrm{f})|^2$ is also known as the power spectrum of x(t). This approach is usually called the "periodogram." As noted in Hamilton (1994), its accuracy seems questionable as the confidence interval for the estimated spectrum is typically very broad. Furthermore, the variance of the periodogram does not tend to zero as the length of the data series tends to infinity. This implies that the periodogram is not a consistent estimator of the population spectrum. Therefore, modified versions of the periodogram have been put forward. For instance, smoothed periodogram estimates have been suggested as a way to reduce the noise of the original estimator and to improve its accuracy. The idea underlying this approach is that $s_x(\omega)$ will be close to $s_x(\lambda)$ when ω is close to λ. This suggests that $s_x(\omega)$ might be estimated with a weighted average of the values of $s_x(\omega)$ for values of λ in a neighborhood around ω, where the weights depend on the distance between ω and λ (Hamilton 1994). The weights are typically determined by a kernel weighting function. Welch's method (Welch 1967) and Childers (1978) constitute another alternative based on a simple idea: instead of estimating a single periodogram for the complete sample, one divides the original sample into subsamples, estimates the periodogram for each subsample, and computes the average periodogram over all subsamples.

The second approach is based on some parameterization of the data-generating process of x(t). Methods belonging to this category are close in spirit to the population spectrum, i.e., to a direct application of Eq. 19.4. Typically some specification based on an ARMA (autoregressive moving average) representation is chosen to represent the temporal dynamics of the variable. The model is then calibrated, i.e., the ARMA coefficients are estimated from the realizations of the process x(t). These estimated coefficients are employed to calculate the spectrum. As long as the autocovariances are reasonably well estimated, the results would also be reasonably close to the true values. A detailed discussion of the various parametric methods (e.g., the covariance, Yule-Walker and Burg methods) is beyond the scope of this introduction, but parametric methods are particularly effective when the length of the observed sample is short. This is due to their ability to distinguish the noise from the information contained in the data.

19.2.1.3 Example

We now turn to the discussion of a simple example. We consider a time series, which has the following dynamic:

$$x(t) = a \cdot \cos\left(\frac{2\pi t}{21}\right) + b \cdot \sin\left(\frac{2\pi t}{63}\right) + \varepsilon(t),$$

where $\varepsilon(t)$ is a random term that follows a normal distribution with mean zero and unit variance. One may observe that the process is driven by two cyclical components, which repeat themselves, respectively, each 21 and 63 units of time.

The full line in Fig. 19.1 shows the first 100 (simulated) realizations of x(t); the dotted lines are for the cos and sin functions. At first glance, it seems difficult to distinguish the realizations of x(t) from a purely random process. Figure 19.2 reports the autocorrelations (left panel) and partial autocorrelations (right panel) of $x(t)$ (upper panel) and of the cos and sin components (bottom panel). Again, it remains difficult, when looking at this figure, to gather conclusive evidence concerning the appropriate model specification for $x(t)$.

On the other hand, results from the Fourier analysis, reported in Fig. 19.3, clearly show that two cyclical components drive the evolution of x(t) and repeat themselves around each 21 and 63 units of time. This demonstrates the effectiveness of Fourier methods for the study of processes featuring cyclical components.

19.2.1.4 Illustration: Home Prices in New York City

We now illustrate how spectral methods can be applied to real economic data series. We consider the Case-Shiller home price index for the city of New York. The dataset covers the period January 1987 to December 2011 on a monthly basis. The upper panel of Fig. 19.4 shows the evolution of the index level over this time period, while the lower panel reports the time series of index returns. Results from the Dickey-Fuller test cannot reject the null hypothesis that the index level series is nonstationary. Application of the Fourier transform requires the series under study to be stationary. We therefore study the spectral properties of the index using the time series of returns rather than the levels themselves.

We also estimate the autocorrelations and partial autocorrelations of the index returns up to 48 lags (i.e., 4 years of observations). These are reported in Fig. 19.5. The structure of both the autocorrelations and the partial autocorrelations indicates that the index returns are significantly autocorrelated and it also suggests a cyclical (or seasonal) behavior of the returns. This observation is in line with previous results from the literature (see Kuo 1996; Gu 2002). In order to gain more insight into the presence of such patterns, we compute the power spectrum of the series using parametric and nonparametric methods displayed in Fig. 19.6. The estimated power spectra returned by the two nonparametric methods (periodogram and Welch) are much noisier than the spectra obtained from the parametric methods (Yule-Walker and Burg). The Welch method also seems to result in an oversmoothed estimate of the power spectrum as compared to the other estimates

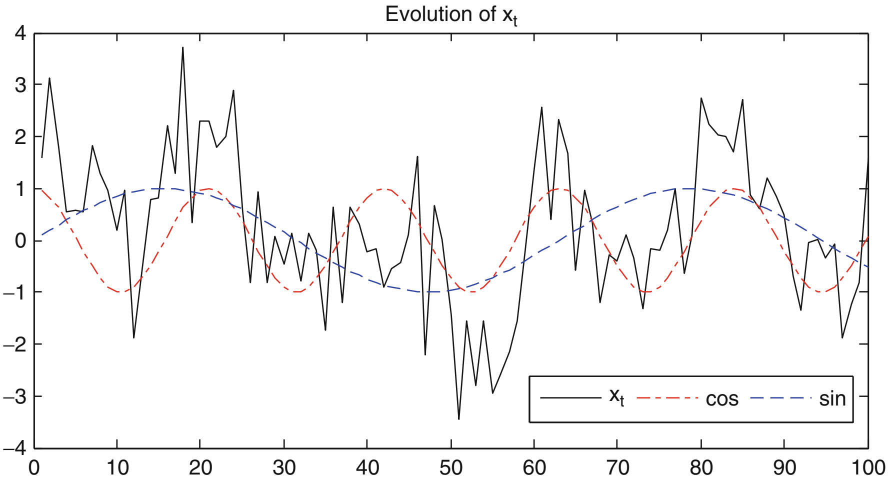


Fig. 19.1 Sample path of $x(t)$

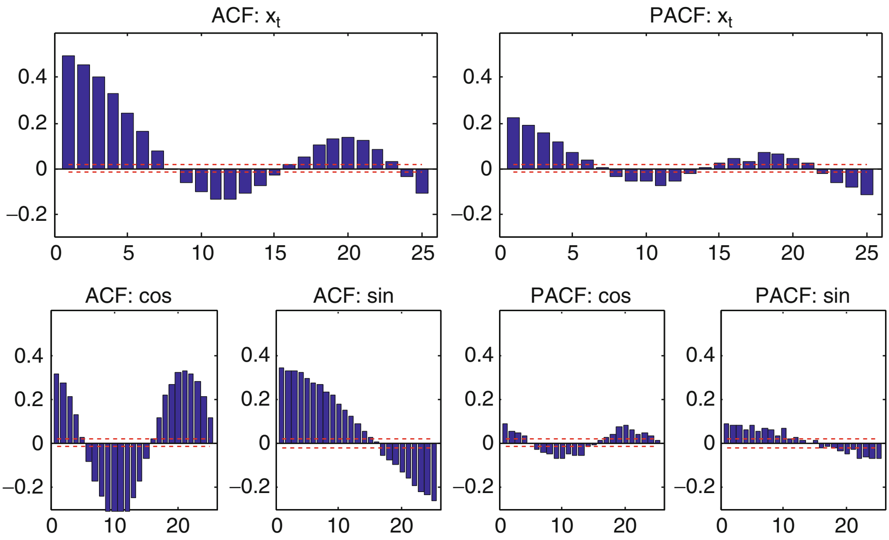


Fig. 19.2 Autocorrelations and partial autocorrelations of $x(t)$. The upper panel reports the autocorrelations (*left panel*) and partial autocorrelations (*right panel*) of $x(t)$. The lower panel shows similar statistics for the cosinus and sinus functions

(e.g., periodogram estimates). On the other hand, the difference between the Yule-Walker and the Burg methods is minimal. Nevertheless, the key message remains remarkably similar: strong seasonalities affect home prices with a frequency of recurrence of 12 months.

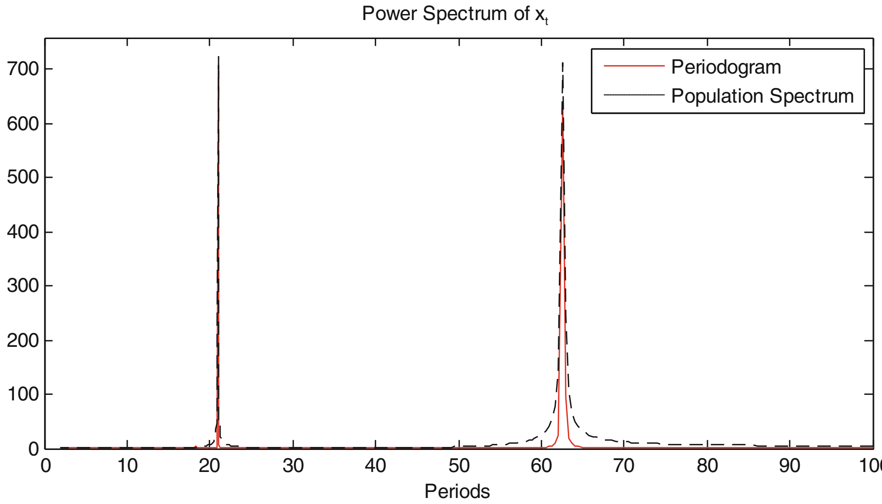


Fig. 19.3 Spectral analysis. The periodogram has been estimated directly from the realizations of $x(t)$. The population spectrum has been estimated on the basis of the theoretical autocovariances of $x(t)$

19.2.2 Filtering Methods

A filter is a mathematical operator that serves to convert an original time series $x(t)$ into another time series $y(t)$. The filter is applied by convolution of the original series $x(t)$ with a coefficient vector w:

$$y(t) = (w * x)(t) = \sum_{k=-\infty}^{\infty} w(k)x(t-k), \tag{19.7}$$

The purpose of this operation is to explicitly identify and to extract certain components from $x(t)$. In the present context, one may want to remove from the original time series some particular features (e.g., trends, business cycles, seasonalities, or noise) that are associated with specific frequency components.

19.2.2.1 Frequency Response Function

Filters in the time domain can be characterized on the basis of their *impulse-response function*, which traces the impact of a one-time unit impulse in $x(t)$ on subsequent values of $y(t)$. Similarly, in the frequency domain, the analysis of the *frequency response function* (or transfer function) of a filter tells us which frequency components the filter captures from the original series. The frequency response function is defined as the Fourier transform of the filter coefficients

$$H(f) = \sum_{k=-\infty}^{\infty} w(k)e^{-i2\pi fk}. \tag{19.8}$$

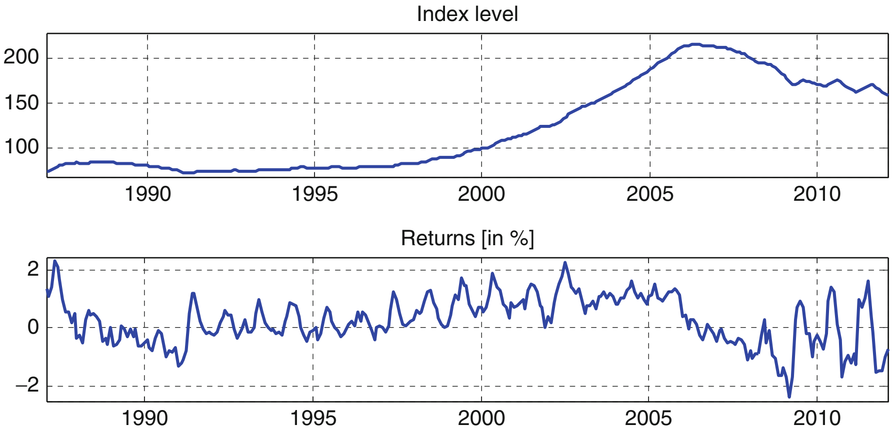


Fig. 19.4 Case-Shiller home price index for the city of New York. Price levels and returns are reported in the *upper* and lower panel respectively.

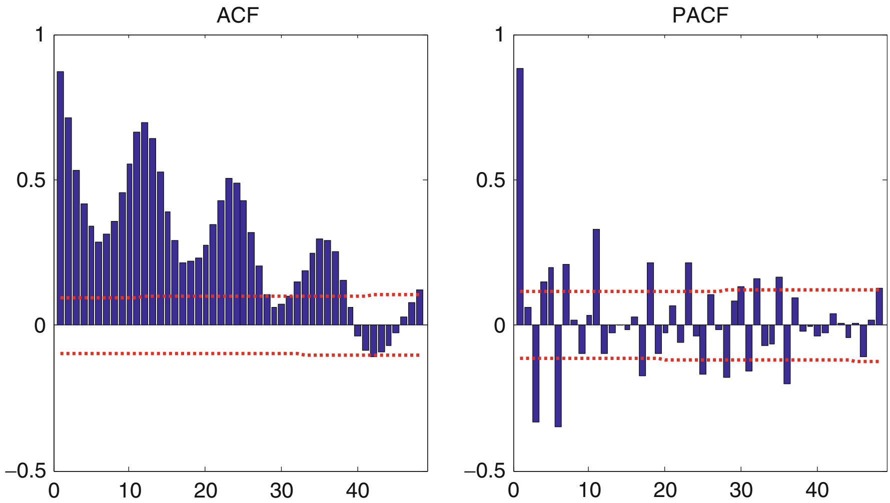


Fig. 19.5 Autocorrelations and partial autocorrelation of the returns on the New York home price index

The frequency response, $H(f)$, can be further split into two parts:

$$H(f) = G(f)e^{i\theta(f)}, \tag{19.9}$$

where $G(f)$ is the *gain* function, $e^{i\theta(f)}$ the *phase* function, and θ the phase angle (or equivalently the argument of $H(f)$). The gain function is the magnitude of the frequency response, i.e., $G(f) = |H(f)|$. If the application of the filter on $x(t)$ results in a phase shift, i.e., if the peaks and lows of $x(t)$ and $y(t)$ have a different timing, the phase angle θ will be different from zero. The use of uncentered moving average

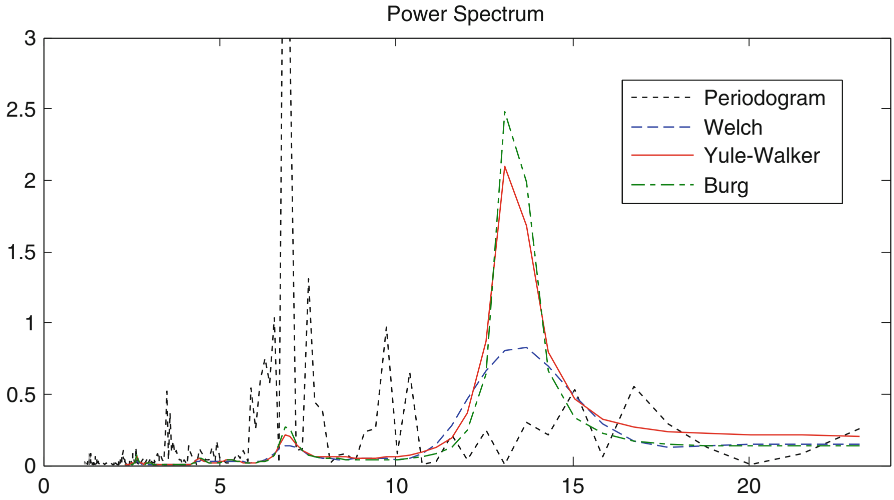


Fig. 19.6 Spectral analysis of the return series. Comparison between parametric (Burg and Yule-Walker) and non-parametric (Periodogram and Welch) methods

filters leads to this (often) undesirable feature because turning points will be recorded earlier in the original series than in the filtered series. On the other hand, centered (symmetric) moving averages have $\theta(f) = 0$; hence, there is no phase shift for this class of filters. For instance, the frequency response of a two-period uncentered moving average filter with coefficients $w(k) = 0.5$ for $k = 0, 1$ is[4]

$$\begin{aligned} H(f) &= \sum\nolimits_{k=0}^{1} 0.5e^{-i2\pi fk} \\ &= 0.5 + 0.5e^{-i2\pi f} \\ &= 0.5\left(e^{i\pi f} + e^{-i\pi f}\right)e^{-i\pi f} \\ &= \cos(\pi f)e^{-i\pi f} \end{aligned}$$

This result shows the existence of a phase shift as the phase angle is $\theta(f) = -\pi f$.[5] Hence, the turning points from the original series will be shifted to the right in the filtered series. On the other hand, a filter with coefficients w(k) = 1/3 for k = −1, 0, 1 has a zero phase angle.

Based on their gain functions, filters can be categorized as follows:

- High-pass filters should be able to capture the high-frequency components of a signal, i.e., the value of their gain function $G(f)$ should equal one for frequencies f close or equal to 1/2.

[4]See also Gençay et al. (2002) who uses a similar example.

[5]In full generality, the phase angle can be computed as $\theta(f) = \arctan\left(\frac{Im[H(f)]}{\mathrm{Re}[H(f)]}\right)$, where $Im[H(f)]$ and $\mathrm{Re}[H(f)]$ are, respectively, the imaginary part and the real part of $H(F)$.

- Low-pass filters should be able to capture the low-frequency components of a signal, i.e., $G(f) = 1$ for f close or equal to 0.
- Band-pass filters should be able to capture a range of frequency components of a signal, i.e., $G(f) = 1$ for $f_{lo} < f < f_{hi}$.
- All-pass filters capture all the frequency components of a signal, i.e., $G(f) = 1$ for $\forall f$. Such filters leave the frequency components of the original signal unaltered.

Filters are commonly used in economics. The Hodrick and Prescott (1997) filter is probably the best known. It has a structure and a gain function which enable it to capture business cycle components.

19.2.2.2 Example

A basic example of a high-pass filter is a filter that takes the difference between two adjacent values from the original series; its coefficients are $w_{hi} = [0.5, -0.5]$. Similarly, the most simple low-pass filter is a 2-period moving average; in this case $w_{lo} = [0.5, 0.5]$. In wavelet theory, $w_{hi}/\sqrt{2}$ and $w_{lo}/\sqrt{2}$ form the Haar wavelet family. In this case, the low-pass filter w_{lo} is basically an averaging filter, while the high-pass filter w_{hi} is a differencing filter. The gain functions for these two filters are reported in Fig. 19.7.

19.2.2.3 Illustration

The full line in the left panels of Fig. 19.8 shows the monthly (unadjusted) returns on the Case-Shiller New York home price index from 1987 to 2011. We add to the top panel the output series resulting from the application of both a centered and an uncentered (causal) 3-period moving average on the original data. The three-filter coefficients have a value of 1/3. This implies that the output series of the centered moving average at time t is basically the average return from months $t - 1$ to $t + 1$. Similarly, the uncentered moving returns the average return from $t - 2$ to t. It is apparent from the figure that the uncentered moving average leads to a phase shift of 1 month. The bottom panel shows the outputs of a 7-period moving average. The filter coefficients are equal to 1/7. Again, we consider both centered and uncentered filters. The use of an uncentered moving average leads to a phase shift of 3 months as compared to the centered moving average.

The right part of the figure reports the gain functions for the 3-period and the 7-period moving averages. The gain function is similar for both the uncentered and the centered moving average. The only element that distinguishes these two filters is indeed their phase function. One may notice that (i) both filters are low-pass filters, and (ii) the longer filter captures more efficiently the low-frequency components of the original signal than the short filter.

19.3 Scale-by-Scale Decomposition with Wavelets

To a large extent, wavelets can be seen as a natural extension to spectral and Fourier analysis as (i) wavelets do not suffer the weaknesses of Fourier analysis, and (ii) wavelets provide a more complete decomposition of the original time series than Fourier analysis does.

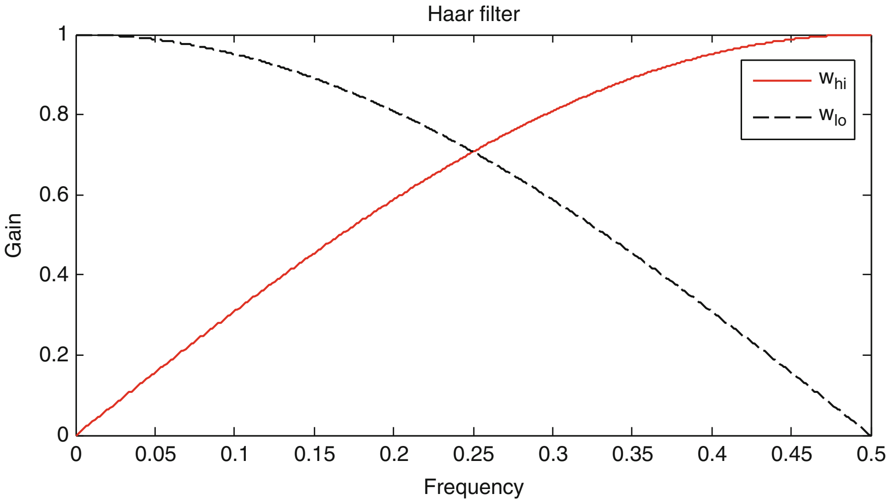


Fig. 19.7 Gain functions of the high-pass (differencing) and low-pass (averaging) Haar filters

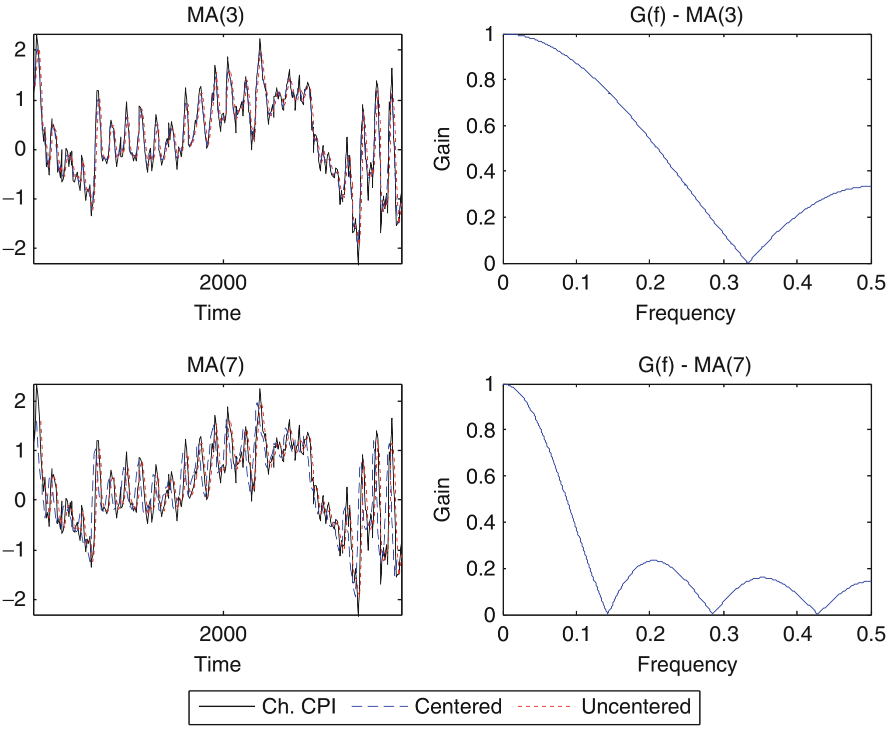


Fig. 19.8 Two moving averages and their frequency responses. *Left panels* show the outputs from a 3-period (*upper panel*) and a 7-period (*lower panel*) moving average filters applied on the returns on the Case-Shiller New York home price index. *Right panels* report the corresponding gain functions

There are some problems with spectral methods and Fourier transforms. Notably, these methods require the data under investigation to be stationary. This is often not the case in economics and finance. In particular, volatility is known to exhibit complicated patterns like jumps, clustering, and long memory. Furthermore, the frequency decomposition delivered by Fourier analysis only makes sense if the importance of the various frequency components remains stable over the sample period. Ex ante, there is good reason to expect this assumption not to hold for a variety of economic and financial variables. For instance, volatility changes are likely to exhibit a different frequency spectrum when trading activity is intense than when the market is quiet. The short-time Fourier transform (which is also known as the Gabor or windowed Fourier transform) has been suggested to overcome these difficulties. The idea is to split the sample into subsamples and to compute the Fourier transform on these subsamples. Hence, this extension achieves a better trade-off between the time and the frequency representation of the original data. Nevertheless, this provides at best a partial solution to the aforementioned issues as the strong restrictions regarding the possible data-generating process over each subsample are maintained.

Wavelets do not make any of these assumptions. Furthermore, wavelets provide a complete decomposition of the original series, which is located both in time and in frequency. From a mathematical viewpoint, a wavelet is a function, which enables to split a given signal into several components, each reflecting the evolution trough time of the signal at a particular frequency. Wavelet analysis has originally been used in signal processing (e.g., image processing and data compression). Its applications to economics and finance are relatively recent. Nevertheless, the range of application of wavelets is potentially wide: denoising and seasonality filtering, decorrelation and estimation of fractionally integrated models, identification of regime shift and jumps, robust estimation of the covariance and correlation between two variables at different time scales, etc.

From a physicist perspective, but with application to time series analysis, Percival and Walden (2000) and Struzik (2001) provide a mathematically rigorous and exhaustive introduction to wavelets. Struzik (2001) particularly emphasizes the unique ability of nonparametric methods (like wavelets) to let the data speak by themselves. Thus, such methods avoid making misleading interpretations of the coefficients obtained from the calibration of misspecified models. Gençay et al. (2002) discuss the use of wavelets for specific purposes in economics and finance and adopt a more intuitive approach (with many illustrations and examples), while Ramsey (2002) surveys the most important properties of wavelets and discusses their fields of application in both economics and finance. Crowley (2007) proposes a genuine guide to wavelets for economists. His article can be considered as a complete and easily understandable toolkit, precisely explaining in which circumstances to use wavelets and how to proceed. Schleicher (2002) is a complementary reference to those already named; Schleicher focuses on some mathematical concepts underlying the use of wavelets and discusses them in details using examples.

This short literature review focuses only on textbook-style references. There are however quite a large amount of economic and financial articles that have employed wavelets for empirical purposes. Ramsey[6] and Gençay[7] and their respective coauthors can be considered as the pioneers of the use of wavelets in economics and finance. Other recent contributions include Nielsen and Frederiksen (2005), Vuorenmaa (2005), Oswiecimka et al. (2006), Elder and Jin (2007), Fan et al. (2007), Fernandez and Lucey (2007), Rua and Nunes (2009), Manchaldore et al. (2010), Curci and Corsi (2012), and Hafner (2012).

19.3.1 Theoretical Background

19.3.1.1 What Is a Wavelet?

As its name suggests, a wavelet is a small wave. In the present context, the term "small" essentially means that the wave grows and decays in a limited time frame. Figure 19.9 illustrates this notion by contrasting the values taken by a simple wavelet function (the Morlet function)[8] and the values of the sin function, which can be considered as a sort of "large" wave. In order to clarify the notion of small wave, we start by introducing a function, which is called the *mother wavelet* and is denoted by $\psi(t)$. This function is defined on the real axis and must satisfy two conditions:

$$\int_{-\infty}^{\infty} \psi(t)dt = 0. \tag{19.10}$$

$$\int_{-\infty}^{\infty} |\psi(t)|^2 dt = 1. \tag{19.11}$$

Taken together, these conditions imply (i) that at least some coefficients of the wavelet function must be different from zero and (ii) that these departures from zero must cancel out. Clearly the sin function does not meet these two requirements. A vast variety of functions meets conditions (19.10) and (19.11). Nevertheless, these conditions are very general and not sufficient for many practical purposes. Therefore, one has to impose additional conditions in order to run a specific analysis with wavelets. One of these conditions is the so-called *admissibility condition*, which states that a wavelet function is admissible if its Fourier transform

$$\Psi(f) = \int_{-\infty}^{\infty} \psi(t)e^{-i2\pi ft}dt, \tag{19.12}$$

is such that

$$C_\Psi = \int_0^{\infty} \frac{|\Psi(f)|^2}{f} df \text{ satisfies } 0 < C_\Psi < \infty. \tag{19.13}$$

These conditions allow reconstructing a function from its continuous wavelet transform (see Percival and Walden (2000) for more details).

[6]See Ramsey et al. (1995), Ramsey and Zhang (1997), and Ramsey (1999).

[7]See Gençay et al. (2003), Gençay et al. (2010), Gençay and Fan (2010), and Gençay and Gradojevic (2011).

[8]The Morlet wavelet is actually similar to a sin curve modulated by a Gaussian envelope.

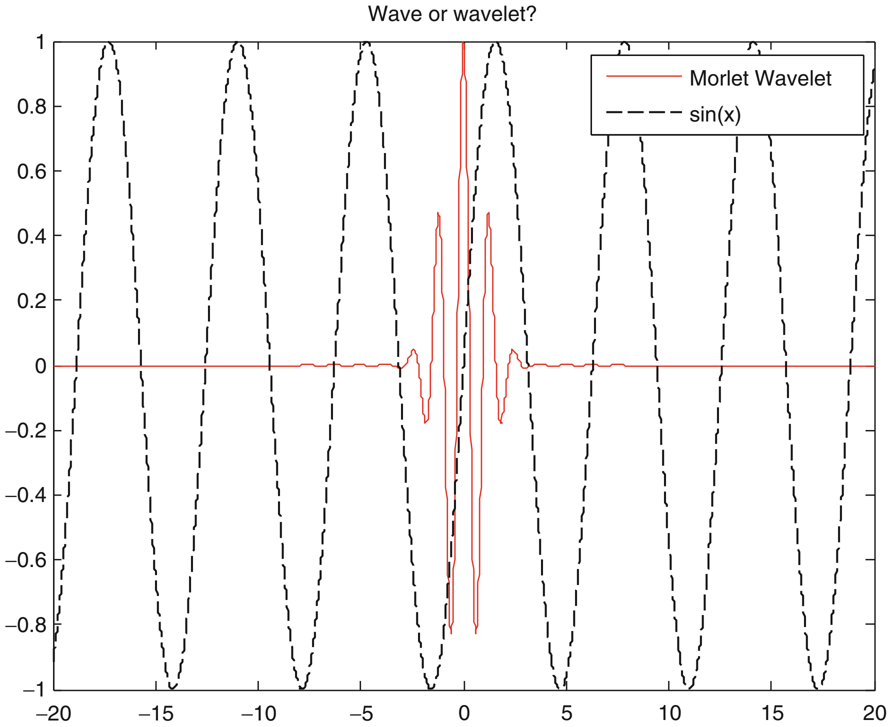


Fig. 19.9 The Morlet wavelet and the sin function

19.3.1.2 The Continuous Wavelet Transform (CWT)

As a starting point, we discuss the CWT. The CWT primarily aims at quantifying the change in a function at a particular frequency and at a particular point in time. In order to be able to achieve this, the mother wavelet $\psi(t)$ is dilated and translated:

$$\psi_{u,s}(t) = \frac{1}{\sqrt{s}}\psi\left(\frac{t-u}{s}\right), \tag{19.14}$$

where u and s are the location and scale parameters. The term $\frac{1}{\sqrt{s}}$ ensures that the norm of $\psi_{u,s}(t)$ is equal to one. The CWT, $W(u, s)$, which is a function of the two parameters u and s, is then obtained by projecting the original function $x(t)$ onto the mother wavelet $\psi_{u,s}(t)$:

$$W(u,s)=\int_{-\infty}^{\infty} x(t)\psi_{u,s}(t)dt. \tag{19.15}$$

To assess the variations of the function on a large scale (i.e., at a low frequency), a large value for s will be chosen and vice versa. By applying the CWT for a continuum of location and scale parameters to a function, one is able to decompose the function under study into elementary components. This is particularly interesting for studying a function with a complicated structure, because this procedure allows extracting a set of "basic" components that have a simpler structure than the original function. By "synthesizing" $W(u, s)$, it is also possible to reconstruct the original function $x(t)$ (see Gençay et al. (2002) for more details).

In empirical applications, several difficulties with the CWT occur. First, it is computationally impossible to analyze a signal using *all* wavelet coefficients. CWT is thus more suitable for studying functions than signals or (economics) time series. Second, as noted by Gençay et al. (2002), $W(u, s)$ is a function of two parameters and as such it contains a high amount of redundant information. We therefore turn to the discussion of the discrete wavelet transform (DWT).

19.3.1.3 The Discrete Wavelet Transform (DWT)

The core difference between the CWT and the DWT is that the latter does not use all translated and dilated versions of the mother wavelet to decompose the original signal (Gençay et al. 2003). The idea is to select u and s such that the information contained in the signal can be summarized in a minimum of wavelet coefficients. This objective is achieved by setting

$$s = 2^{-j} \text{ and } u = k2^{-j},$$

Where j and k are integers representing the set of discrete translations and discrete dilatations. Gençay et al. (2002) refer to this procedure as the *critical sampling* of the CWT. This implies that the wavelet transform of the original function or signal is calculated only at *dyadic* scales, i.e., at scales 2^j. A further implication is that for a time series with T observations, the largest number of scales for the DWT is equal to the integer J such that $J = \lfloor log_2(T) \rfloor = \lfloor log(T)/log(2) \rfloor$. It is not possible to directly apply the DWT if the length of the original series is not dyadic (i.e., if $J < log_2(T) < J + 1$). In such case, one has either to remove some observations or to "complete" the original series in order to have a series of dyadic length. Several methods exist to deal with this kind of boundary problems (see Sect. 19.3.2).

The DWT is based on two discrete wavelet filters, which are called the *mother wavelet* $h_l = (h_l, \ldots, h_{L-1})$ and the *father wavelet* $g_1 = (g_1, \ldots, g_{L-1})$. The mother wavelet is characterized by three basic properties:

$$\sum_{l=0}^{L-1} h_l = 0, \sum_{l=0}^{L-1} {h_l}^2 = 1, \text{and} \sum_{l=0}^{L-1} h_l h_{l+2n} = 0 \text{ for all integers } n \neq 0 \quad (19.16)$$

These three properties ensure that (i) the mother wavelet is associated with a difference operator, (ii) the wavelet transform preserve the variance of the original data, and (iii) a multiresolution analysis can be performed on a finite variance data series. The first property implies that the mother wavelet (also called "differencing function") is a high-pass filter as it measures the deviations from the smooth components. On the other hand, the father wavelet ("scaling function") aims at capturing long-scale (i.e., low-frequency) components of the series and generates the so-called scaling coefficients.[9]

The mother and father wavelets must respect the following conditions:

$$\sum_{l=0}^{L-1} h_l = 0. \tag{19.17}$$

$$\sum_{l=0}^{L-1} g_l = 1. \tag{19.18}$$

The application of both the mother and the father wavelets allows separating the low-frequency components of a time series from its high-frequency components. Furthermore, a band-pass filter can be constructed by recursively applying a succession of low-pass and high-pass filters.

Let's assume that we have observed a sample of size T of some random variable $x(t)$, $\{x(1), x(2), \ldots, x(T)\}$. The *wavelet* and *scaling coefficients* at the first level of decomposition are obtained by convolution of the data series with the mother and the father wavelets:

$$w_1(t) = \sum_{l=0}^{L-1} h_l x(t') \text{ and } v_1(t) = \sum_{l=0}^{L-1} g_l x(t') \tag{19.19}$$

where $t = 0,1, \ldots, T/2 - 1$ and t' the time subscript of x is defined as $t' = 2t + 1 - l \bmod T$. The modulus operator is employed to deal with boundary conditions.[10] It ensures that the time subscript of x stays always positive. If, for some particular values of t and l, the expression $2t + 1 - l$ becomes negative, the application of the modulus operator returns $t' = 2t + 1 - l + T$. Thus, we are implicitly assuming that x can be regarded as periodic. Alternative methods to deal with boundary conditions are discussed thereafter. $w_1(t)$ and $v_1(t)$ are, respectively, the wavelet and the scaling coefficients at the first scale. Hence, $w_1(t)$ corresponds to the vector containing the components of x recorded at the highest frequency. One may notice that the operation returns two series of coefficients that have length $T/2$. To continue the frequency-by-frequency decomposition of the original signal, one typically resorts to what is known as the pyramid algorithm.

19.3.1.4 Pyramid Algorithm

After having applied the mother and father wavelets on the original data series, one has a series of high-frequency components and a series of lower-frequency

[9]The low-pass filter can be directly obtained from the high-pass filter using the quadrature mirror relationship; see Percival and Walden (2000, p. 75).

[10]For two integer a and b, a modulus b is basically the remainder after dividing a by b, i.e., $a \bmod b = a - c \cdot b$ with $c = \lfloor a/b \rfloor$.

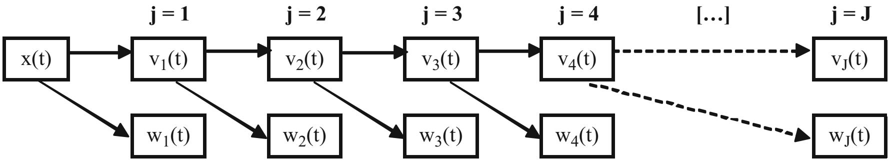


Fig. 19.10 Flowchart of the pyramid algorithm

components. The idea of the pyramid algorithm is to further decompose the (low-frequency) scaling coefficients $v_1(t)$ into high- and low-frequency components:

$$w_2(t) = \sum\nolimits_{l=0}^{L-1} h_l v_1(t') \text{ and } v_2(t) = \sum\nolimits_{l=0}^{L-1} g_l v_1(t'), \tag{19.20}$$

where $t = 0,1, \ldots,T/4 - 1$ and $t' = 2t + 1 - l \bmod T$. After two steps, the decomposition looks like $w = [w_1 \ w_2 \ v_2]$. One can then apply the pyramid algorithm again and again up to scale $J = \lfloor log_2(T) \rfloor$ to finally obtain $w = [w_1 \, w_2 \ldots w_j \, v_\text{j}]$. Figure 19.10 summarizes these steps. One may also apply the algorithm up to scale $J_p < J$ only. This is known as the *partial* DWT.

19.3.1.5 The Maximal Overlap Discrete Wavelet Transform (MODWT)

The standard DWT suffers from three drawbacks. First, it requires a series with a dyadic length. Second, DWT is not shift invariant, i.e., if one shifts the series one period to the right, the multiresolution coefficients will be different. Third, it may introduce phase shifts in the wavelet coefficients: peaks or troughs in the original series may not be correctly aligned with similar events in the multiresolution analysis. To overcome these problems, the MODWT has been proposed. This wavelet transform can handle any sample size, it has an increased resolution at coarser scales (as compared to the DWT), and it is invariant to translation. It also delivers a more asymptotically efficient wavelet variance than the DWT.[11]

The main difference between the DWT and the MODWT lies in the fact that the MODWT considers all integer translations, i.e., $u = k$. This means that the MODWT keeps at each frequency a complete resolution of the series. Whatever the scale considered, the length of the wavelet and scaling coefficient vectors will be equal to the length of the original series. The wavelet and scaling coefficients at the first level of decomposition are obtained as follows:

$$\widetilde{w}_I(t) = \sum\nolimits_{l=0}^{L-1} \widetilde{h}_l x(t') \text{ and } \widetilde{v}_1(t) = \sum\nolimits_{l=0}^{L-1} \widetilde{g}_l x(t'), \tag{19.21}$$

where $t = 0,1,\ldots T$ and $t' = t - l \bmod T$. As for the DWT, the MODWT coefficients for scales $j > 1$ can be obtained using the pyramid algorithm. For instance, $\widetilde{w}_j$ and $\widetilde{v}_j$ are calculated as

[11]See Crowley (2007) for more details about the properties of MODWT.

$$\widetilde{w}_j(t) = \sum_{l=0}^{L-1} \widetilde{h}_l \widetilde{v}_{j-1}(t') \text{ and } \widetilde{v}_j(t) = \sum_{l=0}^{L-1} \widetilde{g}_l \widetilde{v}_{j-1}(t'), \tag{19.22}$$

where $t' = 2^{j-1}\, l \bmod T$.

Using matrix notation, we can conveniently calculate the wavelet and scaling coefficients up to scale J. We first define a matrix $\widetilde{W}$ that is composed of $J + 1$ sub-matrices, each of them T x T:

$$\widetilde{W} = \begin{bmatrix} \widetilde{W}_1 \\ \widetilde{W}_2 \\ \vdots \\ \widetilde{W}_J \\ \widetilde{V}_J \end{bmatrix} \tag{19.23}$$

Each $\widetilde{W}_j$ has the following structure:

$$\widetilde{W}_j = \begin{bmatrix} \widetilde{h}_0/2^{j/2} & 0 & 0 & \cdots & 0 & \widetilde{h}_{L-1}/2^{j/2} & \cdots & \widetilde{h}_1/2^{j/2} \\ \widetilde{h}_1/2^{j/2} & \widetilde{h}_0/2^{j/2} & 0 & \cdots & 0 & 0 & \cdots & \vdots \\ \vdots & \widetilde{h}_1/2^{j/2} & \widetilde{h}_0/2^{j/2} & \cdots & 0 & 0 & \cdots & \widetilde{h}_{L-1}/2^{j/2} \\ \widetilde{h}_{L-1}/2^{j/2} & \vdots & \widetilde{h}_1/2^{j/2} & \cdots & 0 & 0 & \cdots & 0 \\ 0 & \widetilde{h}_{L-1}/2^{j/2} & \vdots & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \widetilde{h}_{L-1}/2^{j/2} & \cdots & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & 0 & \cdots & \widetilde{h}_0/2^{j/2} & 0 & \cdots & 0 \\ 0 & 0 & 0 & \cdots & \widetilde{h}_1/2^{j/2} & \widetilde{h}_0/2^{j/2} & \cdots & 0 \\ 0 & 0 & 0 & \cdots & \vdots & \vdots & \cdots & 0 \\ 0 & 0 & 0 & \cdots & \widetilde{h}_{L-1}/2^{j/2} & \widetilde{h}_{L-2}/2^{j/2} & \cdots & \widetilde{h}_0/2^{j/2} \end{bmatrix} \tag{19.24}$$

$\widetilde{V}_J$ has a similar structure as $\widetilde{W}_J$ but it contains the coefficients associated to the father wavelet instead of the mother wavelet.

We can now directly calculate all wavelet and scaling coefficients via

$$\widetilde{w} = \widetilde{W} x, \tag{19.25}$$

where $\widetilde{w}$ is a vector made up of $J + 1$ length T vectors of wavelet and scaling coefficients, $\widetilde{w}, \ldots, \widetilde{w}_J$ and $\widetilde{v}_j$; i.e., $\widetilde{w} = \left[\widetilde{w}_1 \widetilde{w}_2 \widetilde{w}_J \cdots \widetilde{v}_J\right]^T$.

19.3.1.6 Multiresolution Analysis (MRA)

Multiresolution analysis can be used to reconstruct the original time series x from the wavelet and scaling coefficients, $\widetilde{v}_j$ and $\widetilde{w}_j$. In order to achieve this, one has to apply the inverse MODWT on $\widetilde{v}_j$ and $\widetilde{w}_j$, $j = 1, \ldots, J$.[12]

[12]Our presentation of the multiresolution analysis is restricted to the case of the MODWT. Nevertheless, a very similar procedure exists for the DWT; see Percival and Walden (2000).

$\widetilde{W}$ is an orthonormal matrix, as such $\widetilde{W}^T\widetilde{W} = 1$. Hence, if we multiply both side of Eq. 19.25 by $\widetilde{W}^T$, we get

$$\widetilde{W}^T \widetilde{w} = \widetilde{W}^T\widetilde{W}x = x. \tag{19.26}$$

As $\widetilde{W}^T = \left[\widetilde{W}_1\ \widetilde{W}_2 \cdots \widetilde{W}_J\ \widetilde{V}_J\right]$ and $\widetilde{w} = \left[\widetilde{w}_1\ \widetilde{w}_2\ \widetilde{w}_J \cdots \widetilde{v}_J\right]^T$, we can further rearrange Eq. 19.26 and show that

$$x = \sum\nolimits_{j=1}^{J} \widetilde{W}_j{}^T \widetilde{w}_j + \widetilde{V}_J{}^T \widetilde{v}_J. \tag{19.27}$$

Setting $D_j = \widetilde{W}_j{}^T \widetilde{w}_j$ and $S_j = \widetilde{V}_j{}^T \widetilde{v}_j$, we can reconstruct the original time series as

$$x = D_1 + \ldots + D_J + S_J. \tag{19.28}$$

This "reconstruction" is known as multiresolution analysis (MRA). The elements of S_j are related to the scaling coefficients at the maximal scale and therefore represent the *smooth components* of x. The elements of D_j are the *detail* (or *rough*) *coefficients* of x at scale j.

On the basis of formula (19.28), one may also think of a way to compute an approximation or a smooth representation of the original data. This can be achieved by considering the scaling coefficients and the wavelet coefficients from scale J_s ($J_s < J$) to J only, i.e.:

$$x_S = D_S + \ldots + D_J + S_J. \tag{19.29}$$

Equation 19.29 can be used, for instance, to filter out noise or seasonalities from a time series. In image processing, Eq. 19.29 serves for data compression. Equation 19.27 has been specifically derived for the MODWT but similar results are available for the DWT (see Percival and Walden 2000).

19.3.1.7 Analysis of Variance

On the basis of the wavelet and scaling coefficients, it is also possible to decompose the variance into different frequency components. There are some slight differences between the variance decomposition for the DWT and for the MODWT. We will therefore first present the main results for the DWT and then discuss their extension to the MODWT.

Using the wavelet and scaling coefficients of the discrete wavelet transform, it is possible to decompose the energy of the original series on a scale-by-scale basis:

$$\|x\|^2 = \sum_{t=0}^{T-1} x(t)^2 = \sum_{j=1}^{J} \sum_{t=0}^{T/2^J-1} w_j(t)^2 + v_J(t)^2, \tag{19.30}$$

where $\|x\|^2$ denotes the energy of x. The wavelet coefficients capture the deviations of x from its long-run mean at the different frequency resolutions. Therefore, at scale $j = J = log_2(T)$, the last remaining scaling coefficient is equal to the sample mean of x,

$$E(x) = v_J. \tag{19.31}$$

On this basis, we can express the variance of x as

$$V(x) = E(x^2) - E(x)^2 = \sum_{j=1}^{J} E(w_j{}^2) = \sum_{j=1}^{J} V(w_j), \tag{19.32}$$

where $V(w_j)$ denotes the variance of the wavelet coefficients at scale j.

If we consider the wavelet and scaling coefficients obtained from a partial DWT, the variance of x can be expressed as

$$V(x) = \sum_{j=1}^{J_p} V(w_j) + V(v_{J_p}). \tag{19.33}$$

The variance of the scaling coefficients has to be taken into account because v_{J_p} incorporates deviations of x from its mean at scales $J_p < j < J$.[13]

An alternative way to decompose the energy of x is based on the smooth and detail coefficients of the MRA. As above, $\|x\|^2$ can be computed as the sum of the energy of the smooth and detail coefficients. This approach is, however, valid only for the DWT (See Gençay et al. 2002).

It is important to note that some of the wavelet coefficients involved in Eq. 19.32 are affected by boundary conditions. One should remove the corresponding wavelet coefficients in order to get an unbiased estimator of the wavelet variance:

$$\hat{V}(x) = \sum_{j=1}^{J} \left[\frac{1}{2\lambda_j \hat{T}_j} \sum_{t=L_j'}^{\frac{T}{2^j}-1} w_j(t)^2 \right], \tag{19.34}$$

where λ_j is the scale that is associated to the frequency interval $[1/2^{j+1}\ 1/2^j]$. $L_j' = \lceil (L-2)(1-2^{-j}) \rceil$ is the number of DWT coefficients computed using the boundaries. Hence, $\hat{T}_j = T/2^j - L_j'$ is the number of coefficients unaffected by the boundary.

[13]One may notice that the variance of the scaling coefficient at scale J is 0 as v_J is a scalar (the sample mean of x).

We now turn to the analysis of variance in the context of the MODWT. Equation 19.32 remains perfectly valid. The MODWT keeps the same number of coefficients at each stage of the wavelet transform. The way of dealing with boundary conditions must therefore be adapted. From the detail coefficients of a partial MODWT of order $J_p < log_2(T)$, the wavelet variance can be estimated as follows:

$$\hat{V}(x) = \sum_{j=1}^{J_p} \left[\frac{1}{\hat{T}_j} \sum_{t=L_j-1}^{T-1} \tilde{w}_j(t)^2 \right] \tag{19.35}$$

where $L_j = (2^j - 1)(L - 1) + 1$ is the number of scale λ_j wavelet coefficients, which are affected by boundary conditions. This number also corresponds to the length of the wavelet filter at scale λ_j. $\hat{T}_j = T - L_j + 1$ is thus the number of wavelet coefficients unaffected by the boundary.

19.3.2 Implementation and Practical Issues

19.3.2.1 Choice of a Wavelet Filter

Many different wavelet filters exist with each of them being particularly suitable for specific purposes of analysis. Wavelet filters differ in their properties and in their ability to match with the features of the time series under study. Furthermore, when it comes to implement a *discrete* wavelet transform, one also has to decide about the filter length. Because of boundary conditions, longer filters are well adapted for long time series. The simplest filter is the Haar wavelet, which is basically a difference and average filter of length two. In finance, most researchers have worked either with Daubechies (denoted as "D") or with Least-Asymmetric ("LA") filters of length 4–8. Elder and Jin (2007) and Nielsen and Frederiksen (2005) employ D(4) filters. Gençay et al. (2003, 2010) suggest that the LA (8) wavelet (i.e., a Least-Asymmetric filter of length 8) is a good choice for analyzing financial time series, while Subbotin (2008) uses a LA(4) wavelet. Crowley (2007) argues that the impact of choosing another wavelet filter has a rather limited impact on the distribution of the variance of the time series across the scales.

Depending on the purpose of the analysis, it might be appealing to select a wavelet filter which satisfies one or more of the following properties:

- Symmetry: symmetric filters are appealing as they ensure that there will be no phase shift in the output series. Unfortunately, most wavelets are not symmetric. An exception is the Haar wavelet. The requirement of a symmetric wavelet is, however, less essential if a MODWT is used as it ensures that the original series and its filter coefficients will be aligned.
- Orthogonality: this property refers to the fact that the wavelet and the scaling coefficients contain different information. This is an important feature as it allows for the wavelet decomposition to preserve the energy (variance) of the original series (Crowley 2007). Daubechies and Least-Asymmetric wavelets

meet this requirement, that is, their scaling and wavelet coefficients are orthogonal by construction.

- Smoothness: The degree of smoothness is measured by the number of continuous derivatives of the basis function. As such, the Haar wavelet is the least smooth wavelet. The choice of a more or less smooth filter depends essentially on the data series to be represented. If the original time series is very smooth, then one will opt for a smooth wavelet. For instance, the Haar wavelet is appropriate for the analysis of a pure jump process.
- Number of vanishing moments: The number of vanishing moments of the wavelet function has a direct implication on the ability of the wavelet to account for the behavior of the signal. That is, if a signal has a polynomial structure or if it can be approximated by a polynomial of order q, then the wavelet transform will be able to properly capture this polynomial structure only if it has q vanishing moments. For instance, Daubechies wavelets have a number of vanishing moments which is half the length of the filter. Thus, the Haar and D(8) have, respectively, 1 and 4 vanishing moments.

The last two properties depend not only on the wavelet filter but also on its length. In fact, the most crucial point is probably not to choose the "right" filter but to choose a filter with an appropriate length. Increasing filter length allows better fitting the data. Unfortunately, this also renders the influence of boundary conditions more severe. Hence, a trade-off has to be found.

Problems due to boundary conditions arise in two situations. The first case concerns the DWT. To use the DWT, one requires a time series with a dyadic length. If the series does not meet this requirement, i.e., if its length N is such that $2^j < N < 2^{j+1}$, one has the choice between removing observations until $N = 2^j$ and completing the series such that $N = 2^{j+1}$. Removing data might be the best alternative but it leads to a loss of information.

The second case concerns both the DWT and the MODWT. The wavelet filter has to be applied on all observations, including observations recorded at the beginning ($t = 1$). A problem arises because the convolution operator requires that $L - 1$ observations are available before t. In this case, removing data is useless. Therefore, one has to complete the data series. One solution is to pad each end of the series with zeros; this technique is known as "zero padding." An alternative is to use the fit from a polynomial model to replace nonexisting data at each end of the series ("polynomial approximation"). One may also complete each end of the series either by mirroring the last observations ("mirror" or "reflection") or by taking the values observed at the beginning of the other end of the series ("circular"). The choice depends on the data considered. For instance, if working on stock returns, the use of a "mirror" seems to be the most suitable approach as it accounts for the presence of volatility clustering. Moreover, after the multiresolution decomposition of the original signal, one may obviously discard the coefficients that are affected due to their proximity to the boundaries of the series.

Instead of selecting a priori specific wavelet function, one may also use the so-called optimal transforms. The idea is to choose the wavelet function that minimizes a loss function (e.g., the entropy cost function; see Crowley 2007).

19.3.2.2 Examples of Wavelet Filters and Their Gain Functions

Figure 19.11 shows the coefficients of the Haar, D(4), D(8), and LA(8) wavelets for level $j = 4$. One may observe the very simple structure of the Haar wavelet. When comparing the latter with the D(4) and D(8) filters, it becomes clear that the longer the filter, the smoother it is. The LA(8) looks less asymmetric than the D(8). Nevertheless it is still far from being symmetric.

Studying the frequency response of these filters permits to assess their ability to capture the different frequency components. Figure 19.12 displays the gain function for each wavelet filter at scales 1–4. It is evident from the figure that the longer filters (D(8) and LA(8)) have better frequency localization. The gain functions of the D(8) and LA(8) are similar. In order to make this statement clearer, we contrast the gain functions of the Haar, D(4) and D(8) wavelets at scale $j = 5$ in Fig. 19.13. The D(8) captures much better the components corresponding to frequencies between $1/2^j$ and $1/2^{j+1}$.

19.3.2.3 Example

Let's consider a variable x, whose dynamics is primarily driven by an AR(1) process and three cyclical components:

$$y(t) = 0.90y(t-1) + \sum\nolimits_{s=3}^{5} 5\cos\left(\frac{2\pi t}{s}\right) + \varepsilon(t), \tag{19.36}$$

$\varepsilon(t)$ is an i.i.d. Gaussian process with mean zero and unit variance and $t = 1, \ldots, 10{,}000$. In the absence of seasonalities and noise, the autocorrelation function of y should take value 0.90^k at lag k. Wavelets can be used to remove the impact of both noise and seasonalities. From Eq. 19.36, one may notice that the cyclical components have a period length of 3–5 periods. Hence, they have an impact on frequencies between 1/5 and 1/3.

At scale j, the wavelet detail D_j captures frequencies $1/2^{j+1} \leq f \leq 1/2^j$ and the wavelet smooth S_j captures frequencies $f < 1/2^{j+1}$. If we use a level 2 multiresolution analysis, the wavelet smooth S_2 will thus capture the components of the time series, which have a frequency $f < 1/8$. This means that S_2 will take into account changes in y that are associated with a period length of at least 8 units of time. Therefore, S_2 should keep the AR(1) dynamics of y, while removing its cyclical behavior and noise.

Figure 19.14 reports the autocorrelation coefficients for the original time series y, for the theoretical AR(1) process, and for the wavelet smooth S_2. In order to assess the impact of choosing a different wavelet filter, we use both the Haar and the LA (8) wavelets. The results are very similar even if for short lags the LA(8) seems to provide some improvements over the Haar. All in one, this example demonstrates the ability of a wavelet filter to deal with a complex cyclical structure and with noise.

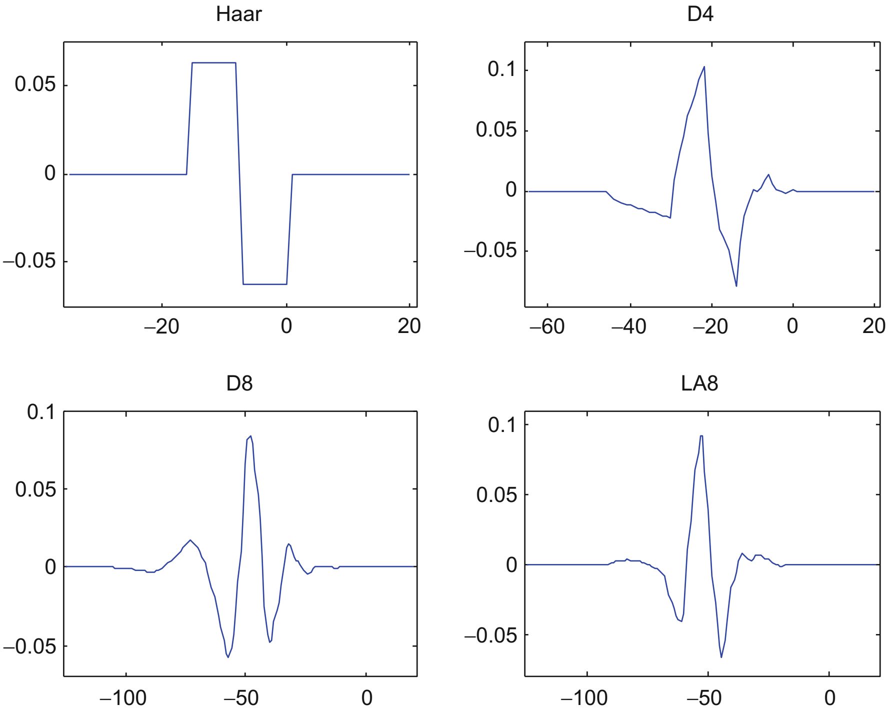


Fig. 19.11 Mother wavelet filters. The Haar, D(4), D(8), and LA(8) filters at scale $j = 4$

19.3.3 Illustration: Home Price Indices for Different US Cities

19.3.3.1 Descriptive Statistics

In this section, we study the evolution of home prices in 12 US cities from January 1987 to February 2012 (302 months).[14] We employ the Case-Shiller home price indices. The data has been gathered from the Standard & Poors website. $P_j(t)$ denotes the price index level for the jth city in period t which has been standardized so that it has a value of 100 in the first month of the sample, i.e., $P_j(t = 0) = (\forall_j)$. We compute the index returns as

$$r_j(t) = \log\left[P_j(t)\right] - \log\left[P_j(t-1)\right].$$

Table 19.1 reports some descriptive statistics for the returns on each of the 12 home price indices. There are huge disparities in performance. The largest price increases are to be found in Portland (+216 % since 1987 but −31 % since the

[14] Los Angeles (LA), San Francisco (SF), Denver (De), Washington (Wa), Miami (Mi), Chicago (Chi), Boston (Bos), Las Vegas (LV), New York (NY), Portland (Po), Charlotte (Cha), and Cleveland (Cl).

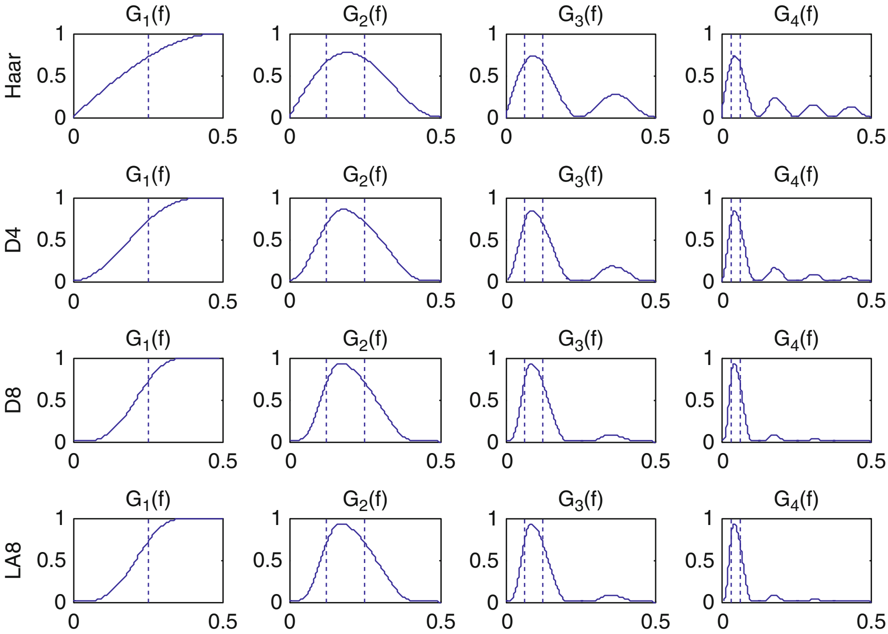


Fig. 19.12 Gain function for the Haar, D(4), D(8), and LA(8) wavelets

beginning of the subprime/financial crisis), Washington (+174 % and −34 % respectively), and Los Angeles (+169 % and −42 %). Las Vegas (+35 % and −62 %), Charlotte (+71 % and −20 %), and Cleveland (+76 % and −24 %) are the worst performers. The volatility of home price changes has been much larger in cities like San Francisco, Los Angeles, Las Vegas, and Miami than in northeast cities (plus Portland and Denver). This higher volatility is, to a large extent, the result from the severe price downturn in these four cities during the last 6 years of the sample. It is worth mentioning that some cities have been less affected by the home price crash than others (e.g., Denver and Boston). The skewness is always negative and the kurtosis is generally smaller than three.

19.3.3.2 Autocorrelations

Figure 19.15 shows the autocorrelations (up to 24 lags, i.e., 2 years) of the index returns. The full lines are for the original series, while the dotted lines show the autocorrelations computed from the smooth coefficients obtained using the MRA from a partial MODWT. We employ a LA(8) filter with $J_p = 3$. Hence, the wavelet smooth (S_3) should capture the frequency components that are associated with a period length of at least 16 months and should therefore be free of seasonal effects. In order to deal with the boundary conditions, we use the reflection method.

Some indices are very affected by seasonal effects (Portland, Boston, Denver, Charlotte, and Cleveland) while other indices are much less affected

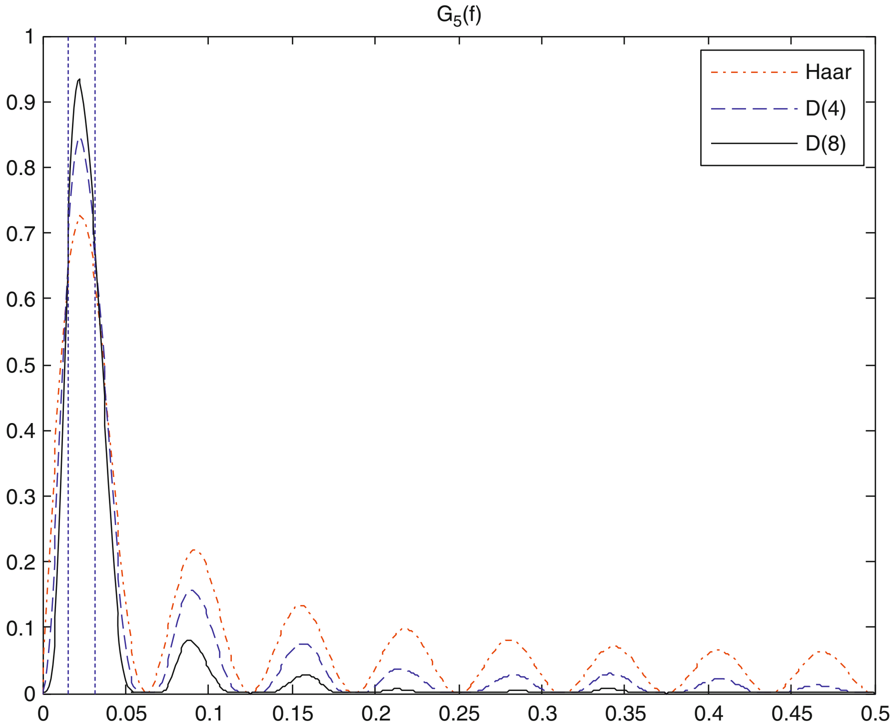


Fig. 19.13 Gain function for the Haar, D(4), and D(8) at scale $j = 5$

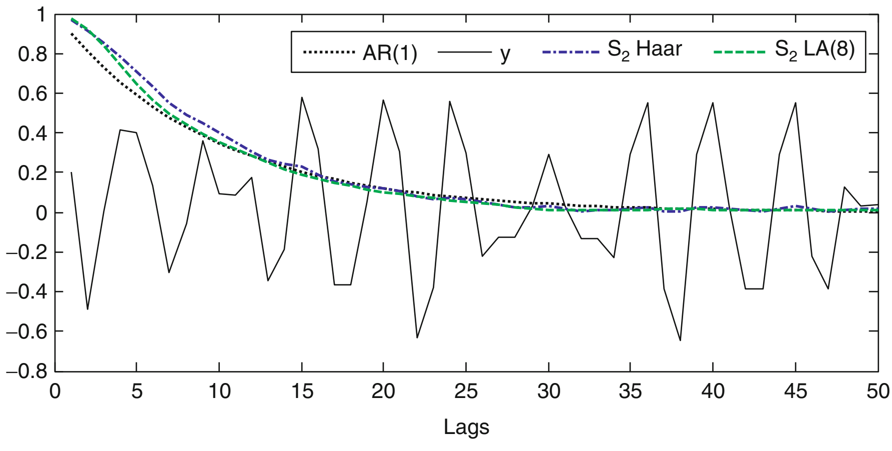


Fig. 19.14 Autocorrelations estimated before and after having removed some specific frequency components of the original time series

Table 19.1 Descriptive statistics for the 12 home price indices

	LA	SF	De	Wa	Mi	Chi	Bos	LV	NY	Po	Cha	Cl
Averag	3.9 %	3.9 %	3.5 %	4.0 %	2.8 %	2.7 %	2.9 %	1.2 %	3.0 %	4.6 %	2.1 %	2.3 %
Std.	4.4 %	4.9 %	2.6 %	3.5 %	4.0 %	3.6 %	3.2 %	4.8 %	2.8 %	3.1 %	2.1 %	3.1 %
Skew.	(0.4)	(0.7)	(0.6)	(0.2)	(1.2)	(1.1)	(0.1)	(0.3)	(0.2)	(0.8)	(0.7)	(0.8)
Kurt.	1.0	1.8	1.1	0.7	2.9	3.1	(0.0)	4.9	(0.1)	2.0	2.1	6.9
Last	159.5	124.6	121.8	175.7	139.5	105.4	146.2	89.9	159.6	129.6	108.1	94.1
High	273.9	218.4	140.3	251.1	280.9	168.6	182.5	234.8	215.8	186.5	135.9	123.5
Bottom	159.2	117.7	120.2	165.9	137.0	105.4	145.8	89.9	159.6	129.6	108.1	94.1

We report the average return, the standard deviation (both annualized and in %), the skewness, and the kurtosis of home price index returns. We also report the index level in the last month of the sample as well as the highest level reached before the crisis and the lowest level reached during the crisis

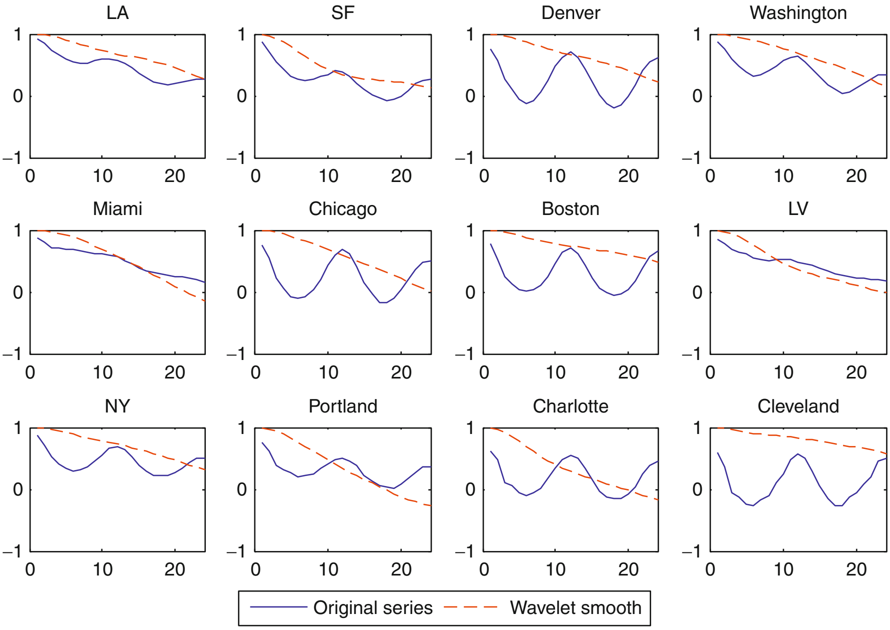


Fig. 19.15 Autocorrelations of $r_j(t)$. This figure shows the autocorrelations for both the original returns series and the wavelet smooth series for each home price index

(Miami, Las Vegas, and Los Angeles). In general the indices that display the least significant seasonal patterns are also those that have been the most affected by the recent crisis. This observation may suggest that the (quasi) absence of these patterns is the result of the predominant role of speculation on price changes in these cities.

One may again observe the ability of wavelets to remove seasonal patterns. The autocorrelations estimated from the wavelet smooth show the long-run temporal dynamics of home prices. In contrast to financial markets, whose evolution is almost unpredictable, home prices are strongly autocorrelated. The autocorrelation remains positive even after 2 years.

19.3.3.3 Variance and Cross-Correlations

Next, we analyze the variance of each index returns series and the correlations between the different indices at a variety of frequencies. We employ a partial MODWT with $J_p = 5$. Hence, at the largest scale, the wavelet coefficients contain information regarding price changes over horizon of 32–64 months. Similarly, the scaling coefficients capture the evolution for periods longer than 64 months. Figure 19.16 shows the distribution of variance across the different frequencies.

From Fig. 19.16, one may notice that most of the variance is due to frequency components associated with scales 3 and 6. On the other hand, variance at scales 1 and 2 is low and may be due to noise. Scales 4–5 are not related to important frequency components, neither to seasonal patterns (scale $j = 3$ nor to business

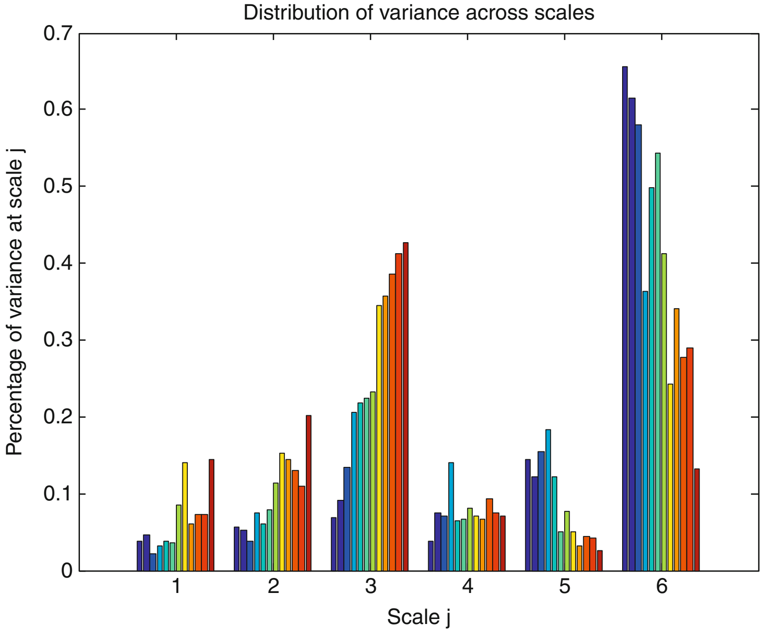


Fig. 19.16 Distribution of the wavelet variance across scales. Each bar corresponds to the variance recorded at a specific scale *j* and for a particular city. Cities are ordered as follows (from *left* to *right*): Miami, Las Vegas, Los Angeles, Washington, New York, San Francisco, Chicago, Portland, Boston, Denver, Charlotte, Cleveland

cycle components (scales $j > 5$). Thus, it comes as no surprise that they do not have much information content.

Scale 3 corresponds to periods of 8–16 months, and as such the variance observed at this scale reflects the importance of seasonal patterns in the dynamics of the various home price indices. As before, we observe that seasonal effects have a very limited impact on prices in Miami; they are also much less important in cities like Las Vegas, Los Angeles, Washington, and New York than in Boston, Denver, Charlotte, and Cleveland. Interestingly the ordering is reversed when considering the variance at scale 6. That is, the index returns series on which seasonalities have a weak impact exhibit the largest percentage of long-term volatility.

Figure 19.17 reports the correlation between the various city indices at different scales. The largest correlations are observed at scales larger than 2. In particular at scale 3, the correlations are very significant. This is because seasonalities affect most indices simultaneously. One may notice that indices 5 and 8 (Miami and Las Vegas) show less correlation with the other indices. At scales 4 and 5, the correlations become highly significant. This demonstrates that the series tend to behave very similarly in the long run.

An extension to this correlation analysis is to study how many (economic) factors are important to explain the structure of the correlation matrix at each

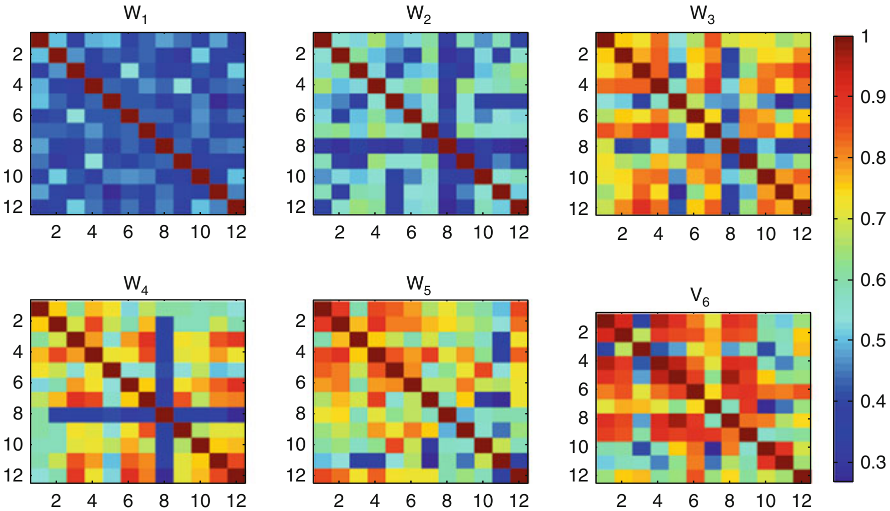


Fig. 19.17 Correlations between home price indices at different frequencies. Panels denoted by W_1 to W_5 show the correlations between the wavelet coefficients at scales $j = 1, \ldots, 5$, while panel denoted V_5 show the correlations between the scaling coefficients at scale $j = 5$. The 12 cities are order as follows: Los Angeles, San Francisco, Denver, Washington, Miami, Chicago, Boston, Las Vegas, New York, Portland, Charlotte, Cleveland

scale. In order to address this question, one may resort to random matrix theory (RMT). A good introduction is provided by Bouchaud and Potters (2004) (See also Sharifi et al. (2004) for a literature review). Here, we concentrate on the main premises of RMT. That is, this approach should help to (i) assess if the correlation coefficients have a genuine information content or if they are merely due to the noise inherent in the data and (ii) estimate the number of factors that are necessary to "explain" the correlation matrix. This is done by comparing the eigenvalues of the empirical correlation matrix with those from a theoretical distribution.

Let's consider a dataset with N time series of length T. We assume that the theoretical random matrix belongs to the ensemble of Wishart matrices. On this basis, we can derive the theoretical distribution of the eigenvalues of the correlation matrix. Under the null of pure randomness, the eigenvalues must be confined within the bounds[15]:

$$\lambda_{\min} = \left(1 + \frac{1}{q} - 2\sqrt{\frac{1}{q}}\right) \quad \text{and} \quad \lambda_{\max} = \left(1 + \frac{1}{q} + 2\sqrt{\frac{1}{q}}\right), \tag{19.37}$$

where $q = \frac{T}{N}$. If the correlation matrix is random, then the probability that any of its eigenvalues lies outside the bounds defined by $[\lambda_{\min}, \lambda_{\max}]$ is zero. Hence, the

[15]See Sharifi et al. (2004) and Kwapien et al. (2007).

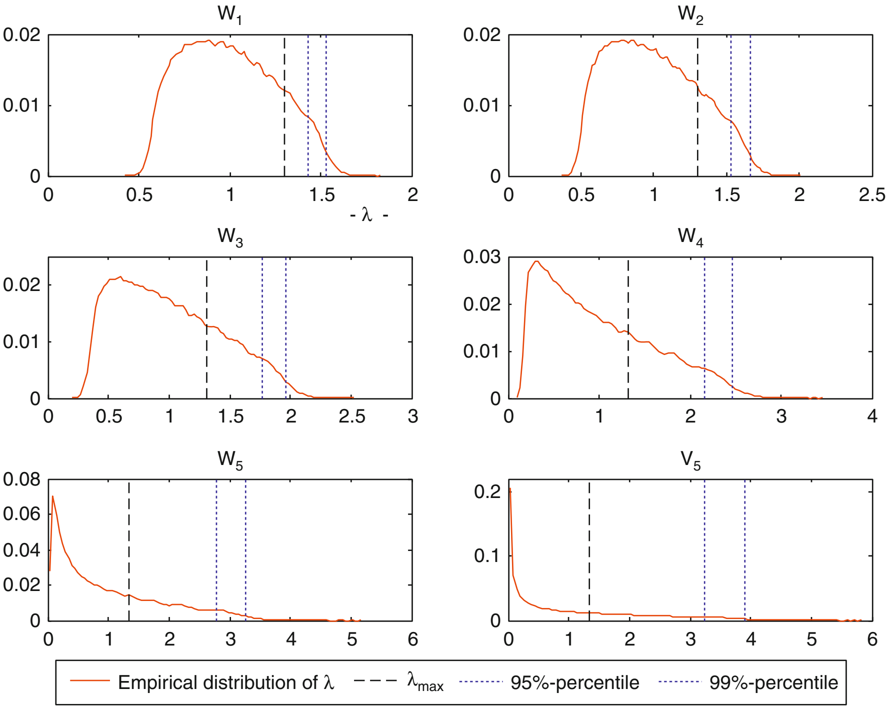


Fig. 19.18 Distribution of the empirical eigenvalues of the correlation matrix

presence of eigenvalues larger than the upper bound can be taken as an evidence that the structure of the correlation matrix is not due to chance, i.e., that there are deviations from RMT. Furthermore, the number of such eigenvalues can be interpreted as corresponding to the number of factors underlying the evolution of the N time series. A potential problem with the RMT approach is that it requires $N \to \infty$ and $T \to \infty$. In our case, these conditions are evidently not fulfilled as $T = 302$ and $N = 12$. To account for this, we also estimate λ_{max} from the empirical distribution of the correlation matrix eigenvalues. To this aim, we resample (without replacement) the original time series of returns and then apply the MODWT on the resampled data and calculate the correlation matrix and its eigenvalues at each scale of the MODWT. This procedure is done 100,000 times. In Fig. 19.18, we report λ_{max} as computed on the basis of Eq. 19.37 and the empirical values of λ_{max} obtained from our simulations. In the latter case, we consider two different values for λ_{max}, which correspond to the 5 %- and, respectively, 1 %-percentile of the empirical cumulative distribution of λ. The theoretical (RMT) maximum eigenvalue is always much lower than its empirical counterparts. This is probably due to the small sample size.

In Table 19.2, we report the eigenvalues of the correlation matrix for each scale of the MODWT. Comparing the eigenvalues λ_k, $k = 1,\ldots,12$ with $\hat{\lambda}_{max}$ demonstrates that there is a single factor underlying the evolution of home

Table 19.2 Eigenvalues of the correlation matrix at each scale of the MODWT

	w_1	w_2	w_3	w_4	w_5	v_5
λ_1	0.54	0.29	0.04	0.04	0.01	0.00
λ_2	0.60	0.30	0.07	0.07	0.02	0.00
λ_3	0.69	0.33	0.08	0.13	0.08	0.00
λ_4	0.78	0.46	0.12	0.15	0.13	0.01
λ_5	0.84	0.48	0.14	0.20	0.16	0.01
λ_6	0.86	0.61	0.18	0.23	0.18	0.05
λ_7	0.99	0.70	0.22	0.25	0.34	0.15
λ_8	1.02	0.77	0.29	0.44	0.50	0.16
λ_9	1.16	0.92	0.38	0.62	0.66	0.37
λ_{10}	1.38	1.06	0.77	0.82	0.69	1.10
λ_{11}	1.49	1.37	1.11	1.33	1.38	1.18
λ_{12}	1.64[a]	4.70[a]	8.60[a]	7.72[a]	7.84[a]	8.98[a]
λ_{max}	1.30	1.31	1.31	1.32	1.34	1.34
$\hat{\lambda}_{max}$ (95 %)	1.43	1.53	1.77	2.16	2.79	3.25
$\hat{\lambda}_{max}$ (99 %)	1.53	1.66	1.96	2.47	3.25	3.89

We report the 12 eigenvalues of the correlation matrix at each scale j as well as the theoretical (RMT) maximum eigenvalue (λ_{max} in the table) and the empirical 95 %- and 99 %-percentiles of the empirical (bootstrapped) cumulative distribution of eigenvalues at scale j ($\hat{\lambda}_{max}(95\%)$ and $\hat{\lambda}_{max}(99\%)$). [a] denotes significance at the 95 %-level

prices. At each scale, the eigenvalue that is attached to this factor is significant at the 99 % level. This unique factor can be interpreted as a sort of national-wide home price index.

19.4 Conclusion

This chapter discusses spectral and wavelet methods. It aims at being an easy-to-follow introduction and it is structured around conceptual and practical explanations. It also offers many supporting examples and illustrations. In particular, the last section provides a detailed case study, which analyzes the evolution of home prices in the USA over the last 20 years using wavelet methodology.

References

Bouchaud, J.-P., & Potters, M. (2004). *Theory of financial risk and derivative pricing: From statistical physics to risk management*. Cambridge: Cambridge University Press.

Childers, D. (1978). *Modern spectrum analysis*. New York: IEEE Press.

Crowley, P. (2007). A guide to wavelets for economists. *Journal of Economics Surveys, 21*, 207–264.

Curci, G., & Corsi, F. (2012). Discrete sine transform for multi-scale realized volatility measures. *Quantitative Finance, 12*, 263–279.

Elder, J., & Jin, H. (2007). Long memory in commodity futures volatility: A wavelet perspective. *The Journal of Futures Markets, 27*, 411–437.

Fan, J., Fan, Y., & Yazhen, W. (2007). Multiscale jump and volatility analysis for high-frequency financial data. *Journal of American Statistical Association, 102*, 618–631.

Fernandez, V., & Lucey, B. (2007). Portfolio management under sudden changes in volatility and heterogeneous investment horizons. *Physica A: Statistical Mechanics and its Applications, 375*, 612–624.

Gençay, R., & Fan, Y. (2010). Unit root tests with wavelets. *Econometric Theory, 26*, 1305–1331.

Gençay, R., & Gradojevic, N. (2011). Errors-in-variables estimation with wavelets. *Journal of Statistical Computation and Simulation, 81*, 1545–1564.

Gençay, R., Selçuk, F., & Whitcher, B. (2002). *An introduction to wavelets and other filtering methods in finance and economics*. San Diego: Academic Press.

Gençay, R., Selçuk, F., & Whitcher, B. (2003). Systematic risk and timescales. *Quantitative Finance, 3*, 108–116.

Gençay, R., Gradojevic, N., Selçuk, F., & Whitcher, B. (2010). Asymmetry of information flow between volatilities across time scales. *Quantitative Finance*, 10, 895–915.

Gu, K. (2002). The predictability of house prices. *Journal of Real Estate Research, 24*, 213–234.

Hafner, C. (2012). Cross-correlating wavelet coefficients with applications to high-frequency financial time series. *Journal of Applied Statistics, 39*, 1363–1379.

Hamilton, J. (1994). *Time series analysis*. Princeton.

Hodrick, R., & Prescott, E. (1997). Postwar U.S. business cycles: An empirical investigation. *Journal of Money, Credit, and Banking, 29*, 1–16.

Iacobucci, A. (2003). *Spectral analysis for economic time series*. Documents de Travail de l'OFCE.

Kuo, C. (1996). Autocorrelation and seasonality in the real estate market. *Journal of Real Estate Finance and Economics, 12*, 139–162.

Kwapien, J., Oswiecimka, P., & Drozdz, S. (2007). Empirics versus RMT in financial cross-correlations. *Acta Physica Polonica B, 38*, 4027–4039.

Lynch, P., & Zumbach, G. (2003). Market heterogeneities and the causal structure of volatility. *Quantitative Finance, 3*, 320–331.

Manchaldore, J., Palit, I., & Soloviev, O. (2010). Wavelet decomposition for intra-day volume dynamics. *Quantitative Finance, 10*, 917–930.

Müller, U., Dacorogna, M., Davé, R., Pictet, O., Olsen, R., & Ward, J. (1995). *Fractals and intrinsic time – A challenge to econometricians* (Working Papers). Olsen and Associates. Zürich

Nielsen, M., & Frederiksen, P. (2005). Finite sample comparison of parametric, semiparametric, and wavelet estimators of fractional integration. *Econometric Reviews, 24*, 405–443.

Oswiecimka, P., Kwapien, J., & Drozdz, S. (2006). Wavelet versus detrended fluctuation analysis of multifractal structures. *Physical Review E, 74*, 016103.

Percival, D., & Walden, A. (2000). *Wavelet methods for time series analysis*. Cambridge: Cambridge University Press.

Ramsey, J. B. (1999). The contribution of wavelets to the analysis of economic and financial data. *Philosophical Transactions of the Royal Society of London A, 357*, 2593–2606.

Ramsey, J. B. (2002). Wavelets in economics and finance: Past and future. *Studies in Nonlinear Dynamics & Econometrics, 3*, 1–29.

Ramsey, J. B., & Zhang, Z. (1997). The analysis of foreign exchange data using waveform dictionaries. *Journal of Empirical Finance, 4*, 341–372.

Ramsey, J. B., Zaslavsky, G., & Usikov, D. (1995). An analysis of U. S. stock price behavior using wavelets. *Fractals, 3*, 377–389.

Rua, A., & Nunes, L. (2009). International comovement of stock market returns: A wavelet analysis. *Journal of Empirical Finance, 16*, 632–639.

Schleicher, C. (2002). *An introduction to wavelets for economists* (Working Paper). Bank of Canada.

Sharifi, S., Crane, M., Shamaie, A., & Ruskin, H. (2004). Random matrix theory for portfolio optimization: A stability approach. *Physica A, 335*, 629–643.

Struzik, Z. (2001). Wavelet methods in financial time-series processing. *Physica A, 296*, 307–319.

Subbotin, A. (2008). *A multi-horizon scale for volatility*. University of Paris-1: CES Working Paper 2008.20.

Vuorenmaa, T. (2005). *A wavelet analysis of scaling laws and long-memory in stock market volatility*, Bank of Finland Research Discussion Paper.

Welch, P. (1967). The use of fast Fourier transforms for the estimation of power spectra: A method based on time averaging over short modified periodograms. *IEEE Transactions on Audio and Electroacoustics, 15*, 70–73.

Composite Goodness-of-Fit Tests for Left-Truncated Loss Samples

20

Anna Chernobai, Svetlozar T. Rachev, and Frank J. Fabozzi

Contents

A. Chernobai (✉)
Department of Finance, M.J. Whitman School of Management, Syracuse University, Syracuse, NY, USA
e-mail: annac@syr.edu

S.T. Rachev
Department of Applied Mathematics and Statistics, College of Business, Stony Brook University, SUNY, Stony Brook, NY, USA

FinAnalytica, Inc, New York, NY, USA
e-mail: rachev@pstat.ucsb.edu

F.J. Fabozzi
EDHEC Business School, EDHEC Risk Institute, Nice, France
e-mail: fabozzi321@aol.com; frank.fabozzi@edhec.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_20,

Abstract

In many financial models, such as those addressing value at risk and ruin probabilities, the accuracy of the fitted loss distribution in the upper tail of the loss data is crucial. In such situations, it is important to test the fitted loss distribution for the goodness of fit in the upper quantiles, while giving lesser importance to the fit in the low quantiles and the center of the distribution of the data. Additionally, in many loss models the recorded data are left truncated with the number of missing data unknown. We address this gap in literature by proposing appropriate goodness-of-fit tests.

We derive the exact formulae for several goodness-of-fit statistics that should be applied to loss models with left-truncated data where the fit of a distribution in the right tail of the distribution is of central importance. We apply the proposed tests to real financial losses, using a variety of distributions fitted to operational loss and the natural catastrophe insurance claims data, which are subject to the recording thresholds of $1 and $25 million, respectively.

Keywords

Goodness-of-fit tests • Left-truncated data • Minimum recording threshold • Loss distribution • Heavy-tailed data • Operational risk • Insurance • Ruin probability • Value at risk • Anderson-Darling statistic • Kolmogorov-Smirnov statistic

20.1 Introduction

In most loss models, the central attention is devoted to studying the distributional properties of the loss data. The shape of the dispersion of the data determines the vital statistics such as the expected loss, variance, and ruin probability, value at risk, or conditional value at risk where the shape in the right tail is crucial. Parametric procedures for testing the goodness of fit (GOF) include the likelihood ratio test and chi-squared test. A standard semi-parametric procedure to test how well a hypothesized distribution fits the data involves applying the in-sample GOF tests that provide a comparison of the fitted distribution to the empirical distribution. These tests, referred to as empirical distribution function (EDF) tests, include the Kolmogorov-Smirnov test, Anderson-Darling test, and the Cramér-von Mises tests. Related works on the discussion of these widely used tests include Anderson and Darling (1952, 1954), D'Agostino and Stephens (1986), and Shorack and Wellner (1986).

In many applications, the data set analyzed is incomplete, in the sense that the observations are present in the loss database only if they exceed a predetermined threshold level. This problem is usually absent in risk models involving market risk and credit risk. However, it is a common problem in operational risk or insurance claims models. In operational risk, banks' internal databases are subject to a minimum recording threshold of roughly \$6,000–\$10,000, and external databases usually collect operational losses starting from \$1 million, BCBS (2003). Similarly, in non-life insurance models, the thresholds are set at \$5 million, \$25 million, or other levels. Consequently, in the analysis of operational losses, recorded loss data are left truncated, and, as a result, it is inappropriate to employ standard GOF tests.

GOF tests for truncated and censored data have been studied by Dufour and Maag (1978), Gastaldi (1993), and Guilbaud (1998), among others. In this paper, we derive the exact formulae for several GOF test statistics that should be applied where there exist incomplete samples with an unknown number of missing data in low quantiles and propose two new statistics to determine the goodness of fit in the upper tail that can be used for loss models where the accuracy of the upper tail estimate is of central concern.

The paper is organized as follows. In Sect. 20.2 we describe the problem of left-truncated samples and explain the necessary adjustments that are required to the GOF tests to make them applicable for the truncated samples. In Sect. 20.3 we review the widely used test statistics for complete samples and derive the exact formulae for the statistics to be used for left-truncated samples. We propose in Sect. 20.4 two new EDF statistics to be used for the situations when the fit in the upper tail is of the central concern. Application of the modified EDF tests to operational loss data, obtained from Zurich IC2 FIRST Database, and the USA natural catastrophe insurance claims data, obtained from Insurance Services Office Inc. Property Claim Services, is presented in Sect. 20.5, with final remarks in Sect. 20.6. Necessary derivations are provided in the Appendix.

20.2 Problem Setup

Suppose we have a left-truncated sample, with the data below a prespecified threshold level H not recorded (not observable). The observable data sample $\mathbf{x} = \{x_1, x_2, \ldots, x_n\}$ has each data point at least as great as H and includes a total of n observations. Let $\{X_{(j)}\}_{1 \le j \le n}$ be a vector of the order statistics, such that $X_{(1)} \le X_{(2)} \le \ldots \le X_{(n)}$. The empirical distribution function of the sample is defined as

$$F_n(x) := \frac{\#\,observations \le x_{(j)}}{n} = \begin{cases} 0 & x < x_{(1)} \\ \dfrac{j}{n} & x_{(j)} \le x < x_{(j+1)}, \quad j = 1, 2, \ldots, n-1 \cdot \\ 1 & x \ge x_n, \end{cases} \tag{20.1}$$

Graphically, the empirical distribution function of an observed data sample is represented as a step function with a jump of size $1/n$ occurring at each recorded sample value. On the other hand, with left-truncated data which is a part of a larger complete data set, the true size of jumps of the EDF at each value of the *complete* data sample would be of size $1/n^c$, $n^c = n + m$ rather than $1/n$, where n^c is the total number of points of the complete data set and m is the unknown number of missing points. In the GOF tests the null hypothesis states that the observed loss sample belongs to a family of truncated distributions, with the parameter specified (simple test) or unspecified (composite test).

We fit a continuous *truncated* distribution F to the data, given that the data exceed or equal to H, and estimate the conditional parameters θ with the maximum likelihood (or an alternative method) by

$$\hat{\theta}_{\mathrm{MLE}} = \arg\max_{\theta} \log\left(\prod_{k=1}^{n} \frac{f_\theta(x_k)}{1 - F_\theta(H)}\right). \tag{20.2}$$

Assuming that F is the true distribution, the estimated number of missing points $\hat{m}$ and the number of observed points n are related as[1]

$$\frac{\hat{m}}{n} = \frac{z_{\mathrm{H}}}{(1 - z_{\mathrm{H}})} \tag{20.3}$$

resulting in

$$\hat{n}^c = n\frac{z_{\mathrm{H}}}{(1 - z_{\mathrm{H}})} + n = \frac{n}{(1 - z_{\mathrm{H}})} \tag{20.4}$$

where $z_{\mathrm{H}} := \hat{F}_\theta(H)$ is the estimated distribution evaluated at the truncation point.

Then, the estimated empirical distribution function $\hat{F}_{n^c}(x)$ of complete data sample is

$$\hat{F}_{n^c}(x) := \frac{\textit{estimated \# observations} \le x_{(j)}}{\hat{n}^c}, \tag{20.5}$$

where the numerator refers to the total number of observations of the complete data sample, not exceeding in magnitude the jth-order statistic of the incomplete (observed) data sample such that $X_{(1)} \le X_{(2)} \le \ldots \le X_{(n)}$. By Eqs. 20.3 and 20.4, Eq. 20.5 becomes

$$\hat{F}_{n^c}(x) = \frac{\hat{m} + j}{n/(1 - z_{\mathrm{H}})} = \frac{n z_{\mathrm{H}} + j(1 - z_{\mathrm{H}})}{n} = z_{\mathrm{H}} + \frac{j}{n}(1 - z_{\mathrm{H}}), \quad j = 1, 2, \ldots, n.$$

[1]More accurately, $\hat{m}$ should be estimated as $\hat{m} = \left\lceil n\frac{z_{\mathrm{H}}}{(1 - z_{\mathrm{H}})}\right\rceil$, but this detail can be ignored for all practical purposes.

Rearranging terms leads to the fitted distribution function of the observed sample of the following form:

$$\hat{F}^*_\theta(x) = \begin{cases} \dfrac{\hat{F}_\theta(x) - \hat{F}_\theta(H)}{1 - \hat{F}_\theta(H)} & x \geq H \\ 0 & x < H, \end{cases} \tag{20.6}$$

so that $F_\theta(X) \sim U[F_\theta(H), 1]$ and $F^*_\theta(X) \sim \mathcal{U}[0, 1]$ under the null hypothesis that the fitted distribution function is true. Therefore, the estimated empirical distribution function of the observed part of the data, using Eq. 20.1, is represented by

$$F_n(x)\left(1 - \hat{F}_\theta(H)\right) + \hat{F}_\theta(H) = \begin{cases} \hat{F}_\theta(H) & x < x_{(1)} \\ \dfrac{j}{n}\left(1 - \hat{F}_\theta(H)\right) + \hat{F}_\theta(H) & x_{(j)} \leq x < x_{(j+1)},\ j = 1, 2, \ldots, n-1 \cdot \\ 1 & x \geq x_n, \end{cases} \tag{20.7}$$

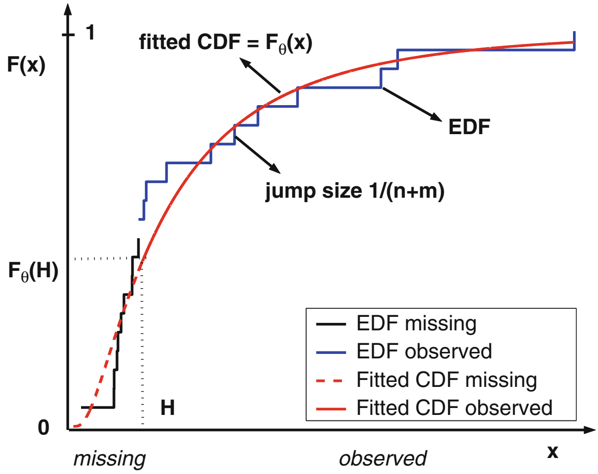
Figure 20.1 gives a visual illustration of the idea we just described. With these modifications, the in-sample GOF tests can be applied to the left-truncated samples.

In this paper we consider tests of a composite hypothesis that the empirical distribution function of an observed incomplete left-truncated loss data sample belongs to a family of hypothesized distributions (with parameters not specified), i.e.,

$$\mathbf{H_0} : F_n(x) \in F^*_\theta(x) \qquad \text{vs.} \qquad \mathbf{H_A} : F_n(x) \notin F^*_\theta(x). \tag{20.8}$$

Under the null Eq. 20.8, $F^*_\theta(X) \sim \mathcal{U}[0, 1]$, and the null is rejected if the p-value is lower than the level α, such as α from 5 % to 10 %. Letting D be the observed value of a GOF statistic (such as Kolmogorov-Smirnov or Anderson-Darling) and d the critical value for a given level α, the p-value is computed as p-value $= P(D \geq d)$. Since the distribution of the statistic is not parameter-free, one way to compute the p-values and the critical values is by means of Monte Carlo simulation, for each hypothesized fitted distribution (Ross 2001). Under the procedure, the observed value D is computed. Then, for a given level α, the following algorithm is applied:

1. Generate a large number of samples I (such as I = 1,000) from the fitted truncated distribution of size n (the number of observed data points), such that these random variates are above or equal to H.
2. Fit a *truncated* distribution and estimate conditional parameters $\hat{\theta}$ for each sample $i = 1, 2, \ldots I$.
3. Estimate the GOF statistic value D_i for each sample $i = 1, 2, \ldots I$.
4. Calculate p-value as the proportion of times the sample statistic values exceed the observed value d of the original sample.
5. Reject $\mathbf{H}_0$ if the p-value is smaller than α.



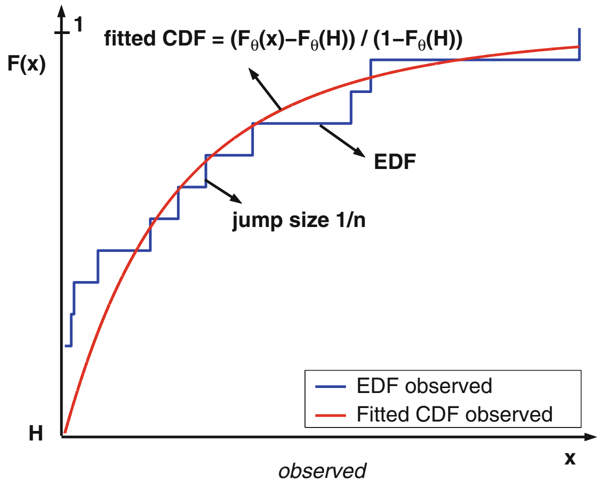


Fig. 20.1 Illustration of the empirical distribution function for data with missing observations below the threshold H and the fitted cumulative distribution function

A p-value of, for example, 0.3, would mean that in 30 % of samples of the same size simulated from the same distribution with the same parameter estimation procedure applied, the test statistic value was higher than the one observed in the original sample.

20.3 EDF Statistics for Left-Truncated Loss Samples

The EDF statistics are based on the vertical differences between the empirical and fitted (truncated) distribution function. They are divided into two classes: (1) the supremum class (such as Kolmogorov-Smirnov and Kuiper statistics) and

(2) the quadratic class (such as Anderson-Darling and Cramér-von Mises statistics). In this section, we derive the exact computing formulae for a number of EDF test statistics, modified so that they can be applied to left-truncated loss samples. For the left-truncated samples, $F^*_\theta(X)$ denotes the null distribution function for left-truncated sample values. The variable $F^*_\theta(X)$ is distributed uniformly over the [0, 1] interval. The variable $F_\theta(X)$ is distributed uniformly over the $[F_\theta(H), 1]$ interval. We reserve some other notations: $z_{\mathrm{H}} := \hat{F}_\theta(H)$ that was defined earlier and $z_j := \hat{F}_\theta\left(x_{(j)}\right)$ for truncated samples. In this section we discuss the Kolmogorov-Smirnov, Kuiper, Anderson-Darling, and the Cramér-von Mises statistics and use an asterisk (*) to denote their left-truncated sample analog.

20.3.1 Supremum Class Statistics

20.3.1.1 Kolmogorov-Smirnov Statistic

A widely used supremum class statistic, the Kolmogorov-Smirnov (*KS*) statistic, measures the absolute value of the maximum distance between the empirical and fitted distribution function and puts equal weight on each observation. Let $\{X_{(j)}\}_{1<j<n}$ be the vector of the order statistics and $X_{(1)} < X_{(2)} < \ldots < X_{(n)}$, such that strict inequalities hold. Usually, such distance is the greatest around the median of the sample. For the left-truncated data samples, the *KS* statistic is expressed as

$$KS^* = \sqrt{n}\sup_x \left|F_n(x) - \hat{F}^*_\theta(x)\right|. \tag{20.9}$$

The *KS** statistic can be computed from

$$KS^{+*} = \sqrt{n}\sup_j\left\{F_n\left(x_{(j)}\right) - \hat{F}^*_\theta\left(x_{(j)}\right)\right\} = \frac{\sqrt{n}}{1-z_{\mathrm{H}}}\sup_j\left\{z_{\mathrm{H}} + \frac{j}{n}(1-z_{\mathrm{H}}) - z_j\right\},$$
$$KS^{-*} = \sqrt{n}\sup_j\left\{\hat{F}^*_\theta\left(x_{(j)}\right) - F_n\left(x_{(j)}\right)\right\} = \frac{\sqrt{n}}{1-z_{\mathrm{H}}}\sup_j\left\{z_j - \left(z_{\mathrm{H}} + \frac{j-1}{n}(1-z_{\mathrm{H}})\right)\right\},$$

and becomes

$$KS^* = \max\{KS^{+*}, KS^{-*}\}. \tag{20.10}$$

20.3.1.2 Kuiper Statistic

The *KS* statistic gives no indication of whether the maximum discrepancy between $F_n(x)$ and $\hat{F}^*_\theta(x)$ occurs when $F_n(x)$ is above $\hat{F}^*_\theta(x)$ or when $\hat{F}^*_\theta(x)$ is above $F_n(x)$. The Kuiper statistic (V) is closely related to the *KS* statistic. It measures the total sum of the absolute values of the two largest vertical deviations of the fitted distribution function from $F_n(x)$, when $F_n(x)$ is above $\hat{F}^*_\theta(x)$ and when $\hat{F}^*_\theta(x)$ is above $F_n(x)$. For left-truncated data samples, it is computed as

$$V^* = KS^{+*} + KS^{-*}, \tag{20.11}$$

with KS^{+*} and KS^{-*} defined in Sect. 20.1.1.

20.3.1.3 Anderson-Darling Statistic

There are two variations of the Anderson-Darling (*AD*) statistic. Its simplest, supremum class version is a variance-weighted *KS* statistic with a weight of $\psi\left(\hat{F}^*_\theta(x)\right) = \left(\hat{F}^*_\theta(x)\left(1-\hat{F}^*_\theta(x)\right)\right)^{-1/2}$ attached to each observation in Eq. 20.9. (The second version will be discussed in the next section.) Under this specification, the observations in the lower and upper tails of the truncated sample are assigned a higher weight. Let $\{X_{(j)}\}_{1\le j\le n}$ be the vector of the order statistics, such that $X_{(1)} \le X_{(2)} \le \ldots \le X_{(n)}$. Then the *AD* statistic is defined for left-truncated samples as

$$AD^* = \sqrt{n}\sup_x \left| \frac{F_n(x) - \hat{F}^*_\theta(\mathtt{x})}{\sqrt{\hat{F}^*_\theta(x)\left(1-\hat{F}^*_\theta(x)\right)}} \right|. \tag{20.12}$$

For left-truncated samples, the computing formula is derived from

$$AD^{+*} = \sqrt{n}\sup_j \left\{ \frac{F_n\left(x_{(j)}\right) - \hat{F}^*_\theta\left(x_{(j)}\right)}{\sqrt{\hat{F}^*_\theta\left(x_{(j)}\right)\left(1-\hat{F}^*_\theta\left(x_{(j)}\right)\right)}} \right\} = \sqrt{n}\sup_j \left\{ \frac{z_H + \frac{j}{n}(1-z_{\mathrm{H}}) - z_j}{\sqrt{\left(z_j - z_{\mathrm{H}}\right)\left(1-z_j\right)}} \right\},$$

$$AD^{-*} = \sqrt{n}\sup_j \left\{ \frac{\hat{F}^*_\theta\left(x_{(j)}\right) - F_n\left(x_{(j)}\right)}{\sqrt{\hat{F}^*_\theta\left(x_{(j)}\right)\left(1-\hat{F}^*_\theta\left(x_{(j)}\right)\right)},} \right\} = \sqrt{n}\sup_j \left\{ \frac{z_j - z_{\mathrm{H}} - \frac{j-1}{n}(1-z_{\mathrm{H}})}{\sqrt{\left(z_j - z_{\mathrm{H}}\right)\left(1-z_j\right)}} \right\}$$

and becomes

$$AD^* = \max\{AD^{+*},\ AD^{-*}\}. \tag{20.13}$$

20.3.2 Quadratic Class Statistics

The quadratic statistics for complete data samples are grouped under the Cramér-von Mises family as

$$Q = n\int_{-\infty}^{\infty} \left(F_n(x) - \hat{F}_\theta(x)\right)^2 \psi\left(\hat{F}_\theta(x)\right) d\hat{F}_\theta(x), \tag{20.14}$$

in which the weight function $\psi\left(\hat{F}_\theta(x)\right)$ is assigned to give a certain weight to different observations, depending on the purpose. For left-truncated samples, we denote the Cramér-von Mises family as $Q*$ and $\hat{F}_\theta(x)$ is replaced by $\hat{F}^*_\theta(x)$:

$$Q^* = n\int_{H}^{\infty} \left(F_n(x) - \hat{F}^*_\theta(x)\right)^2 \psi\left(\hat{F}^*_\theta(x)\right) d\hat{F}^*_\theta(x). \tag{20.15}$$

Depending on the form of the weighting function, the sample observations are given a different weight. $\psi\left(\hat{F}^*_\theta(x)\right) = \left(\hat{F}^*_\theta(x)\left(1-\hat{F}^*_\theta(x)\right)\right)^{-1}$ yields the quadratic Anderson-Darling statistic and $\psi\left(\hat{F}^*_\theta(x)\right) = 1$ yields the Cramér-von Mises statistic.

Derivation of the computing formulae makes use of Eq. 20.1 and involves the Probability Integral Transformation (PIT) technique. For left-truncated samples, this leads to

$$Q^* = n\int_0^1 (F_n(z^*) - z^*)^2 \psi(z^*)dz^* = \frac{n}{1-z_{\mathrm{H}}}\int_{z_{\mathrm{H}}}^1 \left(F_n(z^*) - \frac{z-z_{\mathrm{H}}}{1-z_{\mathrm{H}}}\right)^2 \psi\left(\frac{z-z_{\mathrm{H}}}{1-z_{\mathrm{H}}}\right)dz, \tag{20.16}$$

where $Z^* = \hat{F}^*_\theta(X) = \frac{Z-z_{\mathrm{H}}}{1-z_{\mathrm{H}}}$ with $F^*_\theta(X) \sim \mathcal{U}[0,1]$ under the null, and so $Z = \hat{F}_\theta(X)$ with $F_\theta(X) \sim U[F_\theta(H), 1]$ under the null. $F_n(Z^*) = F_n(X) = F^*_\theta(F_n(X))$ is the empirical distribution function of Z*. $z_{\mathrm{H}} = \hat{F}_\theta(H) = F_\theta\left(\hat{F}_\theta(H)\right)$.

20.3.2.1 Anderson-Darling Statistic

The supremum version of the *AD* statistic was described in Sect. 20.1.3. Another, more generally used, version of this statistic belongs to the quadratic class defined by the Cramér-von Mises family (Eq. 20.15) with the weight function for left-truncated samples of $\psi\left(\hat{F}^*_\theta(x)\right) = \left(\hat{F}^*_\theta(x)\left(1-\hat{F}^*_\theta(x)\right)\right)^{-1}$. Again, with this specification, most weight is being put on the outer left and right quantiles of the distribution, proportional to the appropriate tails. Let $\{X_{(j)}\}_{1\le j\le n}$ be the vector of the order statistics, such that $X_{(1)} \le X_{(2)} \le \ldots \le X_{(n)}$. The computing formula for the *AD* statistic for left-truncated samples becomes (the derivation is given in Appendix)

$$\begin{aligned} AD^{2*} = &- n + 2n\log(1-z_{\mathrm{H}}) - \frac{1}{n}\sum_{j=1}^n (1+2(n-j))\log\left(1-z_j\right) + \ldots \\ &+ \frac{1}{n}\sum_{j=1}^n (1-2j)\log\left(z_j - z_{\mathrm{H}}\right). \end{aligned} \tag{20.17}$$

20.3.3 Cramér-von Mises Statistic

Cramér-von Mises (denoted as W^2) statistic belongs to the Cramér-von Mises family (Eq. 20.15) with the weight function $\psi\left(\hat{F}^*_\theta(x)\right) = 1$. Let $\{X_{(j)}\}_{1\le j\le n}$ be the vector of the order statistics, such that $X_{(1)} \le X_{(2)} \le \ldots \le X_{(n)}$. The computing formula for the statistic for left-truncated samples becomes (the derivation is given in Appendix)

$$W^{2*} = \frac{n}{3} + \frac{nz_{\mathrm{H}}}{1-z_{\mathrm{H}}} + \frac{1}{n(1-z_{\mathrm{H}})}\sum_{j=1}^n (1-2j)z_j + \frac{1}{(1-z_{\mathrm{H}})^2}\sum_{j=1}^n \left(z_j - z_{\mathrm{H}}\right)^2. \tag{20.18}$$

20.4 "Upper Tail" Anderson-Darling Statistic

In practice, there are often cases when it is necessary to test whether a distribution fits the data well in the upper tail and the fit in the lower tail or around the median is of little or less importance. Examples include operational risk and insurance claims modelling, in which goodness of the fit in the tails determines the value at risk, conditional value at risk, and ruin probabilities. Given the Basel II Capital Accord's recommendations, under the loss distribution approach the operational risk capital charge is derived from the value-at-risk measure, which requires an accurate estimate of the upper tail of the loss distribution. Similarly, in insurance, the upper tail of the claim size distribution is central to obtaining accurate estimates of ruin probability. For this purpose, we introduce a statistic, which we refer to as the *upper tail* Anderson-Darling statistic and denote by AD_{up}. We propose two different versions of the statistic.

20.4.1 Supremum Class "Upper Tail" Anderson-Darling Statistic

The first version of AD_{up} belongs to the supremum class EDF statistics. For complete data samples, each observation of the KS statistic is assigned a weight of $\psi\left(\hat{F}_\theta(x)\right) = \left(\left(1 - \hat{F}_\theta(x)\right)\right)^{-1}$. Under this specification, the observations in the upper tail are assigned a higher weight than those in the lower tail. Let $\{X_{(j)}\}_{1 \le j \le n}$ be the vector of the order statistics, such that $X_{(1)} \le X_{(2)} \le \ldots \le X_{(n)}$. Then we define the AD_{up} statistic for complete data samples as

$$AD_{\text{up}} = \sqrt{n} \sup_x \left| \frac{F_n(x) - \hat{F}_\theta(x)}{1 - \hat{F}_\theta(x)} \right|. \tag{20.19}$$

Denoting $z_j := \hat{F}_\theta\left(x_{(j)}\right)$, the computing formula is derived from

$$AD^{+}_{\text{up}} = \sqrt{n} \sup_j \left\{ \frac{\frac{j}{n} - z_j}{1 - z_j} \right\},$$

$$AD^{-}_{\text{up}} = \sqrt{n} \sup_j \left\{ \frac{z_j - \frac{j-1}{n}}{1 - z_j} \right\},$$

and becomes

$$AD_{\text{up}} = \max\left\{ AD^{+*}_{\text{up}},\ AD^{-*}_{\text{up}} \right\}. \tag{20.20}$$

For left-truncated samples, the counterpart of the AD_{up} statistic can be similarly computed using

$$AD_{\text{up}}^{+*} = \sqrt{n}\sup_j\left\{\frac{F_n\left(x_{(j)}\right) - \hat{F}_\theta^*\left(x_{(j)}\right)}{1 - \hat{F}_\theta^*\left(x_{(j)}\right)}\right\} = \sqrt{n}\sup_j\left\{\frac{z_{\text{H}} + \frac{j}{n}(1 - z_{\text{H}}) - z_j}{1 - z_j}\right\},$$
$$AD_{\text{up}}^{-*} = \sqrt{n}\sup_j\left\{\frac{\hat{F}_\theta^*\left(x_{(j)}\right) - F_n\left(x_{(j)}\right)}{1 - \hat{F}_\theta^*\left(x_{(j)}\right)}\right\} = \sqrt{n}\sup_j\left\{\frac{z_j - z_{\text{H}} - \frac{j-1}{n}(1 - z_{\text{H}})}{1 - z_j}\right\},$$

and becomes

$$AD_{\text{up}}^* = \max\left\{AD_{\text{up}}^{+*},\ AD_{\text{up}}^{-*}\right\}. \tag{20.21}$$

20.4.2 Quadratic Class "Upper Tail" Anderson-Darling Statistic

Another way to define the *upper tail* Anderson-Darling statistic is by an integral of the Cramér-von Mises family (Eq. 20.14) with the weighting function[2] of the form $\psi\left(\hat{F}_\theta(x)\right) = \left(1 - \hat{F}_\theta(x)\right)^{-2}$ for complete samples and $\psi\left(\hat{F}_\theta^*(x)\right) = \left(1 - \hat{F}_\theta^*(x)\right)^{-2}$ for left-truncated samples. Such weighting function gives a higher weight to the upper tail and a lower weight to the lower tail. We define this statistic as AD_{up}^2.

Let $\{X_{(j)}\}_{1\le j\le n}$ be the vector of the order statistics, such that $X_{(1)} \le X_{(2)} \le \ldots \le X_{(n)}$. The AD_{up}^2 statistic's general form for complete samples can be expressed as

$$AD_{\text{up}}^2 = n\int_H^\infty \frac{\left(F_n(x) - \hat{F}_\theta(x)\right)^2}{\left(1 - \hat{F}_\theta(x)\right)^2} d\hat{F}_\theta(x). \tag{20.22}$$

For complete data samples $F_\theta(X) \sim U[0, 1]$ under the null hypothesis. If we denote $z_j := \hat{F}_\theta\left(x_{(j)}\right)$, straightforward calculations lead to the following computing formula:

$$AD_{\text{up}}^2 = 2\sum_{j=1}^n \log\left(1 - z_j\right) + \frac{1}{n}\sum_{j=1}^n (1 + 2(n-j))\frac{1}{1 - z_j}.$$

For incomplete left-truncated samples, $F_\theta^*(X)$ is distributed $U[F_\theta(H),\ 1]$ under the null. Applying the PIT technique leads to the computing formula of the

[2]It has been shown that the weighting function $\psi(t) = (1-t)^{-\beta}$ possesses nice asymptotic properties only for β=[0,2) (Deheuvels and Martynov 2003). For β=2, which is the case considered in this chapter, the asymptotic distribution of the test statistic has infinite mean. This is indeed a concern for very large samples (i.e., the asymptotic case n→∞). Yet, because firms' operational loss data samples are typically relatively small, the asymptotic distribution of the test statistic should not generate large concerns. Nevertheless, the results of the proposed quadratic class upper-tail Anderson-Darling test should be treated with caution and with consideration of the properties described above.

Table 20.1 Description of EDF statistics for complete data samples

$\mathbf{H_0}: F_n(x) \in F(x)$ vs. $\mathbf{H_A}: F_n(x) \notin F_\theta(x)$	
Notations: $z_j := \hat{F}_\theta(x_{(j)}),\ j = 1, 2, \ldots, n$	
Statistic	Description and computing formula
KS	$KS = \sqrt{n}\sup_x\left\|F_n(x) - \hat{F}_\theta(x)\right\|$
	Computing formula: $KS = \sqrt{n}\max\left\{\sup_j\left\{\frac{j}{n} - z_j\right\},\ \sup_j\left\{z_j - \frac{j-1}{n}\right\}\right\}$
V	$V = \sqrt{n}\left(\sup_x\left\{F_n(x) - \hat{F}_\theta(x)\right\} + \sup_x\left\{\hat{F}_\theta(x) - F_n(x)\right\}\right)$
	Computing formula: $V = \sqrt{n}\left(\sup_j\left\{\frac{j}{n} - z_j\right\} + \sup_j\left\{z_j - \frac{j-1}{n}\right\}\right)$
AD	$AD = \sqrt{n}\sup_x\left\|\frac{F_n(x) - \hat{F}_\theta(x)}{\sqrt{\hat{F}_\theta(x)\left(1 - \hat{F}_\theta(x)\right)}}\right\|$
	Computing formula: $AD = \sqrt{n}\max\left\{\sup_j\left\{\frac{\frac{j}{n} - z_j}{\sqrt{z_j\left(1 - z_j\right)}}\right\},\ \sup_j\left\{\frac{z_j - \frac{j-1}{n}}{\sqrt{z_j\left(1 - z_j\right)}}\right\}\right\}$
AD_{up}	$AD_{\mathrm{up}} = \sqrt{n}\sup_x\left\|\frac{F_n(x) - \hat{F}_\theta(x)}{1 - \hat{F}_\theta(x)}\right\|$
	Computing formula: $AD_{\mathrm{up}} = \sqrt{n}\max\left\{\sup_j\left\{\frac{\frac{j}{n} - z_j}{1 - z_j}\right\},\ \sup_j\left\{\frac{z_j - \frac{j-1}{n}}{1 - z_j}\right\}\right\}$
AD^2	$AD^2 = n\int_{-\infty}^{\infty}\frac{\left(F_n(x) - \hat{F}_\theta(x)\right)^2}{\hat{F}_\theta(x)\left(1 - \hat{F}_\theta(x)\right)}d\hat{F}_\theta(x)$
	Computing formula: $AD^2 = -n + \frac{1}{n}\sum_{j=1}^{n}(1 - 2j)\log z_j - \frac{1}{n}\sum_{j=1}^{n}(1 + 2(n - j))\log\left(1 - z_j\right)$
W^2	$W^2 = n\int_{-\infty}^{\infty}\left(F_n(x) - \hat{F}_\theta(x)\right)^2 d\hat{F}_\theta(x)$
	Computing formula: $W^2 = \frac{n}{3} + \frac{1}{n}\sum_{j=1}^{n}(1 - 2j)z_j + \sum_{j=1}^{n}{z_j}^2$
AD^2_{up}	$AD^2_{\mathrm{up}} = n\int_{-\infty}^{\infty}\frac{\left(F_n(x) - \hat{F}_\theta(x)\right)^2}{\left(1 - \hat{F}_\theta(x)\right)^2}d\hat{F}_\theta(x)$
	Computing formula: $AD^2_{\mathrm{up}} = \frac{1}{n}\sum_{j=1}^{n}(1 + 2(n - j))\frac{1}{\left(1 - z_j\right)} + 2\sum_{j=1}^{n}\log\left(1 - z_j\right)$

Table 20.2 Description of EDF statistics for left-truncated (threshold = H) data samples

$\mathbf{H_0}: F_n(x) \in F^*_\theta(x)$ vs. $\mathbf{H_A}: F_n(x) \notin F^*_\theta(x)$, $F^*_\theta(x) := \frac{F_\theta(x)-F_\theta(H)}{1-F_\theta(H)}$	
Notations: $z_j := \hat{F}_\theta\left(x_{(j)}\right)$, $z_\mathrm{H} = \hat{F}_\theta(H), j = 1, 2, \ldots, n$	
Statistic	Description and computing formula
*KS**	$KS^* = \sqrt{n}\sup_x\left\lvert F_n(x) - \hat{F}^*_\theta(x)\right\rvert$
	Computing formula: $KS^* = \frac{\sqrt{n}}{1-z_\mathrm{H}}\max\left\{\sup_j\left\{z_\mathrm{H} + \frac{j}{n}(1-z_\mathrm{H}) - z_j\right\}, \sup_j\left\{z_j - \left(z_\mathrm{H} + \frac{j-1}{n}(1-z_H)\right)\right\}\right\}$
*V**	$V^* = \sqrt{n}\left(\sup_x\left\{F_n(x) - \hat{F}^*_\theta(x)\right\} + \sup_x\left\{\hat{F}^*_\theta(x) - F_n(x)\right\}\right)$
	Computing formula: $V^* = \frac{\sqrt{n}}{1-z_\mathrm{H}}\left(\sup_j\left\{z_\mathrm{H} + \frac{j}{n}(1-z_\mathrm{H}) - z_j\right\} + \sup_j\left\{z_j - \left(z_\mathrm{H} + \frac{j-1}{n}(1-z_\mathrm{H})\right)\right\}\right)$
*AD**	$AD^* = \sqrt{n}\sup_x\left\lvert\frac{F_n(x) - \hat{F}^*_\theta(x)}{\sqrt{\hat{F}^*_\theta(x)\left(1-\hat{F}^*_\theta(x)\right)}}\right\rvert$
	Computing formula: $AD^* = \sqrt{n}\max\left\{\sup_j\left\{\frac{z_\mathrm{H} + \frac{j}{n}(1-z_\mathrm{H}) - z_j}{\sqrt{(z_j - z_\mathrm{H})(1-z_j)}}\right\}, \sup_j\left\{\frac{z_j - z_\mathrm{H} - \frac{j-1}{n}(1-z_\mathrm{H})}{\sqrt{(z_j - z_\mathrm{H})(1-z_j)}}\right\}\right\}$
AD^*_up	$AD^*_\mathrm{up} = \sqrt{n}\sup_x\left\lvert\frac{F_n(x) - \hat{F}^*_\theta(x)}{1-\hat{F}^*_\theta(x)}\right\rvert$
	Computing formula: $AD^*_\mathrm{up} = \sqrt{n}\max\left\{\sup_j\left\{\frac{z_\mathrm{H} + \frac{j}{n}(1-z_\mathrm{H}) - z_j}{1-z_j}\right\}, \sup_j\left\{\frac{z_j - z_\mathrm{H} - \frac{j-1}{n}(1-z_\mathrm{H})}{1-z_j}\right\}\right\}$
AD^{2*}	$AD^{2*} = n\int_H^\infty \frac{\left(F_n(x) - \hat{F}^*_\theta(x)\right)^2}{\hat{F}^*_\theta(x)\left(1-\hat{F}^*_\theta(x)\right)}d\hat{F}^*_\theta(x)$
	Computing formula: $AD^{2*} = -n + 2n\log(1-z_\mathrm{H}) - \frac{1}{n}\sum_{j=1}^n (1+2(n-j))\log(1-z_j) + \frac{1}{n}\sum_{j=1}^n (1-2j)\log(z_j - z_\mathrm{H})$
W^{2*}	$W^{2*} = n\int_H^\infty \left(F_n(x) - \hat{F}^*_\theta(x)\right)^2 d\hat{F}^*_\theta(x)$
	Computing formula: $W^{2*} = \frac{n}{3} + \frac{n\ z_\mathrm{H}}{1-z_\mathrm{H}} + \frac{1}{n(1-z_\mathrm{H})}\sum_{j=1}^n (1-2j)z_j + \frac{1}{(1-z_\mathrm{H})^2}\sum_{j=1}^n (z_j - z_\mathrm{H})^2$

(*continued*)

Table 20.2 (continued)

$\mathbf{H_0}: F_n(x) \in F_\theta^*(x)$ vs. $\mathbf{H_A}: F_n(x) \notin F_\theta^*(x)$, $F_\theta^*(x) := \frac{F_\theta(x)-F_\theta(H)}{1-F_\theta(H)}$	
Notations: $z_j := \hat{F}_\theta\left(x_{(j)}\right)$, $z_{\mathrm{H}} = \hat{F}_\theta(H), j = 1, 2, \ldots, n$	
Statistic	Description and computing formula
$AD_{\mathrm{up}}^{2\,*}$	$AD_{\mathrm{up}}^{2*} = n\int_H^\infty \frac{\left(F_n(x) - \hat{F}_\theta^*(x)\right)^2}{\left(1-\hat{F}_\theta^*(x)\right)^2} d\hat{F}_\theta^*(x)$
	Computing formula: $AD_{\mathrm{up}}^{2*} = -2n\log(1-z_{\mathrm{H}}) + 2\sum_{j=1}^n \log\left(1-z_j\right) + \frac{1-z_{\mathrm{H}}}{n}\sum_{j=1}^n (1+2(n-j))\frac{1}{1-z_j}$

AD_{up}^{2*} statistic for left-truncated samples of the following form (the derivation is given in Appendix):

$$AD_{\mathrm{up}}^{2*} = -2n\log(1-z_{\mathrm{H}}) + 2\sum_{j=1}^n \log\left(1-z_j\right) + \frac{1-z_{\mathrm{H}}}{n}\sum_{j=1}^n (1+2(n-j))\frac{1}{1-z_j}.$$

Tables 20.1 and 20.2 summarize the EDF statistics and their computing formulae for complete and left-truncated samples.

20.5 Application to Loss Data

In this section we apply the GOF testing procedure to (1) operational loss data, extracted from an external database, and (2) catastrophe insurance claims data. The operational loss data set was obtained from Zurich IC Squared (IC^2) FIRST Database of Zurich IC Squared (IC^2), an independent consulting subsidiary of Zurich Financial Services Group. The external database is comprised of operational loss events throughout the world. The original loss data cover losses in the period 1950–2002. A few recorded data points were below \$1 million in nominal value, so we excluded them from the analysis, to make it more consistent with the conventional threshold for external databases of \$1 million. Furthermore, we excluded the observations before 1980 because of relatively few data points available (which is most likely due to poor data recording practices). The final data set for the analysis covered losses for the time period between 1980 and 2002. It consists of five types of losses: “relationship” (such as events related to legal issues, negligence, and sales-related fraud), “human” (such as events related to employee errors, physical injury, and internal fraud), “processes” (such as events related to business errors, supervision, security, and transactions), “technology” (such as events related to technology and computer failure and telecommunications), and “external” (such as events related to natural and man-made disasters and external fraud). The loss amounts have been adjusted for inflation using the Consumer Price Index from the U.S. Department of Labor. The numbers of data points of each type are $n = 849, 813, 325, 67$, and 233, respectively.

Table 20.3 Goodness-of-fit tests for operational loss data

	KS	*V*	*AD*	AD_{up}	AD^2	AD^2_{up}	W^2
Exponential							
Relationship	11.0868	11.9973	1.3·107	1.2·1023	344.37	1.2·1014	50.5365
	[<0.005]	[<0.005]	[<0.005]	[<0.005]	[<0.005]	[<0.005]	[<0.005]
Human	14.0246	14.9145	2.4·106	1.1·1022	609.15	3.0·1012	80.3703
	[<0.005]	[<0.005]	[<0.005]	[<0.005]	[<0.005]	[<0.005]	[<0.005]
Processes	7.6043	8.4160	3.7·106	1.7·1022	167.61	6.6·105	22.5762
	[<0.005]	[<0.005]	[<0.005]	[<0.005]	[<0.005]	[<0.005]	[<0.005]
Technology	3.2160	3.7431	27.6434	1.4·106	27.8369	780.50	2.9487
	[<0.005]	[<0.005]	[<0.005]	[<0.005]	[<0.005]	[<0.005]	[<0.005]
External	6.5941	6.9881	4.4·106	2.0·1022	128.35	5.0·107	17.4226
	[<0.005]	[<0.005]	[<0.005]	[<0.005]	[<0.005]	[<0.005]	[<0.005]
Catastrophe	5.5543	5.9282	9.0·106	4.1·1022	72.2643	6.1·1013	13.1717
	[<0.005]	[<0.005]	[<0.005]	[<0.005]	[<0.005]	[<0.005]	[<0.005]
Lognormal							
Relationship	0.8056	1.3341	2.6094	875.40	0.7554	4.6122	0.1012
	[0.082]	[0.138]	[0.347]	[0.593]	[0.043]	[0.401]	[0.086]
Human	0.8758	1.5265	3.9829	1086.16	0.7505	4.5160	0.0804
	[0.032]	[0.039]	[0.126]	[0.462]	[0.044]	[0.408]	[0.166]
Processes	0.6584	1.1262	2.0668	272.61	0.4624	4.0556	0.0603
	[0.297]	[0.345]	[0.508]	[0.768]	[0.223]	[0.367]	[0.294]
Technology	1.1453	1.7896	2.8456	41.8359	1.3778	6.4213	0.2087
	[<0.005]	[0.005]	[0.209]	[0.994]	[<0.005]	[0.067]	[<0.005]
External	0.6504	1.2144	2.1702	316.20	0.5816	2.5993	0.0745
	[0.326]	[0.266]	[0.469]	[0.459]	[0.120]	[0.589]	[0.210]
Catastrophe	0.6854	1.1833	5.3860	1.1·104	0.7044	27.4651	0.0912
	[0.243]	[0.307]	[0.064]	[0.053]	[0.068]	[0.023]	[0.111]
Weibull							
Relationship	0.5553	1.0821	3.8703	2.7·104	0.7073	13.8191	0.0716
	[0.625]	[0.514]	[0.138]	[0.080]	[0.072]	[0.081]	[0.249]
Human	0.8065	1.5439	4.3544	3.2·104	0.7908	8.6610	0.0823
	[0.093]	[0.051]	[0.095]	[0.068]	[0.053]	[0.112]	[0.176]
Processes	0.6110	1.0620	1.7210	2200.75	0.2069	2.2340	0.0338
	[0.455]	[0.532]	[0.766]	[0.192]	[0.875]	[0.758]	[0.755]
Technology	1.0922	1.9004	2.6821	52.5269	1.4536	4.8723	0.2281
	[<0.005]	[<0.005]	[0.216]	[0.944]	[<0.005]	[0.087]	[<0.005]
External	0.4752	0.9498	2.4314	4382.68	0.3470	5.3662	0.0337
	[0.852]	[0.726]	[0.384]	[0.108]	[0.519]	[0.164]	[0.431]
Catastrophe	0.8180	1.5438	5.6345	1.5·104	1.3975	15.8416	0.1965
	[0.096]	[0.041]	[0.043]	[0.028]	[0.007]	[0.025]	[0.006]
Pareto (GPD)							
Relationship	1.4797	2.6084	3.5954	374.68	3.7165	22.1277	0.5209
	[<0.005]	[<0.005]	[0.172]	[>0.995]	[<0.005]	[0.048]	[<0.005]

(*continued*)

Table 20.3 (continued)

	KS	V	AD	AD_{up}	AD^2	AD^2_{up}	W^2
Human	1.4022	2.3920	3.6431	374.68	2.7839	23.7015	0.3669
	[<0.005]	[<0.005]	[0.167]	[>0.995]	[<0.005]	[0.051]	[<0.005]
Processes	1.0042	1.9189	4.0380	148.24	2.6022	13.1082	0.3329
	[<0.005]	[<0.005]	[0.104]	[>0.995]	[<0.005]	[0.087]	[<0.005]
Technology	1.2202	1.8390	3.0843	33.4298	1.6182	8.8484	0.2408
	[<0.005]	[<0.005]	[0.177]	[>0.995]	[<0.005]	[0.067]	[<0.005]
External	0.9708	1.8814	2.7742	151.94	1.7091	8.6771	0.2431
	[0.009]	[0.005]	[0.284]	[0.949]	[<0.005]	[0.106]	[<0.005]
Catastrophe	0.4841	0.8671	2.4299	1277.28	0.3528	4.3053	0.0390
	[0.799]	[0.837]	[0.369]	[0.239]	[0.490]	[0.235]	[0.645]

This table reports goodness-of-fit test statistic values for operational loss data of various risk types. p-values are reported in square brackets and were obtained via 1,000 Monte Carlo simulations

The insurance claims data set covers claims resulting from natural catastrophe events occurred in the United States over the time period from 1990 to 1996. It was obtained from Insurance Services Office Inc. Property Claim Services (PCS). The data set includes 222 losses. The observations are greater than \$25 million in nominal value.

Left-truncated distributions of four types were fitted to each of the data set: exponential, lognormal, Weibull, and Pareto (GPD). Table 20.3 presents the observed statistic values and the p-values for the six data sets (five operational losses and insurance claims), obtained with the testing procedure described in Sect. 20.2.

The results reported in Table 20.3 suggest that fitting heavier-tailed distributions, such as Pareto and Weibull, results in lower values of the GOF statistics, which leads to acceptance of the null for practically all loss types and all criteria of the goodness of fit, viewed from the high p-values. Since the analysis of operational losses deals with estimating the operational value at risk (VaR), it is reasonable to determine the ultimate best fit on the basis of the AD_{up} and AD^2_{up} statistics, introduced in this paper. As can be seen from Table 20.3, these proposed measures suggest a much better fit of the heavier-tailed distributions. Moreover, for the Pareto distribution, while the statistics that focus on the center of the data (Kolmogorov-Smirnov, Kuiper, Cramér-von Mises) do not show a good fit, the AD_{up} and AD^2_{up} statistics indicate that the fit in the upper tail is very good. It should be noted that the Pareto and Weibull distributions very often suggest a superior fit in the upper tail to the lognormal distribution. Yet it is the lognormal distribution that was suggested in 2001 by the Basel Committee (BCBS (2001)).

20.6 Conclusions

In this paper we present a technique for modifying the existing goodness-of-fit test statistics so that they can be applied to loss models in which the available data set is incomplete and is truncated from below. Such left truncation is often present in loss data when the data are being recorded starting from a fixed amount and the data below are not recorded at all. Exact computing formulae for the Kolmogorov-Smirnov, Kuiper, Anderson-Darling, and Cramér-von Mises for the left-truncated samples are presented.

In risk management, it is often vital to have a good fit of a hypothesized distribution in the upper tail of the loss data. It is important in loss models that deal with value at risk, conditional value at risk, and ruin probability. We suggest using two other versions of the Anderson-Darling statistic (which we refer to as the *upper tail* Anderson-Darling statistic) in which the weighting function is proportional to the weight of only the upper tail of the distribution. Supremum and quadratic versions of the statistic are proposed. Such statistic is convenient to use when it is necessary to examine the goodness of fit of a distribution in the right tail of the data, while the fit in the left tail is unimportant.

The technique is applied to check the goodness of fit of a number of distributions using operational loss data and catastrophe insurance claims data sets. From the empirical analysis we conclude that heavier-tailed distributions better fit the data than Lognormal or thinner-tailed distributions in many instances. In particular, the conclusion is strongly supported by the *upper tail* Anderson-Darling tests.

Appendix

Derivation of AD^{2*} Computing Formula

By the PIT technique

$$AD^{2*} = n \int_H^{+\infty} \frac{\left(F_n(x) - \hat{F}^*_\theta(x)\right)^2}{\hat{F}^*_\theta(x)\left(1 - \hat{F}^*_\theta(x)\right)} d\hat{F}^*_\theta(x)$$

$$\underline{\underline{PIT}} \;\; \hat{n}^c \int_{z_{\mathrm{H}}}^{1} \frac{(F_n(z^*)(1 - z_{\mathrm{H}}) + z_{\mathrm{H}} - z)^2}{(z - z_{\mathrm{H}})(1 - z)} dz,$$

where in the original integral $F_n(Z^*) = F^*_\theta(F_n(X)) = F_n(X)$ is the empirical distribution function of $Z^* := \hat{F}^*_\theta(X) = F^*_\theta\left(\hat{F}^*_\theta(X)\right)$ so that $F^*_\theta(\cdot) \sim \mathcal{U}[0, 1]$. Changing variable and using Eq. 20.7, the integral becomes expressed in terms of $z_{\mathrm{H}} := \hat{F}_\theta(H) = F_\theta\left(\hat{F}_\theta(H)\right)$ and $Z := \hat{F}_\theta(X) = F_\theta\left(\hat{F}_\theta(X)\right)$ so that $F_\theta(\cdot) \sim U[z_H,1]$. We estimate $\hat{n}^c$ as $\frac{n}{1-z_{\mathrm{H}}}$.

Using Eq. 20.7, the computing formula is expressed in terms of $z_j := \hat{F}_\theta\left(x_{(j)}\right) = F_\theta\left(\hat{F}_\theta\left(x_{(j)}\right)\right), j = 1, 2, \ldots, n$, as

$$\frac{1}{\hat{n}^c}AD^{2*} = \underbrace{\int_{z_{\mathrm{H}}}^{z_1} \frac{(z-z_{\mathrm{H}})^2}{(z-z_{\mathrm{H}})(1-z)}dz}_{\mathrm{A}} + \underbrace{\sum_{j=1}^{n-1}\int_{z_j}^{z_{j+1}} \frac{\left(\frac{j}{n}(1-z_{\mathrm{H}})+z_{\mathrm{H}}-z\right)^2}{(z-z_{\mathrm{H}})(1-z)}dz}_{\mathrm{B}} + \underbrace{\int_{z_n}^{1} \frac{(1-z)^2}{(z-z_{\mathrm{H}})(1-z)}dz}_{\mathrm{C}}.$$

Separately solving for A, B, and C, we obtain

$$\mathrm{A} = z_{\mathrm{H}} - z_1 + (1-z_{\mathrm{H}})(\log(1-z_{\mathrm{H}}) - \log(1-z_1));$$

$$\begin{aligned}
\mathrm{B} = {} & z_1 - z_n + \frac{1-z_{\mathrm{H}}}{n^2}\sum_{j=1}^{n-1}(n-j)^2\left(\log\left(1-z_j\right) - \log\left(1-z_{j+1}\right)\right) - \ldots \\
& - 2\frac{1-z_{\mathrm{H}}}{n^2}\sum_{j=1}^{n-1} j^2\left(\log\left(z_j - z_{\mathrm{H}}\right) - \log\left(z_{j+1} - z_{\mathrm{H}}\right)\right) \\
= {} & z_1 - z_n + (1-z_{\mathrm{H}})\log(1-z_1) - \frac{1-z_{\mathrm{H}}}{n^2}\sum_{j=1}^{n}(1+2(n-j))\log\left(1-z_j\right) + \ldots \\
& + (1-z_{\mathrm{H}})\log(z_n - z_{\mathrm{H}}) + \frac{1-z_{\mathrm{H}}}{n^2}\sum_{j=1}^{n}(1-2j)\log\left(z_j - z_{\mathrm{H}}\right);
\end{aligned}$$

$$\mathrm{C} = z_n - 1 + (1-z_{\mathrm{H}})(\log(1-z_{\mathrm{H}}) - \log(z_n - z_{\mathrm{H}})).$$

Summing the terms A, B, and C, multiplying by $\hat{n}^c$, and simplifying yields the final computing formula:

$$\begin{aligned}
AD^{2*} = {} & -n + 2n\log(1-z_{\mathrm{H}}) - \frac{1}{n}\sum_{j=1}^{n}(1+2(n-j))\log\left(1-z_j\right) + \ldots \\
& + \frac{1}{n}\sum_{j=1}^{n}(1-2j)\log\left(z_j - z_{\mathrm{H}}\right).
\end{aligned}$$

Derivation of W^{2*} Computing Formula

By the PIT technique

$$\begin{aligned}
W^{2*} &= n\int_{H}^{+\infty}\left(F_n(x) - \hat{F}_\theta^*(x)\right)^2 d\hat{F}_\theta^*(x) \\
&\underline{\underline{PIT}}\ \ \hat{n}^c\int_{z_{\mathrm{H}}}^{1}\frac{(F_n(z^*)(1-z_{\mathrm{H}}) + z_{\mathrm{H}} - z)^2}{(1-z_{\mathrm{H}})^2}dz,
\end{aligned}$$

where in the original integral $F_n(Z^*) = F^*_\theta(F_n(X)) = F_n(X)$ is the empirical distribution function of $Z^* := \hat{F}^*_\theta(X) = F^*_\theta\left(\hat{F}^*_\theta(X)\right)$ so that $F^*_\theta(\cdot) \sim \mathcal{U}[0,1]$. Changing variable and using Eq. 20.7, the integral becomes expressed in terms of $z_{\mathrm{H}} := \hat{F}_\theta(H) = F_\theta\left(\hat{F}_\theta(H)\right)$ and $Z := \hat{F}_\theta(X) = F_\theta\left(\hat{F}_\theta(X)\right)$ so that $F_\theta(\cdot) \sim U[F_\theta(H), 1]$. We estimate $\hat{n}^c$ as $\frac{n}{1-z_{\mathrm{H}}}$.

Using Eq. 20.7, the computing formula is expressed in terms of $z_j := \hat{F}_\theta\left(x_{(j)}\right) = {}_{F\theta}\left(\hat{F}_\theta\left(x_{(j)}\right)\right)$, $j = 1, 2, \ldots, n$, as

$$\frac{(1-z_{\mathrm{H}})^2}{\hat{n}^c} W^{2*} = \underbrace{\int_{z_{\mathrm{H}}}^{z_1} (z_{\mathrm{H}} - z)^2 dz}_{\mathrm{A}} + \underbrace{\sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} \left(\frac{j}{n}(1-z_{\mathrm{H}}) + z_{\mathrm{H}} - z\right)^2 dz}_{\mathrm{B}} + \underbrace{\int_{z_n}^{1} (1-z)^2 dz}_{\mathrm{C}}.$$

Separately solving for A, B, and C, we obtain

$$\mathrm{A} = -\frac{z_{\mathrm{H}}{}^3}{3} + z_{\mathrm{H}}{}^2 z_1 - z_{\mathrm{H}} z_1{}^2 + \frac{z_1{}^3}{3};$$

$$\begin{aligned}
\mathrm{B} = {} & \frac{z_n{}^3}{3} - \frac{z_1{}^3}{3} + z_{\mathrm{H}} z_1{}^2 - z_{\mathrm{H}} z_n{}^2 + z_{\mathrm{H}}{}^2 z_1 + z_{\mathrm{H}}{}^2 z_n + \ldots \\
& + \frac{(1-z_H)^2}{n^2} \sum_{j=1}^{n-1} j^2 \left(z_{j+1} - z_j\right) + \frac{1-z_{\mathrm{H}}}{n} \sum_{j=1}^{n-1} j\left(z_j{}^2 - z_{j+1}{}^2\right) + 2 z_{\mathrm{H}} \frac{1-z_{\mathrm{H}}}{n} \sum_{j=1}^{n-1} j\left(z_{j+1} - z_j\right) \\
= {} & \frac{z_n{}^3}{3} - \frac{z_1{}^3}{3} + z_{\mathrm{H}} z_1{}^2 - z_{\mathrm{H}} z_n{}^2 - z_{\mathrm{H}}{}^2 z_1 + z_{\mathrm{H}}{}^2 z_n + \frac{(1-z_{\mathrm{H}})^2}{n^2} \left(n^2 z_n + \sum_{j=1}^{n} (1-2j) z_j\right) + \ldots \\
& + \frac{1-z_{\mathrm{H}}}{n} \left(\sum_{j=1}^{n} z_j{}^2 - n z_n{}^2\right) + 2 z_{\mathrm{H}} \frac{1-z_{\mathrm{H}}}{n} \left(n z_n - \sum_{j=1}^{n} z_j\right) \\
= {} & (1-z_{\mathrm{H}}) z_n + z_{\mathrm{H}}(1-z_{\mathrm{H}}) z_n - (1-z_{\mathrm{H}}) z_n{}^2 + \frac{(1-z_{\mathrm{H}})^2}{n^2} \sum_{j=1}^{n} (1-2j) z_j + \ldots \\
& + \frac{1-z_{\mathrm{H}}}{n} \sum_{j=1}^{n} z_j{}^2 - 2 z_{\mathrm{H}} \frac{1-z_{\mathrm{H}}}{n} \sum_{j=1}^{n} z_j;
\end{aligned}$$

$$\mathrm{C} = \frac{1}{3} + z_n{}^2 - z_n - \frac{z_n{}^3}{3}.$$

Summing the terms A, B, and C, multiplying by $\frac{\hat{n}^c}{(1-z_{\mathrm{H}})^2}$, and simplifying yields the final computing formula:

$$W^{2*} = \frac{n}{3} + \frac{n z_{\mathrm{H}}}{1-z_{\mathrm{H}}} + \frac{1}{n(1-z_{\mathrm{H}})} \sum_{j=1}^{n} (1-2j) z_j + \frac{1}{(1-z_{\mathrm{H}})^2} \sum_{j=1}^{n} \left(z_j - z_{\mathrm{H}}\right)^2.$$

Derivation of AD^2_{up} Computing Formula

By the PIT technique

$$AD^2_{\text{up}} = n\int_{-\infty}^{+\infty} \frac{\left(F_n(x) - \hat{F}_\theta(x)\right)^2}{\left(1 - \hat{F}_\theta(x)\right)^2} d\hat{F}_\theta(x) \underset{=}{\underline{PIT}} \; n\int_0^1 \frac{\left(F_n(z) - z\right)^2}{(1-z)^2} dz,$$

where $F_n(Z) = F_\theta(F_n(X)) = F_n(X)$ is the empirical distribution function of $Z = \hat{F}_\theta(X) = F\left(\hat{F}_\theta(X)\right)$ so that $F_\theta(\cdot) \sim U[0,1]$.

Using Eq. 20.1, the computing formula is expressed in terms of $z_j := \hat{F}_\theta\left(x_{(j)}\right) = F\left(\hat{F}_\theta\left(x_{(j)}\right)\right), j = 1, 2, \ldots, n$, as

$$\frac{1}{n}AD^2_{\text{up}} = \underbrace{\int_0^{z_1} \frac{z^2}{(1-z)^2} dz}_{\text{A}} + \underbrace{\sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} \frac{\left(\frac{j}{n} - z\right)^2}{(1-z)^2} dz}_{\text{B}} + \underbrace{\int_{z_n}^{1} \frac{(1-z)^2}{(1-z)^2} dz}_{\text{C}}.$$

Separately solving for A, B, and C, we obtain

$$\text{A} = z_1 - 1 + \frac{1}{1-z_1} + 2\log(1-z_1);$$

$$\text{B} = z_n - z_1 - \frac{1}{n^2}\sum_{j=1}^{n-1}(n-j)^2\left(\frac{1}{1-z_j} - \frac{1}{1-z_{j+1}}\right) - \ldots$$

$$-2\frac{1}{n}\sum_{j=1}^{n-1}(n-j)\left(\log\left(1-z_j\right) - \log\left(1-z_{j+1}\right)\right)$$

$$= z_n - z_1 - \frac{1}{1-z_1} + \frac{1}{n^2}\sum_{j=1}^{n}(1+2(n-j))\frac{1}{1-z_j} - 2\log(1-z_1) + 2\frac{1}{n}\sum_{j=1}^{n}\log\left(1-z_j\right);$$

$$\text{C} = 1 - z_n.$$

Summing the terms A, B, and C, multiplying by n, and simplifying yields the final computing formula:

$$AD^2_{\text{up}} = 2\sum_{j=1}^{n}\log\left(1-z_j\right) + \frac{1}{n}\sum_{j=1}^{n}(1+2(n-j))\frac{1}{1-z_j}.$$

Derivation of AD^{2*}_{up} Computing Formula

By the PIT technique

$$AD^{2*}_{\text{up}} = n\int\limits_{H}^{+\infty} \frac{\left(F_n(x) - \hat{F}^*_\theta(x)\right)^2}{\left(1-\hat{F}^*_\theta(x)\right)^2} d\hat{F}^*_\theta(x) \underline{\underline{PIT}}\ \hat{n}^c \int\limits_{z_{\text{H}}}^{1} \frac{(F_n(z^*)(1-z_{\text{H}})+z_{\text{H}}-z)^2}{(1-z)^2} dz,$$

where in the original integral $F_n(Z^*) = F^*_\theta(F_n(X)) = F_n(X)$ is the empirical distribution function of $Z^* := \hat{F}^*_\theta(X) = F^*\left(\hat{F}^*_\theta(X)\right)$ so that $F^*_\theta(\cdot) \sim \mathcal{U}[0,1]$. Changing variable and using Eq. 20.7, the integral becomes expressed in terms of $z_{\text{H}} := \hat{F}_\theta(H) = F\big(\hat{F}_\theta(H)\big)$ and $Z := \hat{F}_\theta(X) = F_\theta\big(\hat{F}_\theta(X)\big)$ so that $F_\theta(\cdot) \sim U[z_{\text{H}},1]$. We estimate $\hat{n}^c$ as $\frac{n}{1-z_{\text{H}}}$.

Using Eq. 20.7, the computing formula is expressed in terms of $z_j := \hat{F}_\theta\left(x_{(j)}\right) = F\left(\hat{F}_\theta\left(x_{(j)}\right)\right), j = 1, 2, \ldots, n$ as

$$\frac{1}{\hat{n}^c} AD^{2*}_{\text{up}} = \underbrace{\int_{z_{\text{H}}}^{z_1} \frac{(z-z_{\text{H}})^2}{(1-z)^2} dz}_{\text{A}} + \underbrace{\sum_{j=1}^{n-1} \int_{z_j}^{z_{j+1}} \frac{\left(\frac{j}{n}(1-z_{\text{H}})+z_{\text{H}}-z\right)^2}{(1-z)^2} dz}_{\text{B}} + \underbrace{\int_{z_n}^{1} \frac{(1-z)^2}{(1-z)^2} dz}_{\text{C}}.$$

Separately solving for A, B and C, we obtain

$$\begin{aligned} \text{A} = {} & z_1 - z_{\text{H}} - (1-z_{\text{H}}) + (1-z_{\text{H}})^2 \frac{1}{1-z_1} - 2(1-z_{\text{H}})\log(1-z_{\text{H}}) \\ & + 2(1-z_{\text{H}})\log(1-z_1); \end{aligned}$$

$$\begin{aligned} \text{B} = {} & z_n - z_1 - \frac{(1-z_{\text{H}})^2}{n^2} \sum_{j=1}^{n-1} (n-j)^2 \left(\frac{1}{1-z_j} - \frac{1}{1-z_{j+1}}\right) - \ldots \\ & - 2\frac{1-z_{\text{H}}}{n} \sum_{j=1}^{n-1} (n-j)\big(\log\big(1-z_j\big) - \log\big(1-z_{j+1}\big)\big) \\ = {} & z_n - z_1 - (1-z_{\text{H}})^2 \frac{1}{1-z_1} + \frac{(1-z_{\text{H}})^2}{n^2} \sum_{j=1}^{n} (1+2(n-j)) \frac{1}{1-z_j} - \ldots \\ & - 2(1-z_{\text{H}})\log(1-z_1) + 2\frac{1-z_{\text{H}}}{n} \sum_{j=1}^{n} \log\big(1-z_j\big); \end{aligned}$$

$$\text{C} = 1 - z_n.$$

Summing the terms A, B, and C, multiplying by $\hat{n}^c$, and simplifying yields the final computing formula:

$$AD_{\mathrm{up}}^{2*} = -2n\log(1 - z_{\mathrm{H}}) + 2\sum_{j=1}^{n} \log\left(1 - z_j\right) + \frac{1 - z_{\mathrm{H}}}{n}\sum_{j=1}^{n} \left(1 + 2(n - j)\right)\frac{1}{1 - z_j}.$$

References

Anderson, T. W., & Darling, D. A. (1952). Asymptotic theory of certain ‘Goodness of Fit’ criteria based on stochastic processes. *The Annals of Mathematical Statistics, 23*, 193–212.

Anderson, T. W., & Darling, D. A. (1954). A test of goodness of fit. *Journal of the American Statistical Association, 49*, 765–769.

BCBS. (2001). *Working paper on the regulatory treatment of operational risk*. Basel, Switzerland: Basel Committee on Banking Supervision. http://www.bis.org

BCBS. (2003). *The 2002 loss data collection exercise for operational risk: summary of the data collected*. Basel, Switzerland: Basel Committee on Banking Supervision. http://www.bis.org

D’Agostino, R., & Stephens, M. (1986). *Goodness-of-fit techniques*. New York/Basel: Marcel Dekker.

Deheuvels, P., & Martynov, G. (2003). Karhunen-Loève expansions for weighted Wiener processes and Brownian bridges via Bessel functions. Progress in Probability, 55, 57–93.

Dufour, R., & Maag, U. R. (1978). Distribution results for modified kolmogorov-smirnov statistics for truncated or censored samples. *Technometrics, 20*, 29–32.

Gastaldi, T. (1993). A kolmogorov-smirnov test procedure involving a possibly censored or truncated sample. *Communications in Statistics: Theory and Method, 22*, 31–39.

Guilbaud, O. (1998). Exact kolmogorov-type test for left-truncated and/or right-censored data. *Journal of American Statistical Association, 83*, 213–221.

Ross, S. (2001). *Simulation* (3rd ed.). Boston: Academic.

Shorack, G., & Wellner, J. (1986). *Empirical processes with applications to statistics*. New York: Wiley.

Effect of Merger on the Credit Rating and Performance of Taiwan Security Firms

21

Suresh Srivastava and Ken Hung

Contents

S. Srivastava (✉)
University of Alaska Anchorage, Anchorage, AK, USA
e-mail: afscs@uaa.alaska.edu

K. Hung
Division of International Banking & Finance Studies, Texas A&M International University, Laredo, TX, USA
e-mail: ken.hung@tamiu.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_21,

Abstract

The effect of a merger on credit rating is investigated by testing the significance of change in a firm's rank based on comprehensive performance score and synergistic gains. We extract principle component factors from a set of financial ratios. Percentage of variability explained and factor loadings are adjusted to get a modified average weight for each financial ratio. This weight is multiplied by the standardized Z value of the variable, and summed a set of variables to get a firm's performance score. Performance scores are used to rank the firm. Statistical significance of difference in pre- and post-merger rank is tested using the Wilcoxon sign rank (double end).

We studied the merger of financial firms after the enactment of Taiwan's Merger Law for Financial Institution in November 2000 to examine synergies produced by merger. Synergistic gains affect corporate credit ratings. After taking into account the large Taiwan market decline from 1999 to 2000, test results show there is no significant operating, market, and financial synergy produced by the merger firms. Most likely explanations for the insignificant rank changes are short observation period and the lack of an adequate sample in this investigation.

We identify and define variables for merger synergy analysis followed by principal component factor analysis, variability percentage adjustment, and performance score calculation. Finally, Wilcoxon sign rank test is used for hypothesis testing. Reader is well referred to the appendix for details.

Keywords

Corporate merger • Financial ratios • Synergy • Economies of scale • Credit rating • Variability percentage adjustment • Principle component factors • Firm's performance score • Standardized Z • Wilcoxon rank test

21.1 Introduction

Corporate credit rating helps in determining the soundness of a financial system. It gives lenders and venture capitalists confidence in making direct investment in domestic and foreign countries. Credit rating determines the probability that the corporation will be able to meet its obligations. There are a number of credit rating agencies worldwide.[1] The top three agencies that deal in credit ratings are Moody's, Standard & Poor's, and Fitch IBCA. Independent objective assessments of the credit worthiness of companies help investors decide the riskiness of the security issued by the firm. Credit rating institutions base their subjective

[1]World leading credit rating agencies are Moody's Investors Service, Standard & Poor's Corp, Fitch Investors Service, Duff & Phelps, Japan Bond Research Institution, Nippon Investors Service, Japan Credit Rating Agency, China Credit Rating Agency, and Taiwan New Economy Newspaper (Weston and Mansinghka 1971; Williamson 1981).

rating judgments by the experiences and expertise of their analyst teams. Their public confidences come from their reputation. Yet different analysts or different credit rating institutions may give different meanings or powers to the likelihood of the same incident. Credit rating agencies often revise their ratings after a corporate event. This research deals with the event of corporate mergers.

A merger is the quickest method of corporate growth. At the end of the twentieth century, there was an upsurge of mergers in Europe, the USA, and Asia. The same was true for Taiwan. Mergers occurred in various industries of Taiwan, with financial industry the most dominant. For example, the merger of Bank of Taiwan, Local Bank, and the Central Credit Bureau was the first large merger initiated by the government. The institution resulting from the merger ranked in the top 70 in the world. Yuan Dao, the number one corporation in the security industry, captured 10 % of the Taiwanese brokerage business after its acquisition of Jing Hua. The Merger Law for Financial Institutions passed in November 2000 and permitted the merger of domestic financial institutions, such as banks, insurance companies, and security companies, and allowed for the establishment of assets Management Corporation. After a merger, operating income, market share, revenues, and total asset increase. But does a merger help to improve corporate quality and competence? When does merger synergy manifest? What are its merits? There are a number of questions that need discussion and analysis. The scope of this paper is limited. We study the effect of mergers on credit rating by testing the significance of change in firm's rank based on comprehensive performance score and synergistic gains.

21.2 Merger Literature

Firms merge with the stated intent of shareholders' wealth maximization. Value maximization is achieved by increased profit, reduced risk, or both. Theories that posit wealth maximization are efficiency theory (synergy), information and signaling theory, and tax advantage.[2] Non-value maximization concerns with agency theory. In this research we focus primarily on efficiency theory, also called synergy theory. Synergy theory implies three components: operation synergy, market synergy, and financial synergy. Operation synergy is produced due to economies of scale, operating reduction, and sharing of management expertise. Under market synergy, a merger reduces market competitors and increases market concentration leading to a monopolization or increased market share. The increased market share produces superior profits via pricing strategy or corporate collusion. The merging corporation increases its market share so as to influence product price

[2]Theory of merger is discussed in Brigham and Daves (2007), Copeland and Weston (1979), Gardner (1996), and Rhodes (1983). Tax incentive is discussed in Weston and Chung (1983).

and quantity and to achieve market synergy. However, corporations in pursuit of market synergy must abide by the fair trade law. Seth (1990) said it was easy to achieve market synergy via horizontal merger instead of pricing strategy or corporate collusion. Financial synergy reduces the systematic risk and lowers the capital cost. There is extensive literature of merger-related studies. The following is a selected sample.

Ansoff et al. (1971) used a questionnaire to collect 12 financial variables: sales, retained earnings, total asset, return on capital, etc. Their research showed that sales management, technology, and R&D produced operating synergy. Weston and Mansinghka (1971) estimated management energy with corporate growth using variables: total asset, sales, net profit, retained earnings, and stock price. Their research indicates that a conglomerate merger gives the corporation a higher growth rate than the growth of other corporations within the same industry. Beattie (1980) used nonsystem risk and stock return to estimate merger synergy. His finding is that after a merger, nonsystem risk declines, yet corporate stock returns do not increase. Hoshino (1982) estimated corporate synergy with seven financial data: the ratio of net worth and total liabilities, ratio between net worth and total assets, liquidity ratio, ratio of interest rate to debt, turnover ratio, ratio of net profit and total liabilities, and ratio of net profit to total assets. His finding is that a merged corporation has increased liquidity, but lower profitability and stability. Muller (1985) used market share percentage to estimate the effect of merger. His finding is that there is no significant change in market share; some companies even lose market share after merger. Sigh and Montgomery (1987) used stock returns to discuss related and non-related merger synergy. Their finding is that a related merger has a higher return. Healy et al. (1992) used the ratio between before-tax capital flow and total assets to estimate merger synergy. Their finding is that the return on operating cash flow improves significantly. Williamson (1981) stated that vertical merger could reduce the communication cost between upper and lower stream corporations, cost of product quality check, storage cost, and delivery cost. Yet Scherer (1980) thought a different market configuration could impact the gains due to economies of scale. The more complete the market competition is, the less is the gain due to economies of scale. On the other hand, Rhodes (1983) thought that the resources configuration within a corporation is complicated and bureaucratic. Hence, the internal capital market is less efficient than the external capital market. In support of financial synergy, Fluck and Lynch (1999) pointed out that after a merger, a corporation could have greater investment opportunities at lower cost than before. Lewellen (1971) suggested the merger could reduce the capital cost. Levy and Sarnat (1970) thought stockholders could achieve risk reduction via portfolio diversification at a lower cost than that of the merger. Higgins and Schall (1975) pointed out the coinsurance effect and hence a reduced cost of bankruptcy. Hence, there is a transfer of wealth from creditors to stockholders from creditors, as stockholders could reissue bonds with a lower interest rate.

Table 21.1 Merged security firms

Merged security firms	Survived corporation	Date of announcement	Date of merger
YuanDa, JinHua, Dafa	YuanDaJinHua	Nov. 29, 1999	July 01, 2000
YuanFu, JiaHe, YongSheng	YuanFu	Jan. 28, 1999	July 24, 2000
JianHong, WanSheng	JianHong	Feb. 11, 2000	Aug. 28, 2000
BaoLai, DaShun, ShiTai, HuaYu	BaoLai	Feb. 19, 2000	Sep. 9, 2000
FuBang, HuangQiu, ZhongRi, JinShan, HuaXin, ShiLing, KuaiLe	FuBang	Feb. 21, 2000	Sep. 9, 2000

13 non-merger firms used to test the model are as follows: TaiYu, DaHua, QunYi, ZhongXin, YongChang, TaiZhen, JinDing, RiShen, DaXin, KangHe, YaZhou, XinBao, and TongYi

21.3 Empirical Analysis

The effect of a merger on credit rating is investigated by (1) testing the significance of change in a firm's rank based on a comprehensive performance score and (2) examining post-merger synergy. Synergistic gains from a merger could be in the form of operating (management) synergy, market synergy, and/or financial synergy. Data for this research cover January 1999 to December 2000. Premerger analysis covers year 1999, and post-merger analysis covers year 2000.[3] Definitions of financial ratios are presented in Appendix 1. The methodology of factor extraction and weight assignment is discussed in Appendix 2. Since the analyses are based on ranks, the usual "*t*-test" method is unsuitable to examine the significance of change in operating, financial, or market performance pre- and post-merger. Hence, we use the Wilcoxon Sign Rank Test that checks whether two sets of ranks, pre- and post-merger, come from the same sample or two samples (Appendix 3).[4] In Table 21.1, we list five acquiring financial institutions and the acquired institutions. The footnote of Table 21.1 lists 13 non-merger firms. These 13 firms are included in the analysis for comparison.

21.3.1 Comprehensive Performance Score

Principle component factor analysis was used to extract common factors with factor loading greater than 1. Using 12-month financial ratio covering 1999, we produced five factors (Table 21.2, Panel A). The variability explained by five extracted factors is 24.458 %, 23.764 %, 13.268 %, 12 %, and 11.031 %.

[3]3-month and 6-month analyses were performed but are omitted as these did not add additional foresight.

[4]References for Statistical and Econometric issues are Johnson and Wichern (2007) and Woolridge (2009).

Table 21.2 Extracted factors and explained variability

		Explained variability		
Factor	Factor loading	Percentage	Cumulative percentage	Adjusted percentage
Panel A: December 1999				
1	4.647	24.458	24.458	28.937
2	4.515	23.764	48.222	28.116
3	2.521	13.268	61.49	15.698
4	2.28	12	73.49	14.198
5	2.096	11.031	84.521	13.051
Total	16.059	84.521	84.521	100
Panel B: December 2000				
1	4.245	22.34	22.34	24.746
2	3.844	20.229	42.569	22.407
3	2.991	15.74	58.309	17.435
4	2.377	12.51	70.819	13.857
5	2.011	10.584	81.403	11.724
6	1.686	8.876	90.279	9.832
Total	17.154	90.279	90.279	100.000

Principle component factors. See Appendix 2

The percentage of variability explained was adjusted to make the total variability explained equal to 100. The adjusted percentage of variability explained by five extracted factors is 28.937 %, 28.116 %, 15.698 %, 14.198 %, and 13.051 %. Using 12-month financial ratio covering 2000, six extracted common factors with factor loading greater than 1 are presented in Table 21.2 (Panel B). The adjusted percentage of variability explained by six extracted factors is 24.746 %, 22.407 %, 17.435 %, 13.857 %, 11.724 %, and 9.832 %. The adjusted variability percentage is multiplied by factor loading and summed over all variables to get the total loading for a variable. Then factor loadings are adjusted such that total factor loadings add to 100 (Eq. 21.2). Initial and adjusted weights for each variable are presented in Table 21.3. Finally, the values of adjusted weights for 1999 and 2000 are averaged and sign modified to reflect a positive or negative variable. It is presented in the last column of Table 21.3 and used to calculate comprehensive performance score. The modified average weight for each variable is multiplied by the standardized Z value of the variable and summed over all the variables to get a firm's performance score and comprehensive rank (Table 21.4). The premerger comprehensive ranks of the five merger firms and 13 non-merger firms are listed in Table 21.4, Panel A. Panel B of Table 21.4 presents comprehensive post-merger comprehensive ranks. Rank changes in Table 21.4 can be summarized as JianHong dropped from rank 1 to 6, YuanDaJinHua rose from rank 3 to 2, YuanFu rose from rank 7 to 1, FuBang dropped from rank 2 to 4, and BaoLai rose from rank 16 to 8. The test of significance of the comprehensive rank differences, for

Table 21.3 Adjusted variable weights for comprehensive performance analysis

Item	Variable	December 1999 Initial weight	December 1999 Adjusted weight	December 2000 Initial weight	December 2000 Adjusted weight	Average weight[a]
Financial structure	X_{i1}	−0.254	−9.010	−0.192	−4.404	6.707
	X_{i2}	0.263	9.329	−0.025	−0.583	4.373
Solvency	X_{i3}	0.388	13.741	0.321	7.358	10.549
	X_{i4}	0.408	14.455	0.323	7.405	10.930
Asset utilization	X_{i5}	0.127	4.489	0.310	7.107	5.798
	X_{i6}	−0.214	−7.575	0.171	3.925	1.825
Profitability	X_{i7}	0.333	11.806	0.396	9.077	10.442
	X_{i8}	0.324	11.472	0.380	8.704	10.088
	X_{i9}	0.327	11.573	0.366	8.399	9.986
	X_{i10}	0.302	10.706	0.372	8.525	9.616
Cash flow	X_{i11}	−0.021	−0.732	0.325	7.459	3.363
	X_{i12}	−0.042	−1.487	0.174	3.977	1.245
	X_{i13}	0.035	1.232	0.207	4.743	2.988
Growth	X_{i14}	0.194	6.891	0.282	6.456	6.674
	X_{i15}	0.214	7.575	0.193	4.434	6.005
Size	X_{i16}	0.185	6.559	0.255	5.833	6.196
	X_{i17}	0.265	9.398	0.273	6.263	7.831
Industry specific	X_{i18}	0.289	10.229	0.201	4.604	7.417
	X_{i19}	−0.301	−10.654	0.031	0.717	4.969
	Total	2.821	100.000	4.363	100.000	

Variable definitions are in Appendix 1

[a]Sign of average variable weights is changed to reflect a positive or negative variable

merger and non-merger firms, was conducted using the Wilcoxon sign rank (double end) test discussed in Appendix 3. Panel A of Table 21.5 indicates that the merger firm's comprehensive rank difference is statistically insignificant. Panel B of Table 21.5 indicates that the non-merger firm's comprehensive rank difference is also statistically insignificant. Plausible explanations for the insignificant rank changes are a short observation period, the lack of adequate samples, and the sharp decline in JianHong's performance score. Another plausible reason may be that some security firms (such as YuanDaJinHua and YuanFu) have already achieved high premerger comprehensive performance scores, and any improvement in post-merger performance did not make a significant difference.

21.3.2 Test of Merger Synergy

We use 12-month data to examine operating synergy, market synergy, and financial synergy. If operating synergy exists, then operating cost ratio will be reduced

Table 21.4 Security firm's comprehensive performance score and rank

Firm	Comprehensive score	Rank	Firm	Comprehensive score	Rank
Panel A: December 1999					
JianHong[a]	112.441	1	YongChang	4.045	10
FuBang[a]	90.757	2	ZhongXin	2.177	11
YuanDaJinHua[a]	88.033	3	TaiYu	−3.648	12
TongYi	57.312	4	DaHua	−4.140	13
QunYi	47.620	5	KangHe	−47.811	14
RiShen	47.457	6	JinDing	−81.604	15
YuanFu[a]	40.041	7	BaoLai[a]	−106.382	16
TaiZhen	21.349	8	YaZhou	−116.505	17
XinBao	12.180	9	DaXin	−163.323	18
Panel B: December 2000					
YuanFu[a]	162.4621	1	TaiZhen	−10.0671	10
YuanDaJinHua[a]	132.0642	2	DaHua	−19.123	11
RiShen	118.3488	3	KangHe	−29.1604	12
FuBang[a]	77.26606	4	YaZhou	−75.6848	13
ZhongXin	50.47429	5	YongChang	−77.5209	14
JianHong[a]	48.7586	6	TaiYu	−96.5883	15
QunYu	32.33424	7	JinDing	−98.3886	16
BaoLai[a]	31.40838	8	XinBao	−101.256	17
TongYi	16.56453	9	DaXin	−150.395	18

[a]means Wilcoxon sign rank test for performance score difference between before after merger is statistically significant

significantly; if market synergy exists, then two other ratios (ratio of operating income to total assets and market share change rate) will increase significantly; and existence of financial synergy is indicated by the decline in variability of operating (business) risk.

Some financial variables of security firms are greatly influenced by market conditions. The operating income, in particular, is greatly influenced by a bull or bear market. In years 1999 and 2000, the Taiwan stock market fell from over 10,000 points to over 5,000 points. In order to reduce the market impact on variables, we adjust the operating cost to operating income ratio and the operating income to total asset ratio. The operating cost ratio and the operating return on assets are normalized by the industry average of the ratio.

Table 21.6 presents a test of significance of operating synergy for merger and non-merger firms. The test result shows there are no significant changes of operating costs ratio for merger firms. But for non-merger firms, there are significant increases of operating cost ratio. Then operating income is adjusted for market drop in 2000, and the significance of operating synergy is tested again. The test result is presented in Table 21.7.

Table 21.5 Test of significance of merger and non-merger firms' comprehensive rating

Firm	December 1999 ranks	December 2000 ranks	Difference	D_i	Class	W
Panel A: merger firms						
JianHong	1	6	−5	5	3	W (+) = 33
YuanDaJinHua	3	2	1	1	1	W (−) = 45
YuanFu	7	1	6	6	4	W = 33
FuBang	2	4	−2	2	2	
BaoLai	16	8	8	8	5	
Panel B: non-merger firms						
TaiYu	12	15	−3	3	6.5	W (+) = 33
DaHua	13	11	2	2	3.5	W (−) = 45
QunYi	5	7	−2	2	3.5	W = 33
ZhongXin	11	5	6	6	11	
YongChang	10	14	−4	4	8.5	
TaiZhen	8	10	−2	2	3.5	
JinDing	15	16	−1	1	1	
RiShen	6	3	3	3	6.5	
DaXin	18	18	0			
KangHe	14	12	2	2	3.5	
YaZhou	17	13	4	4	8.5	
XinBao	9	17	−8	8	12	
TongYi	4	9	−5	5	10	

Wilcoxonsign rank (double end) test indicates that rank difference is statistically *insignificant*, Eq. 21.6

The test of significance of merger synergy is performed by examining two financial ratios: ratio of operating income to total assets (operating return ratio) and market share. Table 21.8 presents the first test of significance of market synergy using operating return ratio. Result shows there are *significant* increases in the ratio of operating income to total assets for both merger and non-merger security firms. This indicates a positive market synergy. We repeat the test of significance after adjusting the operating income to account for the market decline. Results presented in Table 21.9 show there are insignificant differences in the ratio of operating income to total assets for both merger and non-merger security firms. The second test of significance of market synergy using market share change is presented in Table 21.10. Results in Table 21.10 show there are significant increases of market share for merger security firms. This means positive market synergy. For the non-merger firms, there are *significant* declines of market share. Taking the two variables into consideration, we state that positive market synergy is produced by the merger of security firms. We use the variability of operating risk to assess financial synergy. The test results in Table 21.11 show there is insignificant financial synergy for merger firms. However, there is significant change in the variability of operating risk for non-merger security firms.

Table 21.6 Test of significance of merger and non-merger firms' operating cost ratio[a]

Firm	December 1999 score	December 2000 score	Difference	D_i	Class	W
Panel A: merged firms						
JianHong	71.634	67.332	4.302	4.302	2	W (+) = 3
YuanDaJinHua	58.689	78.167	−19.478	19.478	5	W (−) = 12
YuanFu	71.857	79.305	−7.448	7.448	4	W = 3
FuBang	70.040	75.378	−5.338	5.338	3	
BaoLai	92.007	91.503	0.503	0.503	1	
Panel B: non-merger firms						
TaiYu	76.548	122.296	−45.748	45.748	12	W (+) = 3
DaHua	84.481	83.181	1.299	1.299	2	W (−) = 88
QunYi	71.105	73.431	−2.325	2.325	4	W = 3
ZhongXin	79.395	84.665	−5.269	5.269	5	
YongChang	79.471	111.708	−32.238	32.238	11	
TaiZhen	81.915	95.024	−13.109	13.109	8	
JinDing	98.436	107.020	−8.584	8.584	6	
RiShen	66.200	66.167	0.034	0.034	1	
DaXin	121.563	143.634	−22.071	22.071	10	
KangHe	81.531	93.207	−11.676	11.676	7	
YaZhou	104.699	106.014	−1.316	1.316	3	
XinBao	74.861	136.984	−62.123	62.123	13	
TongYi	63.744	85.364	−21.621	21.621	9	

[a]Operating cost/operating income
Wilcoxon sign rank (double end) test indicates that rank difference is statistically *insignificant*, Eq. 21.6

21.4 Conclusion

The effect of a merger on credit rating was examined by testing the significance of change in firm's rank based on a comprehensive performance score and examining post-merger synergy. Synergistic gains from a merger could be in the form of operating synergy, market synergy, and/or financial synergy. Our test showed that merger and non-merger firms' comprehensive performance rank difference was statistically insignificant. Plausible explanations for the insignificant rank changes are a short observation period and the lack of adequate samples. Other plausible explanation may be that some security firms have already achieved high premerger comprehensive performance scores, and any improvement in post-merger performance did not make a significant difference.

We used a standardized score of operating cost ratio to rank firms for their operating synergy. Test results show there is no significant operating synergy for merger and non-merger security firms. We used standardized scores of ratio of operating income to total assets and market share change to rank firms for their market synergy. Test result shows there are *significant* increases in the ratio of

Table 21.7 Test of significance of merger and non-merger firms' adjusted operating cost ratio[a]

Firm	December 1999 score	December 2000 score	Difference	D_i	Class	W
Panel A: merger firms						
JianHong	93.492	79.475	14.017	14.017	4	W (+) = 10
YuanDaJinHua	76.597	92.263	−15.667	15.667	5	W (−) = 5
YuanFu	93.783	93.607	0.176	0.176	1	W = 5
FuBang	91.412	88.972	2.440	2.440	2	
BaoLai	120.081	108.005	12.076	12.076	3	
Panel B: non-merger firms						
TaiYu	99.905	144.351	−44.446	44.446	12	W (+) = 32
DaHua	110.259	98.183	12.076	12.076	9	W (−) = 59
QunYi	92.802	86.674	6.129	6.129	5	W = 32
ZhongXin	103.622	99.933	3.688	3.688	3	
YongChang	103.720	131.854	−28.134	28.134	11	
TaiZhen	106.911	112.161	−5.250	5.250	4	
JinDing	128.472	126.320	2.152	2.152	1	
RiShen	86.400	78.099	8.301	8.301	6	
DaXin	158.656	169.537	−10.881	10.881	7	
KangHe	106.409	110.016	−3.608	3.608	2	
YaZhou	136.646	125.133	11.513	11.513	8	
XinBao	97.704	161.688	−63.985	63.985	13	
TongYi	83.194	100.759	−17.565	17.565	10	

[a]Operating cost/operating income adjusted for the drop in the market
Wilcoxon sign rank (double end) test indicates that rank difference is statistically *insignificant*, Eq. 21.6

operating income to total assets for both merger and non-merger security firms, indicating a positive market synergy. However, after adjusting the operating income market decline, the results indicate insignificant differences in the ratio of operating income to total assets for both merger and non-merger security firms. We used standardized score of variability of operating risk to rank firms for their financial synergy. The test results show there is insignificant financial synergy for merger firms; however, there is significant change in the variability of operating risk for non-merger security firms.

Appendix 1: Variables for Merger Synergy Analysis

Post-merger credit rating of the firm depends on the extent of synergy produced by the merger. Components of merger synergy are operating synergy, market synergy, and financial synergy.

Each synergy component is determined by firm characteristics: financial structure, solvency, asset utilization, profitability, cash flow, growth, scale, and industry-specific ratio. In this study a number of financial ratios are used to assess

Table 21.8 Test of significance of merger and non-merger firms' operating return on assets[a]

Firm	December 1999 score	December 2000 score	Difference	D_i	Class	W
Panel A: merger firms						
JianHong	16.868	17.171	−0.303	0.303	1	W (+) = 0
YuanDaJinHua	13.591	22.133	−8.542	8.542	3	W (−) = 15
YuanFu	14.667	26.617	−11.949	11.949	4	W = 0
FuBang	14.245	16.876	−2.630	2.630	2	
BaoLai	12.429	25.243	−12.814	12.814	5	
Panel B: non-merger firms						
TaiYu	12.881	15.878	−2.997	2.997	12	W (+) = 10
DaHua	15.783	12.746	3.037	3.037	2	W (−) = 81
QunYi	13.893	18.924	−5.031	5.031	4	W = 10
ZhongXin	13.214	26.546	−13.332	13.332	5	
YongChang	16.797	19.702	−2.905	2.905	11	
TaiZhen	12.783	14.188	−1.405	1.405	8	
JinDing	11.884	15.938	−4.054	4.054	6	
RiShen	18.111	24.489	−6.379	6.379	1	
DaXin	15.232	15.240	−0.008	0.008	10	
KangHe	14.146	16.561	−2.414	2.414	7	
YaZhou	16.653	16.268	0.385	0.385	3	
XinBao	13.316	15.023	−1.706	1.706	13	
TongYi	13.732	21.170	−7.438	7.438	9	

[a]Operating income/total assets
Wilcoxon sign rank (double end) test indicates that ratio difference is statistically *significant*, Eq. 21.6

operating performance of the firm before and after the merger. The following section outlines firm characteristics and variables to measure operating performance.

Merger Synergy

Operating synergy – it refers to the improvement of operating efficiency achieved via scale economy, transaction cost economy, and differential efficiency caused by merger. *Market synergy* – increase in market share due to enhanced negotiating power and dominant pricing strategy.

Financial synergy – diversification of financial risk and cost of capital reduction.

Operating Synergy

A. – Financial structure
X_{i1} – Debt ratio
X_{i2} – Ratio of long-term capital to fixed assets

Table 21.9 Test of significance of merger and non-merger firms' adjusted operating return on assets[a]

Firm	December 1999 score	December 2000 score	Difference	D_i	Class	W
Panel A: merger firms						
JianHong	0.0190	0.0154	0.0035	0.0035	2	W (+) = 3
YuanDaJinHua	0.0153	0.0199	−0.0046	0.0046	3	W (−) = 12
YuanFu	0.0165	0.0239	−0.0074	0.0074	4	W = 3
FuBang	0.0160	0.0152	0.0008	0.0008	1	
BaoLai	0.0140	0.0227	−0.0087	0.0087	5	
Panel B: non-merger firms						
TaiYu	0.0145	0.0143	0.0002	0.0002	1	W (+) = 53
DaHua	0.0177	0.0115	0.0063	0.0063	12	W (−) = 38
QunYi	0.0156	0.0170	−0.0014	0.0014	5	W = 33
ZhongXin	0.0149	0.0239	−0.0090	0.0090	13	
YongChang	0.0189	0.0177	0.0012	0.0012	4	
TaiZhen	0.0144	0.0128	0.0016	0.0016	7	
JinDing	0.0134	0.0143	−0.0010	0.0010	2	
RiShen	0.0204	0.0220	−0.0017	0.0017	8	
DaXin	0.0171	0.0137	0.0034	0.0034	9	
KangHe	0.0159	0.0149	0.0010	0.0010	3	
YaZhou	0.0187	0.0146	0.0041	0.0041	11	
XinBao	0.0150	0.0135	0.0015	0.0015	6	
TongYi	0.0154	0.0190	−0.0036	0.0036	10	

[a](Operating income/total assets) adjusted for the drop in the market
Wilcoxon sign rank (double end) test indicates that ratio difference is statistically *insignificant*, Eq. 21.6

B. Solvency
  X_{i3}– Current ratio
  X_{i4}– Quick ratio
C. Asset utilization
  X_{i5} – Operating return on assets
  X_{i6} – Net worth turnover ratio
D. Profitability
  X_{i7} – Return on assets
  X_{i8} – Return on equity
  X_{i9} – Profit margin
  X_{i10} – Earnings per share
E. Cash flow
  X_{i11} – Cash flow to short-term liability
  X_{i12} – 5-year cash flow to debt obligations
  X_{i13} – Retention ratio
F. Growth
  X_{i14} – Growth in revenue
  X_{i15} – Growth in earnings

Table 21.10 Test of significance of merger and non-merger firms' change in market share

Firm	December 1999 score	December 2000 score	Difference	D_i	Class	W
Panel A: merger firms						
JianHong	6.169	5.255	0.914	0.914	1	W (+) = 1
YuanDaJinHua	11.463	17.270	−5.807	5.807	5	W (−) = 14
YuanFu	5.992	7.497	−1.505	1.505	2	W = 1
FuBang	6.007	7.698	−1.692	1.692	3	
BaoLai	5.745	8.076	−2.331	2.331	4	
Panel B: non-merger firms						
TaiYu	1.982	1.475	0.507	0.507	5	W (+) = 80
DaHua	9.589	5.589	4.000	4.000	13	W (−) = 11
QunYi	8.389	8.197	0.192	0.192	1	W = 11
ZhongXin	4.932	6.402	−1.470	1.470	11	
YongChang	3.090	2.559	0.531	0.531	6	
TaiZhen	5.462	4.929	0.532	0.532	7	
JinDing	5.128	4.678	0.450	0.450	4	
RiShen	9.547	7.487	2.061	2.061	12	
DaXin	3.109	1.830	1.280	1.280	10	
KangHe	2.580	2.209	0.371	0.371	3	
YaZhou	2.300	1.462	0.838	0.838	8	
XinBao	1.431	1.178	0.253	0.253	2	
TongYi	7.085	6.210	0.876	0.876	9	

Wilcoxon sign rank (double end) test indicates that ratio difference is statistically *significant*, Eq. 21.6

G. Size
  X_{i16} – Total assets
  X_{i17} – Net worth
H. Industry-specific ratios
  X_{i18} – Consignment to current assets
  X_{i19} – Long-term financing to net worth

Market Synergy

X_{i5} – Operating return on assets turnover
X_{i20} – Market share variability

Financial Synergy

X_{i21} – Operating risk variability

Table 21.11 Test of significance of merger and non-merger firms' operating risk

Firm	December 1999 score	December 2000 score	Difference	D_i	Class	W
Panel A: merger firms						
JianHong	5250.000	−8.178	5258.178	5258.178	5	W (+) =10
YuanDaJinHua	42.250	−29.701	71.951	71.951	2	W (−) = 5
YuanFu	10.324	14.973	−4.649	4.649	1	W = 5
FuBang	1017.647	−17.895	1035.542	1035.542	3	
BaoLai	−93.671	1010.000	−1103.671	1103.671	4	
Panel B: non-merge firms						
TaiYu	1274.074	−205.391	1479.465	1479.465	11	W (+) = 70
DaHua	127.941	−54.839	182.780	182.780	7	W (−) = 21
QunYi	1657.895	1.198	1656.697	1656.697	12	W = 21
ZhongXin	59.877	−3.089	62.965	62.965	3	
YongChang	−249.479	−216.376	−33.103	33.103	1	
TaiZhen	−230.769	−96.078	−134.691	134.691	6	
JinDing	−101.587	−7566.667	7465.079	7465.079	13	
RiShen	−29.612	67.241	−96.853	96.853	4	
DaXin	−58.108	127.016	−185.124	185.124	8	
KangHe	228.378	−72.840	301.218	301.218	9	
YaZhou	−50.000	0.000	−50.000	50.000	2	
XinBao	413.333	−292.208	705.541	705.541	10	
TongYi	64.151	−37.701	101.852	101.852	5	

Wilcoxon sign rank (double end) test indicates that ratio difference is statistically *insignificant*, Eq. 21.6

Appendix 2

Principal Component Factor Analysis of Merger Synergies

Principal component factor analysis is used to examine merger synergies. It is a multivariable statistic method focusing on the relationship between groups of variables. Its purpose is to express the original data structure with fewer factors while keeping most of the information provided by the original data structure.

Factor analysis is composed of two parts: one is common factor, and the other is specific factor. Factor analysis intends to group the variables with same common factors. In other words, it discusses how to break down every variable X_i of P variables $X_1 \sim X_P$ into the linear combination of q common factors fj(q, and $q \leq p$), $j = 1, 2 \ldots q$ and specific factor ε_i. The model is as follows:

$$\begin{aligned} X_1 &= \mu_1 + L_{11}F_1 + L_{12}F_2 + \cdots + L_{1q}F_q + \varepsilon_1 \\ X_2 &= \mu_2 + L_{21}F_1 + L_{22}F_2 + \cdots + L_{2q}F_q + \varepsilon_2 \\ &\vdots \\ X_p &= \mu_p + L_{p1}F_1 + L_{p2}F_2 + \cdots + L_{pq}F_q + \varepsilon_p \end{aligned} \tag{21.1}$$

where $F_1 \ldots, F_q$ are common factors, ε_i is specific factor for variable X_i, and L_{ij} is the factor loading of variable X_i and common factor $F_{j.}$ Factors extracted are independent and also analysis preserves the information in the original variables. There is no overlap of information among principle components. The model should be parsimonious, in the sense that a few principle components should be able to replace the original set of variables.

Variability Percentage Adjustment

All principle components with factor loading greater than 1 are selected. Then weights are assigned to different variables based on percentage of variability explained. The variability percentage adjustment is adjusted as follows:

$$V_j^{adj} = \mathrm{V_j}/\Sigma_j V_j \tag{21.2}$$

where V_j ($j = 1,2\ldots q$) is the percentage of variability explained by factor F_j.

Factor Loading Adjustment

The adjusted variability percentage V_j^{adj} is multiplied by factor loading L_{ij} and summed over all variables ($j = 1,2\ldots q$) to get the total loading fo variable X_i. Then adjust L_i such that total factor loadings add to 100.

$$\begin{aligned} L_i &= \sum V_j^{adj} L_{ij} \\ L_i^{adj} &= 100^* \left(\mathrm{L_i} / \sum \mathrm{L_i}\right) \end{aligned} \tag{21.3}$$

Every variable has a weight, L_i^{adj} before and after the estimation period; we average it to obtain weight W_i. For some variables, X_i, the greater the value of the variable, the better it is for the operating, financial, or merger synergy. Those variables are classified as positive variables. If opposite is true, then those variables are classified as negative variable. Hence adjusted variable weights are redesigned to correctly reflect operating synergy score: $W_i^* = \{W_i \text{ or } -W_i\}$ for positive and negative variables, respectively.

Performance Scores

All variables are not measured in the same unit, so they are standardized as $X_i^* = \{X_i - \mathrm{Ave}(X_i)\}/\sigma$, where $\mathrm{Ave}(X_i)$ is the average and σ is the standard deviation of variable X_i ($i = 1,2\ldots p$). The adjusted variable weight W_i^* multiplied by standardized variable Z_i gives the performance score of variable X_i. Appropriate standardized performance score of the variable is used to rank the firm for its operating or financial or market synergy. The sum of the standardized performance scores over the set of variables gives the comprehensive performance score:

$$S = \sum W_i{}^{*} \cdot Z_i \tag{21.4}$$

where the summation is over the first 19 variables listed in Appendix 1. Then firms are ranked in terms of their respective comprehensive performance score. The greater the total score is, the better is the comprehensive performance rating. On the other hand, the smaller the total score, the worse is the performance rating and lower is its rank.

Appendix 3

Wilcoxon Sign Rank Test

Usual "*t*-test" method is unsuitable to examine the significance of change in operating, financial, or market performance before and after merger as we deal with ranks. We use the Wilcoxon sign rank test which checks whether two sets of ranks, pre- and post-merger, come from the same sample or two samples. Let D be the difference between observed value from the sample and reference value. Then we delete those observations whose value of D is zero and rank incrementally the rest of observations in terms of the absolute value of D. If two or more absolute values are the same, we give each value an appropriate rank, then average those ranks, and use the averaged rank as the rank of the same absolute values.

The statistic analysis method of test is as follows:

The differences between the post-merger synergy ranks (X_i) and corresponding premerger synergy ranks (Y_i), D_i is ranked in descending order. Let R_i be the serial number of D_i (if they have the same rank, then take their average value).

$$\begin{gathered}
\boldsymbol{D_i = X_i - Y_i \qquad i = 1, 2, \cdots n} \\
\boldsymbol{R_i = rank(|D_i|) \quad i = 1, 2, \cdots n} \\
\boldsymbol{W(+) = \sum R_i \quad X_i - Y_i > 0} \\
\boldsymbol{W(-) = \sum R_i \quad X_i - Y_i < 0} \\
\boldsymbol{W = \min(W+, W-)}
\end{gathered} \tag{21.5}$$

where W(+) is the total of absolute value of serial numbers of positive rank changes and W(−) is the total of absolute value serial numbers of negative rank changes. W is the test statistic.

Test Hypotheses

For comprehensive performance ranks:

$H_{0:}$ *Performance rating after merger* = *performance rating before merger* ($\eta_1 = \eta_2$).

$H_{1:}$ *Performance rating after merger* $\neq$ *performance rating before merger* ($\eta_1 \neq \eta_2$).

For operating, market, and financial synergy:

H_0 = *No synergy occurred after merger* ($\eta_1 = \eta_2$).

H_1 = *Synergy occurred after merger* ($\eta_1 \neq \eta_2$).

We undertake a double end test. Under the significant level of α, we find the critical value $W(\alpha)$ from the appropriate statistical table. The null hypothesis is rejected if

$$\begin{aligned} &W \leq (-W(\alpha/2)) \text{ or} \\ &W \geq (W(\alpha/2)) \end{aligned} \tag{21.6}$$

This means that merger has produced significant synergy and performance (and credit) rating has affected. The appropriate variables for operating, market, and financial variables are operating cost ratio, ratio of operating income to total assets and market share, and variability of operating risk, respectively.

References

Ansoff, H. I., Brandenburg, R. G., Portner, F. E., & Radosevich, R. (1971). *Acquisition behavior of U.S. manufacturing firms* (pp. 1946–1965). Nashville: Vanderbilt University Press.

Beattie, D. L. (1980). Conglomerate diversification and performance: A survey and applied economics. *Applied Economics, 12*, 251–273.

Brigham, E. F., & Daves, P. (2007). *Intermediate financial management*. Mason: Thomson South West.

Copeland, T. E., & Weston, J. F. (1979). *Financial theory and corporate policy*. Reading: Addison-Wesley.

Fluck, Z., & Lynch, A. W. (1999). Why do firms merge and then divest? A theory of financial synergy. *Journal of Business, 72*, 319–346.

Gardner, D. C. (1996). *Introduction to mergers and acquisitions*. London: Pitman Publishing.

Healy, P. M., Palepu, K. G., & Ruback, R. S. (1992). Does corporate performance improved after mergers? *Journal of Financial Economics, 31*, 135–175.

Higgins, R. C., & Schall, L. D. (1975). Corporate bankruptcy and conglomerate merger. *The Journal of Finance, 30*, 93–119.

Hoshino, Y. (1982). The performance of corporate mergers in Japan. *Journal of Business Financial and Accounting, 9*, 145–163.

Johnson, R., & Wichern, D. (2007). *Applied multivariate statistical analysis* (6th ed.). Upper Saddle River: Pearson Education.

Levy, H., & Sarnat, M. (1970). Diversification, portfolio analysis and the uneasy case for conglomerate mergers. *The Journal of Finance, 25*, 795–802.

Lewellen, W. G. (1971). A pure financial rationable for conglomerate merger. *The Journal of Finance, 26*, 521–537.

Muller, D. C. (1985). Mergers and market share. *The Review of Economics and Statistics, 67*, 66–74.

Rhodes, S. A. (1983). *Power, empire building, and mergers*. Lexington: Lexington Books.

Scherer, F. M. (1980). *Industrial market structure and economic performance*. Boston: Houghton Mifflin.

Seth, A. (1990). Value creation in acquisitions: A re-examination of performance issues. *Strategic Management Journal, 11*, 90–115.

Singh, P., & Montegomery, T. (1987). Using simulated mergers to evaluate corporate diversification strategies. *Strategic Management Journal, 8*, 377–386.

Standard & Poor's Corporation. (1996). *Standard & Poor's corporation ratings criteria*. New York: McGraw Hill.

Weston, J. F., & Chung, K. S. (1983). Do merger make money? A reach summary. *Mergers & Acquisitions, 18*, 40–48.

Weston, J. F., & Mansinghka, S. K. (1971). Test of the efficiency: Performance of conglomerate firms. *The Journal of Finance, 26*, 919–939.

Williamson, O. E. (1981). The modern corporation: Origins, evolution, attributes. *Journal of Economic Literature, 14*, 1537–1568.

Woolridge, J. (2009). *Introductory econometrics: A modern approach* (4th ed.). Mason: Cengage Learning.

On-/Off-the-Run Yield Spread Puzzle: Evidence from Chinese Treasury Markets 22

Rong Chen, Hai Lin, and Qianni Yuan

Contents

Abstract

In this chapter, we document a negative on-/off-the-run yield spread in Chinese Treasury markets. This is in contrast with a positive on-/off-the-run yield spread in most other countries and could be called an "on-/off-the-run yield spread puzzle." To explain this puzzle, we introduce a latent factor in the pricing of Chinese off-the-run government bonds and use this factor to model the yield

We thank the National Natural Science Foundation of China (Grant No. 71101121 and No. 70971114) for financial support.

R. Chen (✉) • Q. Yuan
Department of Finance, Xiamen University, Xiamen, China
e-mail: aronge@xmu.edu.cn; qnyuan@gmail.com

H. Lin
School of Economics and Finance, Victoria University of Wellington, Wellington, New Zealand
e-mail: hai.lin@vuw.ac.nz; cfc1080@gmail.com

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_22,

difference between Chinese on-the-run and off-the-run issues. We use the nonlinear Kalman filter approach to estimate the model. Regressions results suggest that liquidity difference, market-wide liquidity condition, and disposition effect (unwillingness to sell old bonds) could help explain the dynamics of a latent factor in Chinese Treasury markets. The empirical results of this chapter show evidence of phenomena that are quite specific in emerging markets such as China.

The Kalman filter is a mathematical method named after Rudolf E. Kalman. It is a set of mathematical equations that provides an efficient computational (recursive) means to estimate the state of a process, in a way that minimizes the mean of the squared error. The nonlinear Kalman filter is the nonlinear version of the Kalman filter which linearizes about the current mean and covariance. The filter is very powerful in several aspects: it supports estimations of past, present, and even future states, and it can do so even when the precise nature of the modeled system is unknown.

Keywords
On-/off-the-run yield spread • Liquidity • Disposition effect • CIR model • Nonlinear Kalman filter • Quasi-maximum likelihood

22.1 Introduction

It is well known that there exists an on-the-run phenomenon in worldwide Treasury markets. This phenomenon refers to the fact that just-issued (on-the-run or new) government bonds of a certain maturity are generally traded at a higher price or lower yield than previously issued (off-the-run or old) government bonds maturing on similar dates. For example, Amihud and Mendelson (1991), Warga (1992), Kamara (1994), Furfine and Remolona (2002), Goldreich et al. (2005), and Pasquariello and Vega (2009) report the existence of positive on-/off-the-run yield spread in the US Treasury market with different frequency data. Mason (1987) and Boudoukh and Whitelaw (1991, 1993) provide similar evidence in Japan. In spite of different opinions on the information content of the yield spread, there is no disagreement in the literature that the on-/off-the-run yield spread in Treasury markets is significantly positive.

Academics have proposed many theories to explain the positive on-/off-the-run yield spread. Early studies directly attribute this spread to the liquidity difference between new bonds and old bonds (Amihud and Mendelson 1991; Warga 1992; Kamara 1994). More recent work provides some other possible explanations, such as different tax treatment (Strebulaev 2002), specialness in the repo market[1] (Krishnamurthy 2002), the value of future liquidity (Goldreich et al. 2005), search

[1]Specialness refers to the phenomenon that loans collateralized by on-the-run bonds offer lower interest rates than their off-the-run counterparts in repo markets.

costs (Vayanos and Weill 2008), and market frictions of information heterogeneity and imperfect competition among informed traders (Pasquariello and Vega 2009). No matter what arguments are proposed, however, the important role that liquidity plays in the positive on-/off-the-run yield spread and the liquidity premium hypothesis (Amihud and Mendelson 1986) has never been denied. It is widely accepted, by both practitioners and academics, that off-the-run bonds with a lower liquidity level tend to have a higher yield than otherwise similar, yet more liquid, on-the-run bonds.

In this chapter, we document a negative on-/off-the-run yield spread in Chinese Treasury markets, which is contrary to the usual on-the-run phenomenon in other countries and could be called the on-/off-the-run yield spread puzzle in China. Guo and Wu (2006) and Li and He (2008) report that on-/off-the-run yield spread in Chinese Treasury markets is significantly positive, but they did not match the on-the-run and off-the-run bonds correctly. For example, they compare the yield of a just-issued 7-year government bond with that of a previously issued 7-year government bond that has a different maturity. This is not consistent with the calculation of the usual on-/off-the-run yield spread. In order to calculate the spread correctly, we need to match the bonds in terms of maturity date. For example, we compare a just-issued 1-year government bond and a previously issued government bond maturing on similar dates. The maturities are both about 1 year and the durations are close to each other.

To explain on-/off-the-run yield spread puzzle, we introduce a latent factor in the pricing of Chinese off-the-run bonds. This latent factor is used to model the yield difference between on-the-run bonds and off-the-run bonds. We employ a nonlinear Kalman filter to estimate the model and examine the temporal properties of the latent factor. We find that the liquidity premium hypothesis still holds in Chinese Treasury markets. In particular, the change of the latent factor is positively related to the liquidity difference between off-the-run and on-the-run bonds and positively related to the market-wide liquidity condition. Both findings are consistent with the liquidity premium hypothesis. On the other hand, disposition effect (unwillingness to sell old bonds in bear markets) dramatically changes the sign of the yield spread and causes the puzzle. The change of latent factor in Chinese Treasury markets is negatively related to 7-day repo rates. When interest rates go up and the returns of the bond markets are negative, the holders of off-the-run bonds are reluctant to realize loss and will not sell their bonds, which consequently leads to a relatively low yield level of off-the-run bonds and a negative on-/off-the-run yield spread.

Our article makes several contributions to the literature. We document a negative on-/off-the-run yield spread in China. We introduce a latent factor to explain the yield difference between on-the-run bonds and off-the-run bonds and employ the nonlinear Kalman filter in estimation. Our basic ideas are in line with Longstaff et al. (2005) and Lin et al. (2011). We also provide evidence of irrational investor behavior that is quite specific in emerging Treasury markets such as China. In China, the liquidity premium hypothesis still holds, whereas the existence of disposition effect causes the puzzle.

The chapter is organized as follows. In Sect. 22.2, we present a pricing model of government bonds that introduces a latent factor for the off-the-run bonds.

In Sect. 22.3, we describe the data, report the estimation results, and perform a variety of regression analyses. We present our conclusions in Sect. 22.4.

22.2 Bond Pricing Models

22.2.1 On-the-Run Bond Pricing Model

We use the Cox et al. (1985, CIR) model to price the Chinese on-the-run government bonds. The CIR model has been a benchmark interest rate model because of its analytical tractability and other good properties. In this model, the risk-free short rate, r_t, is assumed to follow a square-root process as

$$dr_t = \kappa(\theta - r_t)dt + \sigma_r\sqrt{r_t}dW_{r,t}, \tag{22.1}$$

under the risk-neutral measure Q. κ is the speed of mean reversion, θ is the long-term mean value, σ_r is the volatility parameter of r_t, and W_r denotes a standard Brownian motion under Q. Such specification allows for both mean reversion and conditional heteroskedasticity and guarantees that interest rates are nonnegative.

At time t, the price of an on-the-run government bond maturing at t_M could be written as

$$P_t^{on} = E_t^Q\left[\sum_{m=1}^{M} C_m \exp\left(-\int_t^{t_m} r_s ds\right)\right], \tag{22.2}$$

where P_t^{on} is the on-the-run bond price, C_m is the cash flow payments at time t_m, and M is the total number of cash flow payments. That is, the price of a government bond is the expected present value of the cash flow payments under the risk-neutral measure.

Solving Eq. 22.2 gives

$$P_t^{on} = \sum_{m=1}^{M} C_m A_{m,t} \exp(-B_{m,t} r_t), \tag{22.3}$$

where

$$A_{m,t} = \left[\frac{2h\exp\{(\kappa+h)(t_m-t)/2\}}{2h+(\kappa+h)(\exp\{(t_m-t)h\}-1)}\right]^{2\kappa\theta/\sigma_r^2},$$

$$B_{m,t} = \frac{2(\exp\{(t_m-t)h\}-1)}{2h+(\kappa+h)(\exp\{(t_m-t)h\}-1)},$$

and

$$h = \sqrt{\kappa^2 + 2\sigma_r^2}.$$

22.2.2 Off-the-Run Bond Pricing Model

In order to model the on-/off-the-run yield spread in Chinese Treasury markets, we next incorporate a latent component, l, into the pricing model of Chinese off-the-run government bonds and extend (22.2) to

$$P_t^{off} = E_t^Q\left[\sum_{m=1}^{M} C_m \exp\left(-\int_t^{t_m} (r_s + l_s)ds\right)\right], \tag{22.4}$$

where P_t^{off} is the off-the-run bond price. Similar to Longstaff et al. (2005) and Lin et al. (2011), we assume that under the risk-neutral measure Q,

$$dl_t = \sigma_l dW_{l,t}, \tag{22.5}$$

where W_l is a standard Brownian motion independent of W_r under Q and σ_l is the volatility parameter.

Given the stochastic processes in (22.1) and (22.5), we can obtain the analytical solution for the pricing formula of (22.4),

$$P_t^{off} = \sum_{m=1}^{M} C_m A_{m,t} \exp\left(D_{m,t} - B_{m,t} r_t - (t_m - t)l_t\right), \tag{22.6}$$

where $D_{m,t} = \frac{\sigma_l^2 (t_m - t)}{6}$ and other notations are the same as in (22.3).

22.3 Data and Empirical Estimation

22.3.1 Data Summary

We use the price of Chinese government bonds in the interbank market to estimate the pricing model. The data are from the RESSET dataset. There are two main bond markets in China. One is the interbank bond market, while the other is the exchange bond market. The interbank bond market is a quote-driven over-the-counter market, and the participants are mainly institutional investors. Its outstanding value and trading volume account for over 90 % of Chinese bond markets. The number of bonds traded in the interbank market is at least three times that traded in the exchange market. This is very important for our empirical study, since we need enough bonds to match new and old ones.

Table 22.1 Summary statistics

	Off-the-run		On-the-run		
	Mean	Std	Mean	Std	Difference
One year					
Yield (%)	2.24	0.92	2.31	0.72	−0.07
Modified duration (years)	0.68	0.25	0.73	0.22	−0.05
Age (years)	4.3	2.92	0.26	0.22	4.04[a]
Coupon (%)	2.95	2.73	1.39	1.46	1.63[a]
Three years					
Yield (%)	2.72	0.78	2.83	0.79	−0.11[b]
Modified duration (years)	2.47	0.33	2.48	0.31	−0.01
Age (years)	3.05	1.45	0.38	0.32	2.67[a]
Coupon (%)	3.39	2.32	2.89	0.62	0.50

This table reports the summary statistics of the Chinese 1-year and 3-year on-the-run and off-the-run government bonds between December 2003 and February 2009. The coupon rate and yield are in percentages, while age and modified duration are in years. This table also reports the difference between the off-the-run bonds and the on-the-run bonds

[a]and[b] indicate statistical significance at the 5 % and 1 % level, respectively

Our sample period is from December 2003 to February 2009. We use monthly data and choose the actively traded government bonds with 1 year and 3 years to maturity. Thus, for each month in our sample period, a most recently issued 1-year and 3-year government bonds are selected as the on-the-run bonds, and we get another old bond maturing on similar dates to match each on-the-run bond.[2] Altogether, 63 matched pairs of 1-year government bonds and 63 matched pairs of 3-year government bonds are included in the final sample.

Table 22.1 reports the summary statistics for the sample bonds. As expected, the modified durations of off-the-run and on-the-run bonds are quite close to each other. This means if interest rate risk is the only risk factor, these bonds should be traded at similar yields. To examine whether the same on-the-run phenomenon exists in Chinese Treasury markets, we compute the on-/off-the-run yield spread as

$$\Delta y_{M,t} = y_{M,t}^{off} - y_{M,t}^{on}, \tag{22.7}$$

where $y_{M,t}^{off}$ and $y_{M,t}^{on}$ are the time t yield of the off-the-run bond and the on-the-run bond maturing at t_M, respectively. In our sample, $t_M - t$ is equal to 1 year or 3 years. Figure 22.1 plots the time series of $\Delta y_{M,t}$ of the 1-year bond and 3-year bond.

Table 22.1 also reports the means of on-/off-the-run yield spreads and their statistical significance. Both the 1-year on-/off-the-run yield spread and the 3-year on-/off-the-run yield spread are negative, which is inconsistent with the findings in other markets. Moreover, the 3-year on-/off-the-run yield spread is significantly

[2]The main reason for the data selection comes from the concern of trading activity. Trading in Chinese Treasury markets is not active, especially in the earlier period.

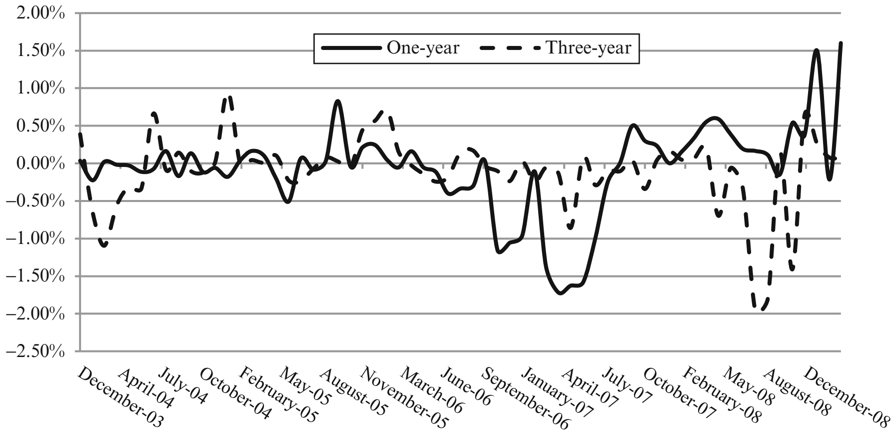


Fig. 22.1 Time series of Chinese on-/off-the-run yield spread. This figure plots the yield difference between the off-the-run and on-the-run Chinese government bonds between December 2003 and February 2009

negative at the 10 % level, which suggests the existence of a significantly negative on-/off-the-run yield spread and remains a puzzle. The 1-year on-/off-the-run yield spread is negative but not significant. However, this spread has some noise, since the difference of coupon rate between the on-the-run issues and off-the-run issues is 1.63 % and significant at the 1 % level. This difference could affect the significance of the on-/off-the-run spread at the 1-year level. Generally speaking, we find evidence of a negative on-/off-the-run spread, which is contrary to the established fact of a positive on-/off-the-run yield spread in most other countries and could be called the on-/off-the-run yield spread puzzle in China.

22.3.2 Empirical Methodology

To explain the on-/off-the-run yield spread puzzle in China, we use the CIR model to price the on-the-run issues and the CIR model with the latent factor to price the off-the-run issues. We first estimate the parameters of the CIR model using all on-the-run bonds. Given the parameters of the CIR process, we further estimate the parameters of the latent factor using off-the-run bonds. Thus, the latent factor represents the yield difference between on-the-run and off-the-run bonds.

In our empirical study, we employ the Kalman filter to estimate the parameters.[3] The standard Kalman filter is not appropriate here because it requires linear state functions and the measurement functions, while Eqs. 22.3 and 22.6 are nonlinear.

[3] See Hamilton (1994) for an explanation of Kalman filter.

Table 22.2 Estimates of pricing models

	On-the-run issues $dr_t = \kappa(\theta - r_t)dt + \sigma_r\sqrt{r_t}dW_{r,t}$		Off-the-run issues $dr_t = \kappa(\theta - r_t)dt + \sigma_r\sqrt{r_t}dW_{r,t}$ $dl_t = \sigma_l dW_{l,t}$	
The state function	$r_t = \gamma + \phi r_{t-1} + \varepsilon_t$		$r_t = \gamma + \phi r_{t-1} + \varepsilon_t l_t = l_{t-1} + \sigma_l(\eta_t - \eta_{t-1})$	
The measurement function	$\mathbf{y}_t^{on} = \boldsymbol{\alpha}_t^{on} + \boldsymbol{\beta}_t^{on} r_t + \boldsymbol{\omega}_t^{on}$		$\mathbf{y}_t^{off} = \boldsymbol{\alpha}_t^{off} + \boldsymbol{\beta}_t^{off} r_t + \boldsymbol{\xi}_t^{off} l_t + \boldsymbol{\omega}_t^{off}$	
κ	0.089(6.403)[a]			
θ	0.023(6.486)[a]			
σ_r	0.063(247.921)[a]			
σ_l			0.009(8.530)[a]	
var(ε_t)	0.000008			
var(η_t)			0.000007	
	One year	Three years	One year	Three years
var($\boldsymbol{\omega}_t^{on}$)	0.000035	0.000020		
var($\boldsymbol{\omega}_t^{off}$)			0.000011	0.0000006
RMSE	0.0045	0.0038	0.0062	0.0058
MAD	0.0596	0.0517	0.0705	0.0672

This table reports the estimate results of pricing models for on-the-run and off-the-run Chinese government bonds. The parameters of the CIR model are estimated from monthly on-the-run bond data. These parameters are then used to estimate the latent factor in the monthly off-the-run bond data. We use a nonlinear Kalman filter approach to estimate the parameters. $\mathbf{y}_t^{on} = \begin{bmatrix} y_{1,t}^{on} \\ y_{3,t}^{on} \end{bmatrix}$, $\mathbf{y}_t^{off} = \begin{bmatrix} y_{1,t}^{off} \\ y_{3,t}^{off} \end{bmatrix}$, $\boldsymbol{\omega}_t^{on} = \begin{bmatrix} \omega_{1,t}^{on} \\ \omega_{3,t}^{on} \end{bmatrix}$, and $\boldsymbol{\omega}_t^{off} = \begin{bmatrix} \omega_{1,t}^{off} \\ \omega_{3,t}^{off} \end{bmatrix}$, where subscript *1* and *3* denote 1-year and 3-year bonds and superscript *on* and *off* denote on-the-run and off-the-run bonds, respectively. The numbers in parentheses are *t* values

[a]indicates statistical significance at the 1 % level respectively. RMSE and MAD are root mean square error and mean absolute deviation, respectively

We therefore use the nonlinear Kalman filter for estimation. The details of the nonlinear Kalman filter with its Matlab codes are reported in the Appendix.

After we estimate the parameters, we then study the dynamic of the latent component and examine its temporal properties to explore the explanations of the on-/off-the-run yield spread puzzle in China.

22.3.3 Estimation Results

22.3.3.1 Estimation Results of On-the-Run Issue

The left-hand column of Table 22.2 reports the estimation results of the CIR model using on-the-run bonds. As shown, all the parameters are significant at the 1 % level. The long-term mean value, the speed of mean reversion, and the volatility parameter of r are 0.023, 0.089, and 0.063, respectively. These results are reasonable and close to the results of other research on dynamic models in the Chinese interest rate (Hong et al. 2010).

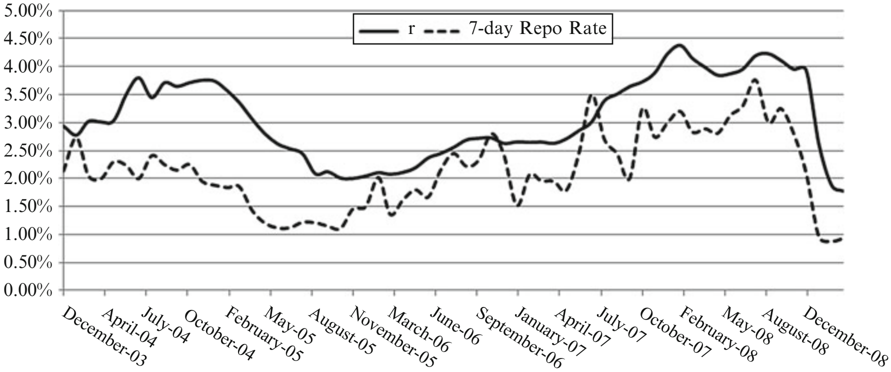


Fig. 22.2 Time series of implied risk-free interest rate and 7-day repo rate in Chinese interbank market. This figure plots the time series of implied risk-free interest rate estimated from the CIR model using Chinese on-the-run government bonds, and the time series of 7-day repo rate in Chinese interbank market

Figure 22.2 plots the time series of r estimated from the model. Most of the time, the value of r is in the interval between 2 % and 4 %. We also plot the time series of the 7-day repo rate in the Chinese interbank market. Similar trends in these two curves suggest that r does capture the change of the market interest rate.

22.3.3.2 Estimation Results of Off-the-Run Issue

With all the parameters obtained for r, we next estimate the parameters of l using the data of off-the-run bonds.[4] The right-hand column of Table 22.2 reports the estimation results. The parameter of σ_l is significant at the 1 % level.

Figure 22.3 plots the time series of l estimated from the data. As we can see from the figure, most of l are negative. We conduct the t-test and find the average of l is significantly negative at the 5 % level (the t-statistic is -017). Since l represents the yield difference between off-the-run bonds and on-the-run bonds, negative l provides further, strong evidence of the on-/off-the-run yield spread puzzle in Chinese Treasury markets.

22.3.4 Regression Analysis

The analysis so far reveals that on average, the off-the-run bonds are traded at a higher price or lower yield than the on-the-run bonds in Chinese Treasury markets, which is hard to explain rationally. We next explore the information contained in this negative yield spread by examining the temporal properties of the latent component, l.

[4]We also estimate the parameters of r and l jointly using the on-the-run and off-the-run data together and find the results are quite similar.

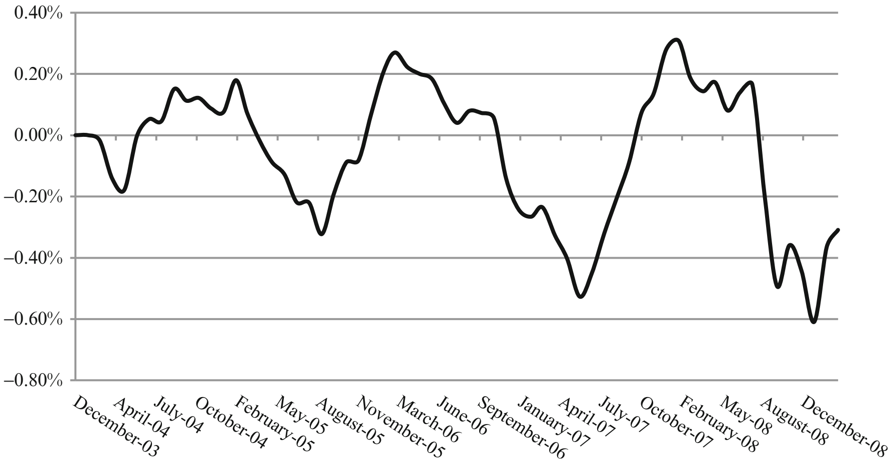


Fig. 22.3 Time series of latent factor. This figure plots the time series of latent factor estimated from the Chinese off-the-run government bonds

In order to explain the temporal properties of the latent component, we introduce several variables. One is the turnover ratio difference between on-the-run issues and off-the-run issues as a measure of liquidity difference, and it is used to examine whether an on-/off-the-run yield spread in Chinese Treasury markets is related to the difference in liquidity conditions. The turnover ratio difference between on-the-run issues and off-the-run issues is defined as

$$TR_t = \frac{\left(TR_{1,t}^{off} - TR_{1,t}^{on}\right) + \left(TR_{3,t}^{off} - TR_{3,t}^{on}\right)}{2} \times 10^{-5},$$

where TR is the turnover ratio, the subscript *1* and *3* denote 1-year bonds and 3-year bonds, and the superscript *on* and *off* denote on-the-run bonds and off-the-run bonds.

Pasquariello and Vega (2007, 2009) find that the release of macroeconomic news changes liquidity, and hence the on-/off-the-run yield spread, in the US Treasury market. For example, when macroeconomic news brings more funds into the bond market, market-wide liquidity conditions will be better, investors might trade old bonds more actively, and the yield difference between off-the-run bonds and on-the-run bonds might decrease, and vice versa. Similarly, we introduce the percentage change of a broad money supply measure, ΔM2, as a proxy of market-wide liquidity conditions to examine whether there is covariation of the latent component with changes in market-wide liquidity conditions. We use one lagged ΔM2 to examine the impact of macroeconomic conditions on Δl_t.

The last factor we investigate is investors' behavior. It is observed that in Chinese bond markets, there exists a "disposition effect." Bond holders are reluctant to realize loss and will not sell old bonds if they have a loss from the investment. Consequently, old bonds might be traded at a lower yield than new bonds. In China, the 7-day repo market is one of the most active bond markets, and

Table 22.3 Time regression results

	(1)	(2)	(3)	(4)
	$\Delta l_t = \beta_0 + \beta_1 TR_t + \varepsilon_t$	$\Delta l_t = \beta_0 + \beta_1 \Delta M2_{t-1} + \varepsilon_t$	$\Delta l_t = \beta_0 + \beta_1 R_{t-1} + \varepsilon_t$	$\Delta l_t = \beta_0 + \beta_1 TR_t + \beta_2 \Delta M2_{t-1} + \beta_3 R_{t-1} + \varepsilon_t$
Intercept	4.330(0.253)	−0.001(−2.349)[a]	0.001(2.348)[a]	0.0005(0.905)
TR_t	0.904(1.625)[b]			−0.231(−0.390)
$\Delta M2_{t-1}$		0.037(2.263)[a]		0.027(1.663)[b]
R_{t-1}			−0.057(−2.703)[c]	−0.045(−1.930)[a]
Adj. R^2	0.043	0.084	0.114	0.163

This table reports the results of regressing the change of latent component, Δl_t, on the on-/off-the-run turnover ratio difference (TR_t), the lagged percentage change of M2 ($\Delta M2_{t-1}$), and the lagged 7-day repo rate (R_{t-1}). The numbers in parentheses are t values

[a],[b], and [c] indicate statistical significance at the 10 %, 5 %, and 1 % level, respectively

Univariate regression of Δl_t on the turnover ratio difference

$\Delta l_t = \beta_0 + \beta_1 TR_t + \varepsilon_t$.

Univariate regression of Δl_t on the lagged percentage change of the money supply:

$\Delta l_t = \beta_0 + \beta_1 \Delta M2_{t-1} + \varepsilon_t$.

Univariate regression of Δl_t on the lagged 7-day repo rate:

$\Delta l_t = \beta_0 + \beta_1 R_{t-1} + \varepsilon_t$.

Multivariate regression of Δl_t on the turnover ratio difference, lagged percentage change of the money supply, and lagged 7-day repo rate:

$\Delta l_t = \beta_0 + \beta_1 TR_t + \beta_2 \Delta M2_{t-1} + \beta_3 R_{t-1} + \varepsilon_t$

the change of the 7-day repo rate is a good measure of market conditions. When the 7-day repo rate goes up, the bond investment will generate a loss for the investors and the disposition effect might occur. We use the interbank 7-day repo rates as the proxy of the market interest rate to investigate whether the on-/off-the-run yield spread puzzle in China is related to this irrational behavior. If the investors are rational and the liquidity premium hypothesis holds, the liquidity of the whole market will decrease in the bear market, and the old bonds will be traded at a higher yield. Thus, detecting the response of the on-/off-the-run yield spread to the change of the 7-day repo rate could help us distinguish whether the disposition effect or the liquidity factor dominates. Similarly, we use one lagged 7-day repo rate in the time series regression.

In what follows, we first run univariate time series regressions of Δl_t against each variable and then conduct multivariate regression analysis against all the three factors. The regression models are specified as follows:

Univariate regression of Δl_t on the turnover ratio difference, TR_t:

$$\Delta l_t = \beta_0 + \beta_1 TR_t + \varepsilon_t.$$

Univariate regression of Δl_t on the lagged percentage change of the money supply, $\Delta M2_{t-1}$:

$$\Delta l_t = \beta_0 + \beta_1 \Delta M2_{t-1} + \varepsilon_t.$$

Univariate regression of Δl_t on the lagged 7-day repo rate, R_{t-1}:

$$\Delta l_t = \beta_0 + \beta_1 R_{t-1} + \varepsilon_t.$$

Multivariate regression of Δl_t on the turnover ratio difference, lagged percentage change of money supply, and lagged 7-day repo rate:

$$\Delta l_t = \beta_0 + \beta_1 TR_t + \beta_2 \Delta M2_{t-1} + \beta_3 R_{t-1} + \varepsilon_t.$$

22.3.4.1 Univariate Regression Analysis

Columns (1), (2), and (3) of Table 22.3 report the results of univariate time series regressions of Δl_t against the turnover ratio difference, the lagged percentage change of the money supply, and the lagged 7-day repo rate, respectively. As shown, all coefficients are significant, indicating that all these factors are useful to explain the change of the on-/off-the-run yield spread in China. It is an interesting finding that is worth further exploring.

The coefficients for TR and the lagged ΔM2 are significantly positive at the 10 % level and the 5 % level, respectively. That means the latent component contains information about liquidity conditions. In particular, since l_t is negative most of the time, the positive sign of coefficients implies that when the explanatory variable increases, the on-/off-the-run yield spread will increase and move close to zero. In other words, there will be less difference between off-the-run bond yields and on-the-run bond yields. The significantly positive coefficient of TR suggests that if the on-/off-the-run turnover ratio difference increases, that is, the liquidity of old bonds becomes better, the yield difference between off-the-run bonds and on-the-run bonds declines, and vice versa. This is consistent with the liquidity premium hypothesis. Similarly, the significantly positive coefficient of lagged ΔM2 reveals that if the money supply increases and market-wide liquidity improves, the yield difference between old bonds and new bonds declines, and vice versa. Both results provide evidence of a liquidity premium in Chinese government bond yields. The liquidity premium hypothesis still holds in Chinese Treasury markets despite the existence of the on-/off-the-run yield spread puzzle.

The lagged 7-day repo rate is significantly negatively related to Δl_t. That is, when the market interest rate goes up and the bond investors have a loss, the yield difference between old bonds and new bonds increases and the on-/off the run spread becomes more negative. This indicates the disposition effect dominates the effect of liquidity. When investors have a loss, the unwillingness to sell old bonds leads to a lower yield for old bonds and a negative on-/off-the-run yield spread. The regression against the lagged 7-day repo rate has the largest adjusted R square, which implies the disposition effect could better explain the change in the on-/off-the-run yield spread in China than the liquidity condition.

22.3.4.2 Multivariate Regression Analysis

Column (4) of Table 22.3 reports the results of multivariate regression. The coefficient of TR is not significant any more, indicating that after controlling for market-wide liquidity conditions and the disposition effect, the liquidity difference has no influence on the on-/off-the-run yield spread. On the other hand, market-wide liquidity and the disposition effect are still significant at the 10 % level.

The adjusted R square of the multivariate is about 16 %. This suggests that the on-/off-the-run spread could be partly explained by the change in market-wide liquidity conditions and the disposition effect.

22.4 Conclusion

In this chapter, we document a negative on-/off-the-run yield spread in Chinese Treasury markets. This is contrary to the positive on-/off-the-run yield spread found in most other countries and could be called the "on-/off-the-run yield spread puzzle." We introduce a latent factor into the pricing formula of off-the-run bonds to capture the yield spread and estimate this factor by the nonlinear Kalman filter. The result confirms the existence of the puzzle.

To reveal the information content of the negative on-/off-the-run yield spread, we perform univariate and multivariate time series regressions of the change of the latent factor against the turnover ratio difference between off-the-run issues and on-the-run issues (a measure of the liquidity difference between off-the-run issues and on-the-run issues), the lagged percentage change of M2 (a measure of market-wide liquidity conditions), and the lagged 7-day repo rates (a measure of disposition effect). We find that the liquidity premium hypothesis still holds in Chinese Treasury markets. The yield spread, however, is dominated by the irrational disposition effect. When the investors have a loss from the bond investment, they are more reluctant to sell old bonds, which leads to a higher price and a lower yield for old bonds and hence causes the puzzle.

Our study is an attempt to explore the coexistence of a standard theoretical hypothesis and irrational behavior in emerging Treasury markets such as China. These markets have been a topic of interest increasingly, as the role of the emerging markets in the global economy becomes more and more important.

Appendix 1: Nonlinear Kalman Filter

Let $y_{M,t}^{on}$ represent the time t yield of an on-the-run government bond maturing at t_M. Equation 22.3 could be written as

$$P_t^{on} = \sum_{m=1}^{M} C_m A_{m,t} \exp\left(-B_{m,t} r_t\right) = \sum_{m=1}^{M} C_m \exp\left(-y_{M,t}^{on}(t_m - t)\right). \tag{22.8}$$

As shown, $y_{M,t}^{on}$ is a nonlinear function of r_t, which is inconsistent with the requirements of the standard Kalman filter that state functions and measurement functions should be linear. So we use the extended (nonlinear) Kalman filter to linearize nonlinear functions. The idea is to employ the Taylor expansions around the estimate at each step. That is, we express $y_{M,t}^{on}$ as

$$y_{M,t}^{on}(r_t) \approx y_{M,t}^{on}\left(\hat{r}_{t|t-1}\right) + \frac{\partial y_{M,t}^{on}}{\partial r_t}\Big|_{r_t=\hat{r}_{t|t-1}} \cdot \left(r_t - \hat{r}_{t|t-1}\right), \tag{22.9}$$

where $\hat{r}_{t|t-1}$ is the estimate of r_t at time t–1.

To get $\frac{\partial y_{M,t}^{on}}{\partial r_t}$, we calculate the first-order derivative of P_t^{on} with respect to r_t,

$$\frac{\partial P_t^{on}}{\partial r_t} = -\sum_{m=1}^{M} C_m A_{m,t} B_{m,t} \exp\left(-B_{m,t} r_t\right) = \frac{\partial P_t^{on}}{\partial y_{M,t}^{on}} \frac{\partial y_{M,t}^{on}}{\partial r_t}. \tag{22.10}$$

Thus, we have

$$\frac{\partial y_{M,t}^{on}}{\partial r_t} = \frac{\sum_{m=1}^{M} C_m A_{m,t} B_{m,t} \exp\left(-B_{m,t} r_t\right)}{\sum_{m=1}^{M} C_m (t_m - t) \exp\left(-y_{M,t}^{on} \cdot (t_m - t)\right)}. \tag{22.11}$$

Given $\hat{r}_{t|t-1}$, we can use Eq. 22.8 to calculate $y_{M,t}^{on}\left(\hat{r}_{t|t-1}\right)$ and then use Eq. 22.11 to get $\frac{\partial y_{M,t}^{on}}{\partial r_t}$

Finally, the linearized measurement model for the on-the-run issues at time t is

$$\mathbf{y_t^{on}} = \boldsymbol{\alpha}_\mathbf{t}^\mathbf{on} + \boldsymbol{\beta}_\mathbf{t}^\mathbf{on} r_t + \boldsymbol{\omega}_\mathbf{t}^\mathbf{on} \tag{22.12}$$

where

$$\mathbf{y_t^{on}} = \begin{bmatrix} y_{1,t}^{on} \\ y_{3,t}^{on} \end{bmatrix},$$

$$\boldsymbol{\alpha}_\mathbf{t}^\mathbf{on} = \begin{bmatrix} y_{1,t}^{on}\left(\hat{r}_{t|t-1}\right) - \hat{r}_{t|t-1} \cdot \frac{\partial y_{1,t}^{on}}{\partial r_t}\Big|_{r_t = \hat{r}_{t|t-1}} \\ y_{3,t}^{on}\left(\hat{r}_{t|t-1}\right) - \hat{r}_{t|t-1} \cdot \frac{\partial y_{3,t}^{on}}{\partial r}\Big|_{r_t = \hat{r}_{t|t-1}} \end{bmatrix},$$

$$\boldsymbol{\beta}_\mathbf{t}^\mathbf{on} = \begin{bmatrix} \frac{\partial y_{1,t}^{on}}{\partial r_t}\Big|_{r_t = \hat{r}_{t|t-1}} \\ \frac{\partial y_{3,t}^{on}}{\partial r_t}\Big|_{r_t = \hat{r}_{t|t-1}} \end{bmatrix},$$

and $\boldsymbol{\omega}_\mathrm{t}^\mathrm{on}$ is the error term,

$$\boldsymbol{\omega}_\mathbf{t}^\mathbf{on} = \begin{bmatrix} \omega_{1,t}^{on} \\ \omega_{3,t}^{on} \end{bmatrix},$$

where the subscript of *1* and *3* represent the 1-year bonds and 3-year bonds, while the superscript *on* refers to on-the-run bonds.

After we get the measurement function, the third step is to rewrite (22.1) as a discrete state function,

$$r_t = \gamma + \phi r_{t-1} + \varepsilon_t, \tag{22.13}$$

where

$$\gamma = \theta(1 - \exp(-\kappa \cdot \Delta t)),$$
$$\phi = \exp(-\kappa \cdot \Delta t),$$

and ε_t is the error term of r_t and Δt is the size of the time interval in the discrete sample. In our study, $\Delta t = 0.0833$. The conditional mean and conditional variance of r_t are

$$\begin{aligned} \hat{r}_{t|t-1} &= \theta(1 - \exp(-k\Delta t)) + \exp(-k\Delta t) \cdot r_{t-1} \\ Var\left(r_{t|t-1}\right) &= \sigma_r^2\left(\frac{1 - \exp(-k\Delta t)}{k}\right)\left(\frac{1}{2}\theta(1 - \exp(-k\Delta t) + \exp(-k\Delta t) \cdot r_{t-1})\right). \end{aligned} \tag{22.14}$$

Similarly, the state functions of the off-the-run issues are

$$\begin{aligned} r_t &= \gamma + \phi r_{t-1} + \varepsilon_t \\ l_t &= l_{t-1} + \sigma_l e_t. \end{aligned} \tag{22.15}$$

The conditional mean and conditional variance of l_t are l_{t-1} and $\sigma_l^2\Delta t$, respectively.

The corresponding measurement function is

$$\mathbf{y_t^{off}} = \boldsymbol{\alpha}_\mathbf{t}^\mathbf{off} + \boldsymbol{\beta}_\mathbf{t}^\mathbf{off} r_t + \boldsymbol{\xi}_\mathbf{t}^\mathbf{off} l_t + \boldsymbol{\omega}_\mathbf{t}^\mathbf{off}, \tag{22.16}$$

where *off* refers to off-the-run bonds,

$$\mathbf{y_t^{off}} = \begin{bmatrix} y_{1,t}^{off} \\ y_{3,t}^{off} \end{bmatrix},$$

$$\boldsymbol{\alpha}_\mathbf{t}^\mathbf{off} = \begin{bmatrix} y_{1,t}^{off}\left(\hat{r}_{t|t-1}, \hat{l}_{t|t-1}\right) - \hat{r}_{t|t-1} \cdot \dfrac{\partial y_{1,t}^{off}}{\partial r_t}\Big|_{r_t=\hat{r}_{t|t-1}} - \hat{l}_{t|t-1} \cdot \dfrac{\partial y_{1,t}^{off}}{\partial l_t}\Big|_{l_t=\hat{l}_{t|t-1}} \\ y_{3,t}^{off}\left(\hat{r}_{t|t-1}, \hat{l}_{t|t-1}\right) - \hat{r}_{t|t-1} \cdot \dfrac{\partial y_{3,t}^{off}}{\partial r_t}\Big|_{r_t=\hat{r}_{t|t-1}} - \hat{l}_{t|t-1} \cdot \dfrac{\partial y_{3,t}^{off}}{\partial l_t}\Big|_{l_t=\hat{l}_{t|t-1}} \end{bmatrix},$$

$$\boldsymbol{\beta}_\mathbf{t}^\mathbf{off} = \begin{bmatrix} \dfrac{\partial y_{1,t}^{off}}{\partial r_t}\Big|_{r_t=\hat{r}_{t|t-1}} \\ \dfrac{\partial y_{3,t}^{off}}{\partial r_t}\Big|_{r_t=\hat{r}_{t|t-1}} \end{bmatrix},$$

$$\boldsymbol{\xi}_\mathbf{t}^\mathbf{off} = \begin{bmatrix} \dfrac{\partial y_{1,t}^{off}}{\partial l_t}\Big|_{l_t=\hat{l}_{t|t-1}} \\ \dfrac{\partial y_{3,t}^{off}}{\partial \mathrm{l}_t}\Big|_{l_t=\hat{l}_{t|t-1}} \end{bmatrix},$$

$$\boldsymbol{\omega}_{\mathbf{t}}^{\mathbf{off}} = \begin{bmatrix} \omega_{1,t}^{off} \\ \omega_{3,t}^{off} \end{bmatrix},$$

$$\frac{\partial y_{M,t}^{off}}{\partial r_t} = \frac{\sum_{m=1}^{M} C_m A_{m,t} B_{m,t} \exp\left(D_{m,t} - B_{m,t} r_t - (t_m - t) l_t\right)}{\sum_{m=1}^{M} C_m (t_m - t) \exp\left(-y_{M,t}^{off} \cdot (t_m - t)\right)},$$

$$\frac{\partial y_{M,t}^{off}}{\partial l_t} = \frac{\sum_{m=1}^{M} C_m A_m (t_m - t) \exp\left(D_{m,t} - B_{m,t} r_t - (t_m - t) l_t\right)}{\sum_{m=1}^{M} C_m (t_m - t) \exp\left(-y_{M,t}^{off} \cdot (t_m - t)\right)},$$

and $\hat{l}_{t|t-1}$ is the estimate of l_t at time t–1.

Once we get the state functions and the measurement functions, we employ the regular iterative prediction-update procedure and the method of quasi-maximum likelihood to estimate the parameters. When estimating the parameters of the off-the-run issues, we use just the parameters γ and ϕ estimated from the on-the-run issues to identify σ_l.

Appendix 2: Matlab Codes

Codes for the On-the-Run Bonds

```
%*************define the likelihood function**************%
  function [logfun v1 zz QQ RR rr]=kalfun(param)
  k=param(1);
  theta=param(2);
  sigm=param(3);
  sigm2=param(4);
  v1=zeros(63,2);
  v=zeros(2,1);
  rr=zeros(63,1);
  zz=zeros(63,2);
  RR=0;
  QQ=0;
  load data.mat
  z1=data(:,1); % YTMs of one-year bonds
  z3=data(:,3); % YTMs of three-year bonds
  c1=data(:,2); % cash flows of one-year bonds
  c3=data(:,4:6); % cash flows of three-year bonds
```

```
  couponrate=data(:,7); % the coupon rate of three-year
bonds
  z=[z1,z3];
  %gam=sqrt(k^2+2*theta^2);
  gam=sqrt(k^2+2*sigm^2);
  tao=[1 2 3]';
  dt=1/12;
  a=zeros(3,1);
  b=zeros(3,1);
  for j=1:3
      a(j)=log((2*gam*exp(k*j/2+gam*j/2))/((k+gam)*(exp
(gam*j)-1)+2*gam)^(2*k*theta/sigm^2));
      b(j)=(2*exp(gam*j)-2)/((k+gam)*(exp(gam*j)-1)
+2*gam);
  end
  r_(1)=theta; %the initial value of r
  A= exp(-k*dt);
  P_=(1/(1-A^2))*
  (sigm^2*(1-exp(-k*dt))/k)*(theta*(1-exp(-k*dt))/2+r_
(1)*exp(-k*dt)); %the initial value of P
  C=theta*(1-exp(-k*dt)); %r(i)=C+A*r(i-1)
  zm=zeros(63,2);     %the prediction of YTM
  R=sigm2*[1 0;0 sqrt(1/3)]; %the covariance of measure-
ment functions
  logfun=0;
  for i=1:63
  Q=(sigm^2*(1-exp(-k*dt))/k)*(theta*(1-exp(-k*dt))/2
+r_(i)*exp(-k*dt)); %the conditional variance of state
functions
  pz1(i)=(c1(i)*b(1)*exp(a(1)-b(1)*r_(i)))/c1(i)*exp
(-z1(i)*1);    %the partial derivative of one year z against r
  pz3(i)=sum(c3(i,:)'.*b.*exp(a-b*r_(i)))/sum(c3(i,:)'.
*tao.*exp(-z3(i)*tao)); %the partial derivative of three
year z against r
  P1=c1(i)*exp(a(1)-b(1)*r_(i));   %the prediction price
of one-year bond
  P3=sum(c3(i,:)'.*exp(a-b*r_(i)));  % the prediction
price of three-year bond
  %zm1=bndyield(P1,c1(i),'20-Jan-1997','20-Jan-
1998',1);
  zm1=-log(P1/c1(i));   %nonlinear measurement function
for one-year bonds
  zm3=bndyield(P3,couponrate(i),'20-Jan-1997','20-Jan-
2000',1); %nonlinear measurement function for three-year
bonds
```

```
  H=[pz1(i) pz3(i)]';
  %C1=[zm1 zm3]'-H*r_(i);
  %zm(i,:)= C1+H*r_(i);
  %zm(i,:)= C1+H*r(i);
  zm(i,:)=[zm1 zm3]';   %the prediction of YTMs
  v=z(i,:)'-zm(i,:)'; %the error of measurement functions
  v1(i,:)=v';      % the error between the prediction and the
real value
  F=H*P_*H'+R;      %the kalman gain
  if det(F)<=0
      logfun=0;
      return
  end
  rr(i,:)=r_(i);
  zz(i,:)=zm(i,:);
  r(i)=r_(i)+P_*H'*inv(F)*v; %update r
  P=P_-P_*H'*inv(F)*H*P_;   %update P
  ll=-0.5*log(det(F))-0.5*v'*inv(F)*v;      %likelihood
function
  logfun=logfun+ll;
    r_(i+1)=A*r(i)+C;   %predict r
    P_=A*P*A'+Q;        %predict P
  end
  QQ=Q;
  RR=R;
  logfun=-logfun;
  function covv=covirance(param)
  covv=zeros(4,4);
  for i=1:4
    for j=1:4
    parama=param;
    paramb=param;
    paramab=param;
    parama(i)=param(i)*1.01;
    paramb(j)=param(j)*0.99;
    paramab(i)=param(i)*1.01;
    paramab(j)=paramab(j)*0.99;
    ua=kalfun(parama);
    db=kalfun(paramb);
    udab=kalfun(paramab);
    kk=kalfun(param);
          covv(i,j)=(ua+db-kk-udab)/((0.01*param(i))*
(0.01*param(j)));
  end
  end
```

Codes for the Off-the-Run Bonds

```
%************define the likelihood function *************%
  function [logfun LL zz v1 QQ RR]=kalfunL(paramL)
  sigm3=paramL(1);
  sigm4=paramL(2);
  load dataL.mat
  z1=dataL(:,1); %YTMs of one-year bonds
  z3=dataL(:,3); % YTMs of three-year bonds
  c1=dataL(:,2); % cash flows of one-year bonds
  c3=dataL(:,4:6); % cash flows of three-year bonds
  couponrate=dataL(:,7); %the coupon rate of three-year
bonds
  r=dataL(:,8); %the estimated r in the CIR model
  % the estimated parameters in the CIR model
  k=0.08899;
  theta=0.022659;
  sigm=0.063329;
  gam=sqrt(k^2+2*sigm^2);
  z=[z1,z3];
  tao=[1 2 3]';
  dt=1/12;
  a=zeros(3,1);
  b=zeros(3,1);
  e=zeros(3,1);
  zz=zeros(63,2);
  v1=zeros(63,2);
  v=zeros(2,1);
  RR=0;
  QQ=0;
  for j=1:3
      a(j)=log((2*gam*exp(k*j/2+gam*j/2))/((k+gam)*(exp
(gam*j)-1)+2*gam)^(2*k*theta/sigm^2));
          b(j)=(2*exp(gam*j)-2)/((k+gam)*(exp(gam*j)-1)
+2*gam);
    e(j)=(sigm3^2*tao(j)^3)/6;
  end
  L_(1)=0; %the initial value of L
  v1=zeros(63,2);
  v=zeros(2,1);
  P_=0;
  zm=zeros(63,2);     %the prediction of YTM
  R=sigm4*[1 0;0 sqrt(1/3)];  %the covariance of measure-
ment functions
  logfun=0;
```

```
  for i=1:63
  Q=sigm3^2*dt;       %the conditional variance of state
functions
  pz1(i)=(c1(i)*exp(a(1)-b(1)*r(i)+e(1)-L_(i)*1))/c1(i)
*exp(-z1(i)*1);%the partial derivative of one year z against r
  pz3(i)=sum(c3(i,:)'.*tao.*exp(a-b*r(i)+e-L_(i)
*tao))/sum(c3(i,:)'.*tao.*exp(-z3(i)*tao));    %the par-
tial derivative of three year z against r
  P1=c1(i)*exp(a(1)-b(1)*r(i)+e(1)-L_(i)*1);   %the pre-
diction price of one-year bond
  P3=sum(c3(i,:)'.*exp(a-b*r(i)+e-L_(i)*tao));     % the
prediction price of three-year bond
  %zm1=bndyield(P1,c1(i),'20-Jan-1997','20-Jan-
1998',1);
  zm1=-log(P1/c1(i));    %nonlinear measurement function
for one-year bonds
  zm3=bndyield(P3,couponrate(i),'20-Jan-1997','20-Jan-
2000',1);  %nonlinear measurement function for three-year
bonds
  H=[pz1(i) pz3(i)]';
  %C1=[zm1 zm3]'-H*r_(i);
  %zm(i,:)= C1+H*r_(i);
  %zm(i,:)= C1+H*r(i);
  zm(i,:)=[zm1 zm3]';   %the prediction of YTMs
  v=z(i,:)'-zm(i,:)';        %the  error  of  measurement
functions
  v1(i,:)=v';       % the error between the prediction and
the real value
  F=H*P_*H'+R;      %the kalman gain
  if det(F)<=0
       logfun=0;
       return
  end
  LL(i,:)=L_(i);
  zz(i,:)=zm(i,:);
  L(i)=L_(i)+P_*H'*inv(F)*v;  %update r
  P=P_-P_*H'*inv(F)*H*P_;     %update P
  ll=-0.5*log(det(F))-0.5*v'*inv(F)*v;
    logfun=logfun+ll;
    L_(i+1)=L(i);   %predict r
     P_=P+Q;       %predict P
  end
  QQ=Q;
  RR=R;
  logfun=-logfun;
```

```
  function covvL=coviranceL(paramL)
  covvL=zeros(2,2);
  for i=1:2
    for j=1:2
    parama=paramL;
    paramb=paramL;
    paramab=paramL;
    parama(i)=paramL(i)*1.0000001;
    paramb(j)=paramL(j)*0.9999999;
    paramab(i)=paramL(i)*1.0000001;
    paramab(j)=paramab(j)*0.9999999;
    ua=kalfunL(parama);
    db=kalfunL(paramb);
    udab=kalfunL(paramab);
    kk=kalfunL(paramL);
      covvL(i,j)=(ua+db-kk-udab)/((0.0000001*paramL(i))*
(0.0000001*paramL(j)));
  end
  end
```

References

Amihud, Y., & Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics, 17*, 223–249.

Amihud, Y., & Mendelson, H. (1991). Liquidity, maturity, and the yields on US treasury securities. *The Journal of Finance, 46*, 1411–1425.

Boudoukh, J., & Whitelaw, R. (1991). The benchmark effect in the Japanese government bond market. *The Journal of Fixed Income, 1*, 52–59.

Boudoukh, J., & Whitelaw, R. (1993). Liquidity as a choice variable: A lesson from the Japanese government bond market. *Review of Financial Studies, 6*, 265–292.

Cox, J., Ingersoll, J. E., & Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica, 53*, 385–407.

Furfine, C., & Remolona, E. (2002). What's behind the liquidity spread? On-the-run and off-the-run US treasuries in autumn 1998, *BIS Quarterly Review*, *6*, 51–58.

Goldreich, D., Hanke, B., & Nath, P. (2005). The price of future liquidity: Time-varying liquidity in the US treasury market. *Review of Finance, 9*, 1–32.

Guo, H., & Wu, K. (2006). Liquidity premium analysis of treasury bonds market in Shanghai stock exchange. *Finance and Economics (In Chinese), 217*, 23–29.

Hamilton, J. D. (1994). *Time series analysis*. Princeton: Princeton University Press.

Hong, Y., Lin, H., & Wang, S. (2010). Modeling the dynamics of Chinese spot interest rates. *Journal of Banking and Finance, 34*, 1047–1061.

Kamara, A. (1994). Liquidity, taxes, and short-term treasury yields. *Journal of Financial and Quantitative Analysis, 29*, 403–417.

Krishnamurthy, A. (2002). The bond/old-bond spread. *Journal of Financial Economics, 66*, 463–506.

Li, D., & He, Y. (2008). Study on the price gap between on-the-run and off-the-run treasury securities of Shanghai stock exchange. *Journal of Shanxi Finance and Economics University (In Chinese), 30*, 80–85.

Lin, H., Liu, S., & Wu, C. (2011). Dissecting corporate bond and CDS spreads. *The Journal of Fixed Income, 20*, 7–39.

Longstaff, F., Mithal, S., & Neis, E. (2005). Corporate yield spreads: Default risk or liquidity? New evidence from the credit default swap market. *The Journal of Finance, 60*, 2213–2253.

Mason, R. (1987). The 10-year bond markets. *Credit Suisse First Boston*, *CSFB Research*.

Pasquariello, P., & Vega, C. (2007). Informed and strategic order flow in the bond markets. *Review of Financial Studies, 20*, 1975–2019.

Pasquariello, P., & Vega, C. (2009). The on-the-run liquidity phenomenon. *Journal of Financial Economics, 92*, 1–24.

Strebulaev, I. (2002). *Liquidity and asset pricing: Evidence from the US Treasury securities market* (Working paper). London Business School, SSRN.

Vayanos, D., & Weill, P. (2008). A search-based theory of the on-the-run phenomenon. *The Journal of Finance, 63*, 1361–1398.

Warga, A. (1992). Bond returns, liquidity, and missing data. *Journal of Financial and Quantitative Analysis, 27*, 605–617.

Factor Copula for Defaultable Basket Credit Derivatives

23

Po-Cheng Wu, Lie-Jane Kao, and Cheng-Few Lee

Contents

Abstract

In this article, we consider a factor copula approach for evaluating basket credit derivatives with issuer default risk and demonstrate its application in a basket credit linked note (BCLN). We generate the correlated Gaussian random

P.-C. Wu (✉) • L.-J. Kao
Department of Finance and Banking, Kainan University, Taoyuan, ROC, Taiwan
e-mail: pcwu@mail.knu.edu.tw; ljkao@mail.knu.edu.tw

C.-F. Lee
Department of Finance and Economics, Rutgers Business School, Rutgers, The State University of New Jersey, Piscataway, NJ, USA

Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
e-mail: cflee@business.rutgers.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_23,

numbers by using the Cholesky decomposition, and then the correlated default times can be decided by these random numbers and the reduced-form model. Finally, the fair BCLN coupon rate is obtained by the Monte Carlo simulation. We also discuss the effect of issuer default risk on BCLN. We show that the effect of issuer default risk cannot be accounted for thoroughly by considering the issuer as a new reference entity in the widely used one-factor copula model, in which constant default correlation is often assumed. A different default correlation between the issuer and the reference entities affects the coupon rate greatly and must be taken into account in the pricing model.

Keywords

Factor copula • Issuer default • Default correlation • Reduced-form model • Basket credit derivatives • Cholesky decomposition • Monte Carlo simulation

23.1 Introduction

Structural and reduced-form models are the two main approaches for modeling default risk. The structural model (Merton 1974) defines default as occurring when a firm's asset value falls below its debt. The reduced-form model (Jarrow and Turnbull 1995), also known as the intensity model, views the default event as an unexpected exogenous stochastic event. It estimates the intensity of the default occurrence by using market data.

However, whether by the structural or reduced-form model, obtaining the joint distribution of default times among a set of assets will be very complicated. Li (1999, 2000) first introduces the copula function (Sklar 1959) to simplify the estimation of the joint distribution. Li assumes that the default events of reference entities follow a Poisson process and sets the dependence structure as a Gaussian copula function. Finally, he performs Monte Carlo simulation to obtain the default times. The copula approach is the main approach for multi-name credit derivatives pricing in the last decade. Mashal and Naldi (2003) use the copula approach to analyze how the default probabilities of the protection sellers and buyers affect basket default swap (BDS) spreads. While pricing the single-name credit default swap (CDS) with counterparty risk based on the continuous-time Markov model, Walker (2006) indicates that using a time-dependent correlation coefficient can improve the market-standard Gaussian copula approach. By connecting defaults through a copula function, Brigo and Chourdakis (2009) find that when the counterparty risk is involved, both the default correlation and credit spread volatility impact the contingent CDS value.

In the implementation of the Gaussian copula using Monte Carlo simulation, the computational complexity increases with the number of reference entities. Thus the factor copula method, which makes the default event conditional on independent state variables, is introduced to deal with these problems. Andersen et al. (2003) find that one or two factors provide sufficient accuracy for the empirical correlation matrices one encounters in credit basket applications. Hull and White (2004) employ a multifactor copula model to price the *k*th-to-default swap and collateralized debt obligation (CDO). Moreover, Laurent and Gregory (2005) use one-factor Gaussian copula to simplify the dependence structure of reference entities and apply this approach to price BDS and CDO. Wu (2010) develops three alternative approaches to price the basket credit linked note (BCLN) with issuer default risk using only one correlation parameter. Wu et al. (2011) analyze how issuer default risk impacts the BCLN coupon rate by an implied default correlation between the issuer and the reference entities.

On the other hand, acceleration techniques such as the importance sampling method and others are used to improve the simulation efficiency. Chiang et al. (2007) and Chen and Glasserman (2008) apply the Joshi-Kainth algorithm (Joshi and Kainth 2004), and Bastide et al. (2007) use the Stein method (Stein 1972) for the multi-name credit derivative pricing to reduce variance of the simulation results.

This article constructs a factor copula framework to evaluate defaultable basket credit derivatives. The effect of the default correlation between the issuer and the reference entities is considered in the proposed model. Its application in BCLN with issuer default risk is also demonstrated. This study shows that the default correlation between the issuer and the reference entities plays an important role in the decision of fair BCLN coupon rate.

The remainder of this article is organized as follows. Section 23.2 reviews the factor copula model and shows the proposed basket credit derivative pricing model with issuer default event. Subsequently, Section 23.3 introduces the BCLN and demonstrates how to price it when issuer default risk exists. Section 23.4 presents the results of numerical analysis. Conclusions are finally drawn in Section 23.5.

23.2 Factor Copula with Issuer Default Risk

The most widely used copula function is the Gaussian copula and its definition is as follows:

$$C^{Ga}(u_1,\ u_2,\ \ldots,u_N) = \Phi_R\left(\phi^{-1}(u_1),\ \phi^{-1}(u_2),\ \ldots,\phi^{-1}(u_N)\right) \tag{23.1}$$

where Φ_R denotes a multivariate cumulative normal (Gaussian) distribution, R represents the correlation coefficient matrix, and ϕ^{-1} is the inverse function of one dimensional cumulative Gaussian distribution.

Consider a credit portfolio which contains N reference entities, the default times of each reference entity are $\tau_1,\tau_2,\ldots,\tau_N$, respectively. According to the reduced-form model, each reference entity default follows a Poisson process. The cumulative default probability before time t is

$$F_i(t) = P(\tau_i \le t) = 1 - e^{-\lambda_i t},\ i = 1, 2,\ \ldots, N \tag{23.2}$$

where λ_i is the hazard rate of the reference entity i. Because $F_i(t) \sim U(0,1)$, applying the Gaussian copula obtains the multivariate joint distribution of default times, as follows:

$$F(\tau_1, \tau_2, \ldots, \tau_N) = \Phi_R\left(\phi^{-1}(F_1(\tau_1)),\ \phi^{-1}(F_2(\tau_2)),\ \ldots,\ \phi^{-1}(F_N(\tau_N))\right) \tag{23.3}$$

Let X_i represent the Gaussian random variable corresponding to the default time of the reference entity i. In the one-factor copula model, the default time of reference entity i depends on a common factor ε_Y and a firm specific risk factor ε_{xi}. Both ε_Y and ε_{Xi} are independent standard Gaussian variables. The details of one-factor copula model are given in Appendix 1. Thus X_i can be created via Cholesky decomposition, as follows:

$$X_i = \rho_{X_iY}\varepsilon_Y + \sqrt{1 - \rho_{X_iY}{}^2}\varepsilon_{X_i}, \qquad i = 1, 2, \ldots, N \tag{23.4}$$

where ρ_{XiY} denotes the correlation coefficient between the reference entity X_i and the common factor ε_Y. How to apply the Cholesky decomposition for generating correlated random variables is shown in Appendix 2.

One-factor Gaussian copula model with constant pairwise correlations has become the standard market model. In the standard market model, all ρ_{X_iY} in Eq. 23.4 are equal to ρ, then the constant pairwise correlation $\rho_{X_iX_j}\,(i \neq j)$ will be ρ^2. Let $X_1 = \phi^{-1}(F_1(\tau_1)), X_2 = \phi^{-1}(F_2(\tau_2)), \ldots, X_N = \phi^{-1}(F_N(\tau_N))$, by mapping τ_i and X_i, we can simulate the default time of the reference entity i using the following equation:

$$\tau_i = F_i^{-1}(\phi(X_i)) = \frac{-\ln(1 - \phi(X_i))}{\lambda_i}, \qquad i = 1, 2, \ldots, N \tag{23.5}$$

When issuer default risk is involved in the basket credit derivative, a natural way is to view it as one reference entity of the derivative holder's credit portfolio, as discussed in Wu (2010). Wu (2010) assumes the issuer default time is determined by a Gaussian random variable Z. Like the reference entity variable X_i, Z is decided by the common factor ε_Y and the issuer's specific risk factor ε_Z. Both ε_Y and ε_Z are independent standard Gaussian variables. Because the issuer is viewed as one additional reference name in the portfolio, the correlation coefficient between Z and ε_Y is also ρ. In this approach, Z and X_i are formulated as follows:

$$Z = \rho\varepsilon_Y + \sqrt{1 - \rho^2}\varepsilon_Z \tag{23.6}$$

$$X_i = \rho\varepsilon_Y + \sqrt{1 - \rho^2}\varepsilon_{X_i}, \qquad i = 1, 2, \ldots, N \tag{23.7}$$

In the above approach, the default correlation between the issuer and the reference entities will be fixed to ρ^2, which is always positive. Thus, it is not flexible enough to deal with the default correlation. The default correlation between the issuer and

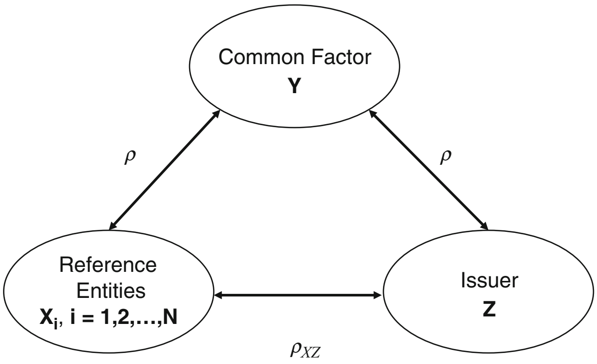


Fig. 23.1 The default correlations between the issuer, reference entities, and the common factor in the proposed model

the reference entities has a different impact on the fair coupon rate, and this needs to be considered in credit derivative pricing.

Suppose that the default correlation between the issuer and the reference entities is ρ_{XZ}. The relationship between the issuer, reference entities, and the common factor in the proposed model is shown in Fig. 23.1. Given that the three random variables ε_Y, ε_Z, and ε_{Xi} are independent of each other, X_i, which is the Gaussian random variable corresponding to the default time of the reference entity i, can be obtained by the Cholesky decomposition. Thus, when the default correlation between the issuer and the reference entities is incorporated into the pricing model, Z and X_i should be formulated as follows:

$$Z = \rho\varepsilon_Y + \sqrt{1-\rho^2}\varepsilon_Z \tag{23.8}$$

$$X_i = \rho\varepsilon_Y + \frac{\rho_{XZ}-\rho^2}{\sqrt{1-\rho^2}}\varepsilon_Z + \sqrt{1-\rho^2-\left(\frac{\rho_{XZ}-\rho^2}{\sqrt{1-\rho^2}}\right)^2}\varepsilon_{X_i} \tag{23.9}$$

To obtain a real number value of X_i, the following criteria must be satisfied.

$$\rho^2 + \left(\frac{\rho_{XZ}-\rho^2}{\sqrt{1-\rho^2}}\right)^2 \leq 1 \tag{23.10}$$

By rearranging the above equation, the criteria can be written as follows:

$$2\rho^2 + \rho_{XZ}^2 - 2\rho^2\rho_{XZ} \leq 1 \tag{23.11}$$

According to the above settings, the correlation coefficient between the reference entity X_i and X_j will be

$$
\begin{aligned}
\rho_{X_i X_j} &= \frac{\mathrm{Cov}\left(X_i, X_j\right)}{\sigma_{X_i}\sigma_{X_j}} \\
&= \mathrm{Cov}\left(\begin{array}{l}
\rho\varepsilon_Y + \dfrac{\rho_{XZ}-\rho^2}{\sqrt{1-\rho^2}}\varepsilon_Z \\
+\sqrt{1-\rho^2-\left(\dfrac{\rho_{XZ}-\rho^2}{\sqrt{1-\rho^2}}\right)^2}\,\varepsilon_{X_i}, \\
\rho\varepsilon_Y + \dfrac{\rho_{XZ}-\rho^2}{\sqrt{1-\rho^2}}\varepsilon_Z \\
+\sqrt{1-\rho^2-\left(\dfrac{\rho_{XZ}-\rho^2}{\sqrt{1-\rho^2}}\right)^2}\,\varepsilon_{X_j}
\end{array}\right) \\
&= \rho^2 + \frac{\left(\rho_{XZ}-\rho^2\right)^2}{1-\rho^2} \\
&= \frac{\rho^2+\rho_{XZ}^2-2\rho_{XZ}\rho^2}{1-\rho^2}
\end{aligned}
\tag{23.12}
$$

23.3 Pricing a Defaultable Basket Credit Linked Note

Basket credit linked note (BCLN) is a kind of basket credit derivative product. It is a note with a price or coupon linked to credit events of reference entities (obligations). The conventional form of BCLN is the *k*th-to-default BCLN. The BCLN holder (the protection seller) pays the notional principal to the BCLN issuer (the protection buyer) at the start of the contract and receives the coupon payments until either the *k*th default or the contract maturity, whichever occurs earlier. If the *k*th default occurs before contract maturity, the BCLN holder receives the recovered value of the reference entity from the BCLN issuer. Otherwise, the BCLN holder receives the notional principal back on contract maturity. In derivative markets, the issuer default risk is attracting considerable attention because of the recent financial turmoil and collapses of large financial institutions. If the BCLN issuer defaults, the BCLN holder will not receive the recovered value of the reference entity as the credit event happens nor the notional amount at the contract maturity. The coupon payments also cease due to the issuer default. Thus the issuer default results in a large loss. Therefore, it is important to incorporate issuer default risk in BCLN pricing to obtain a reasonable coupon rate.

If the issuer default risk is not considered, the value of a *k*th-to-default BCLN with N reference entities, of which the notional principal is one dollar, can be written as follows:

$$BCLN = E^Q \begin{bmatrix} c \times \sum_{i=1}^{T} e^{-rt_i} I(t_i < \tau_k) \\ + \delta_k \times e^{-r\tau_k} \times I(\tau_k \le t_T) \\ + e^{-rt_T} \times I(\tau_k > t_T) \end{bmatrix} \tag{23.13}$$

where the coupon rate is c and is paid annually. The coupon payment dates are $t_i, i = 1, 2, \cdots, T$. The maturity date is t_T. Furthermore, τ_k is the kth default time, and $\tau_1 < \tau_2 < \cdots < \tau_N$. δ_k is the recovery rate of the kth default reference entity. Thus δ_k denotes the recovery value, which the issuer pays to the BCLN holder on the kth default. The discount rate is $r\%$. Finally, Q denotes the risk-neutral probability measure, and $I(.)$ is an indicator function.

Let the above equation equals one, the equation can be rewritten as:

$$\begin{aligned} & c \times E^Q \left[\sum_{i=1}^{T} e^{-rt_i} I(t_i < \tau_k) \right] \\ & = E^Q \begin{bmatrix} 1 - \delta_k \times e^{-r\tau_k} \times I(\tau_k \le t_T) \\ - e^{-rt_T} \times I(\tau_k > t_T) \end{bmatrix} \end{aligned} \tag{23.14}$$

Here we employ the Monte Carlo simulation, which is described in Appendix 3, to obtain the initial fair BCLN coupon rate c as follows:

$$c = \frac{\sum_{s=1}^{W} \begin{bmatrix} 1 - \delta_k^s \times e^{-r\tau_k^s} \times I\left(\tau_k^s \le t_T\right) \\ - e^{-rt_T} \times I\left(\tau_k^s > t_T\right) \end{bmatrix}}{\sum_{s=1}^{W} \left[\sum_{i=1}^{T} e^{-rt_i} I\left(t_i < \tau_k^s\right) \right]} \tag{23.15}$$

where W represents the number of simulation runs. τ_k^s represents the kth default time, and δ_k^s denotes the recovery rate of the kth default reference entity at the sth simulation, respectively.

When the issuer default risk is involved, whether the issuer default occurs before or after the kth default must be taken into account. This article defines $\hat{\tau}$ as the issuer default time and $\hat{\delta}$ as the issuer recovery rate. The BCLN holder gets back the recovered value of the reference obligation if the kth default occurs before both the issuer default time $\hat{\tau}$ and maturity date t_T. If the issuer default occurs before the kth default and maturity date, the issuer will not provide the BCLN holder with the redemption proceeds and stop the coupon payments. In this situation, the notional principal multiplied by the issuer recovery rate is returned to the BCLN holder. To obtain all of the notional principal back, both the kth default time and the issuer default time must be later than the contract maturity date. Thus, the value of a kth-to-default BCLN with issuer default risk is modified as follows:

$$BCLN = E^{Q}\begin{bmatrix} c \times \sum_{i=1}^{T} e^{-rt_i} I(t_i < \min(\tau_k, \hat{\tau})) \\ + \delta_k \times e^{-r\tau_k} \times I(\tau_k < \min(\hat{\tau}, t_T)) \\ + \hat{\delta} \times e^{-r\hat{\tau}} \times I(\hat{\tau} < \min(\tau_k, t_T)) \\ + e^{-rt_T} \times I(t_T < \min(\tau_k, \hat{\tau})) \end{bmatrix} \tag{23.16}$$

Therefore, the fair value of the coupon rate c with issuer default risk is

$$c = \frac{\sum_{s=1}^{W} \begin{bmatrix} 1 - \delta_k^s \times e^{-r\tau_k^s} \times I\left(\tau_k^s < \min(\hat{\tau}^s, t_T)\right) \\ -\hat{\delta} \times e^{-r\hat{\tau}^s} \times I\left(\hat{\tau}^s < \min\left(\tau_k^s, t_T\right)\right) \\ -e^{-rt_T} \times I\left(t_T < \min\left(\tau_k^s, \hat{\tau}^s\right)\right) \end{bmatrix}}{\sum_{s=1}^{W} \left[\sum_{i=1}^{T} e^{-rt_i} I\left(t_i < \min\left(\tau_k^s, \hat{\tau}^s\right)\right)\right]} \tag{23.17}$$

where $\hat{\tau}^s$ represents the issuer default time at the sth simulation.

23.4 Numerical Analysis

The numerical example presented here is a 5-year BCLN with three reference entities; all of which are with notional principal one dollar, hazard rate 5 %, and recovery rate 30 %. Furthermore, the coupon is paid annually; the hazard rate and recovery rate of the issuer are 1 % and 30 %, respectively. Sixty thousand runs of Monte Carlo simulation are executed to calculate the coupon rates, and the results are shown in Tables 23.1, 23.2 and 23.3.

As we can see in Tables 23.1, 23.2 and 23.3, when issuer default risk is considered by viewing it as one reference entity of the credit portfolio (column II), the BCLN coupon rate increases compared to those without issuer default risk (column I) for $k = 1$–3. This is reasonable because the existence of issuer default risk increases the risk of holding a BCLN; thus, the holder will demand a higher coupon rate.

When the default correlations between the issuer and the reference entities are considered as in the proposed model, the BCLN coupon rates with issuer default risk (column III) are greater than those without issuer default risk (column I) for $k = 2$ and 3 in Tables 23.2 and 23.3. However, when $k = 1$ in Table 23.1, most of the BCLN coupon rates with issuer default risk are lower than those without issuer default risk, especially when the issuer and the reference entities are highly negatively or positively correlated. This result shows that the BCLN coupon rates with issuer default risk are not necessarily greater than those without issuer default risk. Moreover, from Figs. 23.2,

Table 23.1 First-to-default BCLN coupon rates without and with issuer default risk. (I) Default free: Issuer default risk is not included in the pricing model. (II) In credit portfolio: The issuer is viewed as one reference entity of the credit portfolio. The default correlations between the issuer and the reference entities are fixed to ρ^2, which is always positive. (III) The proposed model: The default correlation between the issuer and the reference entities is ρ_{XZ}, which may be positive or negative

	(I)	(II)	(III) The proposed model						
						ρ_{XZ}			
ρ	Default free	In credit portfolio	−0.9	−0.6	−0.3	0	0.3	0.6	0.9
−0.9	8.0195 %	8.0346 %	–	–	–	–	–	–	7.6700 %
−0.8	9.2625 %	9.3471 %	–	–	–	–	7.0292 %	9.3593 %	7.8852 %
−0.7	10.2417 %	10.4279 %	–	–	–	7.2895 %	10.2124 %	10.1562 %	7.9327 %
−0.6	11.0469 %	11.3501 %	–	–	–	10.5546 %	11.4061 %	10.5488 %	7.9452 %
−0.5	11.7604 %	12.1928 %	–	–	10.2184 %	12.0213 %	12.0974 %	10.7840 %	7.9650 %
−0.4	12.3336 %	12.8846 %	–	8.8110 %	11.8856 %	12.9609 %	12.5139 %	10.9275 %	7.9665 %
−0.3	12.7890 %	13.4509 %	–	10.6576 %	12.9594 %	13.5479 %	12.7852 %	11.0273 %	7.9659 %
−0.2	13.1304 %	13.8645 %	7.5062 %	11.7035 %	13.5910 %	13.9236 %	12.9839 %	11.0791 %	7.9894 %
−0.1	13.3336 %	14.1181 %	8.7846 %	12.2398 %	13.9263 %	14.1327 %	13.1152 %	11.1038 %	7.9758 %
0	13.3875 %	14.1785 %	9.0738 %	12.3865 %	14.0210 %	14.1785 %	13.1205 %	11.1105 %	7.9933 %
0.1	13.3547 %	14.1215 %	8.7274 %	12.1743 %	13.8962 %	14.1373 %	13.0764 %	11.1036 %	8.0044 %
0.2	13.1236 %	13.8336 %	7.4400 %	11.5964 %	13.4923 %	13.8863 %	12.9678 %	11.0703 %	7.9941 %
0.3	12.7744 %	13.4073 %	–	10.5344 %	12.8035 %	13.4871 %	12.7940 %	10.9995 %	7.9702 %
0.4	12.2933 %	12.8209 %	–	8.6831 %	11.7668 %	12.8436 %	12.5034 %	10.9036 %	7.9704 %
0.5	11.7091 %	12.1252 %	–	–	10.1270 %	11.9387 %	12.0415 %	10.7603 %	7.9715 %
0.6	10.9854 %	11.2806 %	–	–	–	10.4846 %	11.3343 %	10.5102 %	7.9448 %
0.7	10.1941 %	10.3750 %	–	–	–	7.1962 %	10.1661 %	10.1106 %	7.9210 %
0.8	9.2061 %	9.2802 %	–	–	–	–	6.9644 %	9.2933 %	7.8398 %
0.9	7.9788 %	7.9892 %	–	–	–	–	–	–	7.6319 %

Table 23.2 Second-to-default BCLN coupon rates without and with issuer default risk. (I) Default free: Issuer default risk is not included in the pricing model. (II) In credit portfolio: The issuer is viewed as one reference entity of the credit portfolio. The default correlations between the issuer and the reference entities are fixed to ρ^2, which is always positive. (III) The proposed model: The default correlation between the issuer and the reference entities is ρ_{XZ}, which may be positive or negative

	(I)	(II)	(III) The proposed model						
						ρ_{XZ}			
ρ	Default free	In credit portfolio	−0.9	−0.6	−0.3	0	0.3	0.6	0.9
−0.9	5.2169 %	5.2585 %	–	–	–	–	–	–	5.2870 %
−0.8	4.9714 %	5.1490 %	–	–	–	–	5.9481 %	5.1998 %	5.2213 %
−0.7	4.7203 %	5.0406 %	–	–	–	6.1463 %	5.3390 %	4.9617 %	5.2041 %
−0.6	4.4753 %	4.9077 %	–	–	–	5.5463 %	4.9738 %	4.8270 %	5.2135 %
−0.5	4.2476 %	4.7844 %	–	–	5.8264 %	5.1061 %	4.7372 %	4.7592 %	5.1736 %
−0.4	4.0568 %	4.6587 %	–	6.1656 %	5.3827 %	4.8129 %	4.6044 %	4.7055 %	5.1541 %
−0.3	3.8888 %	4.5440 %	–	5.7880 %	5.0664 %	4.6078 %	4.4986 %	4.6566 %	5.1422 %
−0.2	3.7766 %	4.4490 %	6.3028 %	5.5060 %	4.8417 %	4.4584 %	4.4238 %	4.6153 %	5.1375 %
−0.1	3.6917 %	4.3717 %	6.1079 %	5.3332 %	4.7028 %	4.3759 %	4.3711 %	4.6020 %	5.1434 %
0	3.6646 %	4.3412 %	6.0386 %	5.2612 %	4.6477 %	4.3412 %	4.3441 %	4.5867 %	5.1507 %
0.1	3.6833 %	4.3559 %	6.0959 %	5.3033 %	4.6836 %	4.3606 %	4.3478 %	4.5939 %	5.1636 %
0.2	3.7480 %	4.4025 %	6.2475 %	5.4584 %	4.7916 %	4.4244 %	4.3710 %	4.5969 %	5.1732 %
0.3	3.8817 %	4.5070 %	–	5.7235 %	4.9942 %	4.5580 %	4.4430 %	4.6288 %	5.1900 %
0.4	4.0242 %	4.5995 %	–	6.0943 %	5.3165 %	4.7513 %	4.5276 %	4.6679 %	5.1908 %
0.5	4.2155 %	4.7249 %	–	–	5.7533 %	5.0596 %	4.6854 %	4.7051 %	5.2016 %
0.6	4.4092 %	4.8338 %	–	–	–	5.4952 %	4.9002 %	4.7565 %	5.2005 %
0.7	4.6447 %	4.9559 %	–	–	–	6.0639 %	5.2704 %	4.8851 %	5.1929 %
0.8	4.9138 %	5.0873 %	–	–	–	–	5.8734 %	5.1404 %	5.2065 %
0.9	5.1952 %	5.2336 %	–	–	–	–	–	–	5.2602 %

Table 23.3 Third-to-default BCLN coupon rates without and with issuer default risk. (I) Default free: Issuer default risk is not included in the pricing model. (II) In credit portfolio: The issuer is viewed as one reference entity of the credit portfolio. The default correlations between the issuer and the reference entities are fixed to ρ^2, which is always positive. (III) The proposed model: The default correlation between the issuer and the reference entities is ρ_{XZ}, which may be positive or negative

	(I)	(II)	(III) The proposed model						
						ρ_{XZ}			
ρ	Default free	In credit portfolio	−0.9	−0.6	−0.3	0	0.3	0.6	0.9
−0.9	3.5790 %	3.7167 %	–	–	–	–	–	–	3.8150 %
−0.8	3.0115 %	3.3632 %	–	–	–	–	5.1041 %	3.4122 %	3.6937 %
−0.7	2.6429 %	3.1441 %	–	–	–	5.2536 %	3.3862 %	3.1083 %	3.6545 %
−0.6	2.3993 %	2.9853 %	–	–	–	3.4885 %	3.0143 %	2.9859 %	3.6223 %
−0.5	2.2319 %	2.8735 %	–	–	3.7706 %	3.0416 %	2.8644 %	2.9228 %	3.6225 %
−0.4	2.1107 %	2.7835 %	–	4.5560 %	3.2013 %	2.8419 %	2.7752 %	2.8858 %	3.6088 %
−0.3	2.0280 %	2.7177 %	–	3.6782 %	2.9511 %	2.7344 %	2.7243 %	2.8614 %	3.5933 %
−0.2	1.9809 %	2.6597 %	5.3797 %	3.3040 %	2.8003 %	2.6635 %	2.6743 %	2.8273 %	3.5886 %
−0.1	1.9499 %	2.6251 %	4.5085 %	3.1384 %	2.7242 %	2.6247 %	2.6480 %	2.8155 %	3.5889 %
0	1.9359 %	2.6020 %	4.3246 %	3.0819 %	2.6916 %	2.6020 %	2.6360 %	2.8013 %	3.5745 %
0.1	1.9470 %	2.6101 %	4.5012 %	3.1331 %	2.7072 %	2.6095 %	2.6356 %	2.7996 %	3.5754 %
0.2	1.9719 %	2.6320 %	5.3262 %	3.2787 %	2.7677 %	2.6357 %	2.6482 %	2.8099 %	3.5761 %
0.3	2.0171 %	2.6668 %	–	3.6289 %	2.8962 %	2.6780 %	2.6788 %	2.8205 %	3.5859 %
0.4	2.0883 %	2.7214 %	–	4.4910 %	3.1651 %	2.7836 %	2.7186 %	2.8426 %	3.5879 %
0.5	2.2105 %	2.8231 %	–	–	3.7405 %	2.9960 %	2.8098 %	2.8894 %	3.6007 %
0.6	2.3793 %	2.9471 %	–	–	–	3.4615 %	2.9825 %	2.9468 %	3.6257 %
0.7	2.6279 %	3.1099 %	–	–	–	5.2006 %	3.3573 %	3.0784 %	3.6345 %
0.8	2.9783 %	3.3144 %	–	–	–	–	5.0265 %	3.3628 %	3.6818 %
0.9	3.5583 %	3.6858 %	–	–	–	–	–	–	3.7992 %

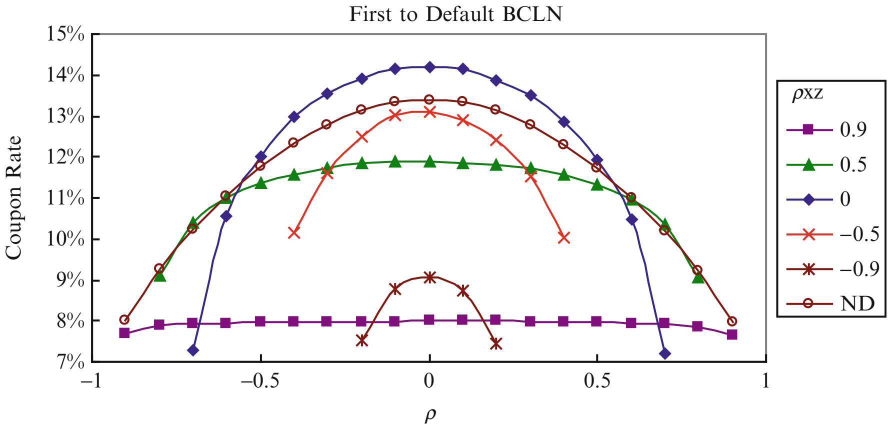


Fig. 23.2 First-to-default BCLN coupon rates under various default correlations between the common factor and reference entities/issuer (ρ)

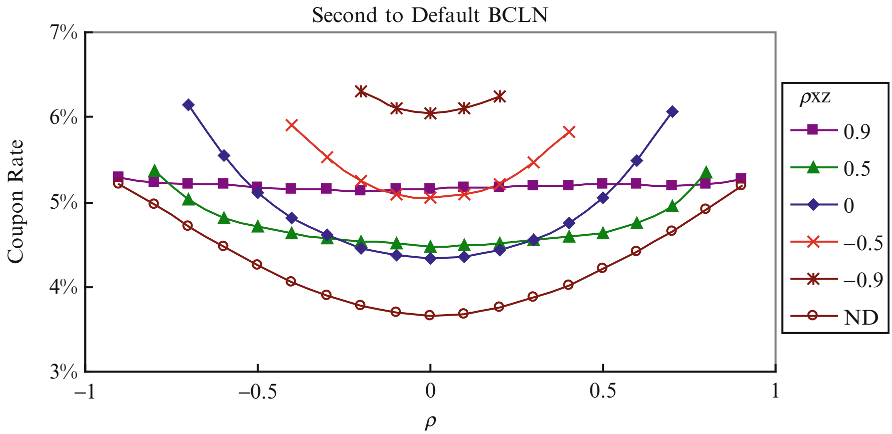


Fig. 23.3 Second-to-default BCLN coupon rates under various default correlations between the common factor and reference entities/issuer (ρ)

23.3 and 23.4, we find that when the correlation between the issuer and the reference entities approaches a strongly positive correlation ($\rho_{XZ} = 0.9$), the BCLN coupon rate curve becomes flatter and less sensitive to the common factor.

23.5 Conclusion

This article applies a factor copula approach for evaluating basket credit derivatives with issuer default risk. The proposed model considers the different effects of

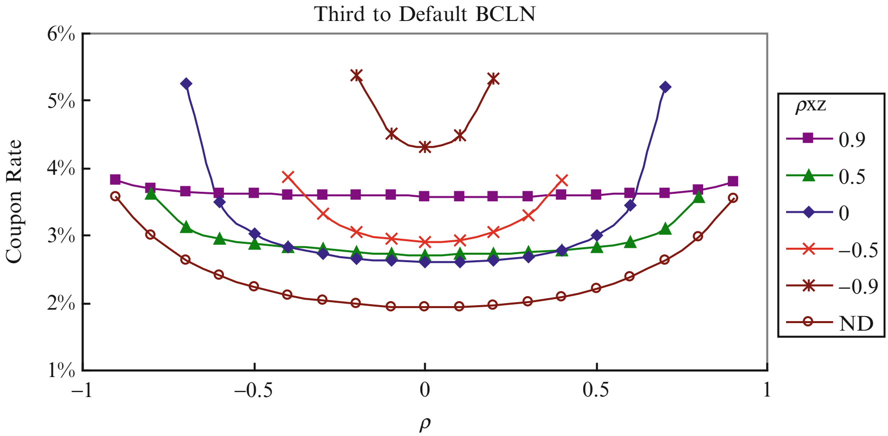


Fig. 23.4 Third-to-default BCLN coupon rates under various default correlations between the common factor and reference entities/issuer (ρ)

the default correlation between the issuer and the reference entities. A numerical example of the proposed model on BCLN is demonstrated and discussed in the article. The example shows that viewing the issuer default as a new reference entity cannot reflect the effect of issuer default risk thoroughly. The different default correlation between the issuer and the reference entities affects the coupon rate greatly and must be taken into account in credit derivative pricing.

Appendix 1: Factor Copula

In a factor copula model, we assume that different variables depend on some common factors. The most widely used model in finance is the one-factor Gaussian copula model.

One-Factor Gaussian Copula Model

Let $S_i(t) = P(\tau_i > t)$ and $F_i(t) = P(\tau_i \le t)$ be the marginal survival and marginal default distributions, respectively. Let Y be the common factor and f its density function. Assume the default times are conditionally independent, given the common factor Y. $q_t^{i|Y} = P(\tau_i > t \mid Y)$ and $p_t^{i|Y} = P(\tau_i \le t \mid Y)$ are the conditional survival and conditional default distributions, respectively. According to the law of iterated expectations, the joint survival and default distribution functions are as follows:

$$\begin{aligned} &S(t_1, t_2, \ldots, t_n) \\ &\quad = P(\tau_i > t_1, \tau_2 > t_2, \ldots, \tau_n > t_n) \\ &\quad = \int \prod_{i=1}^{n} q_t^{i|y} f(y) dy \end{aligned} \tag{23.18}$$

$$\begin{aligned} &F(t_1, t_2, \ldots, t_n) \\ &\quad = P(\tau_i \le t_1, \tau_2 \le t_2, \ldots, \tau_n \le t_n) \\ &\quad = \int \prod_{i=1}^{n} p_t^{i|y} f(y) dy \end{aligned} \tag{23.19}$$

In the one-factor Gaussian copula, the credit variable X_i is Gaussian distributed. X_i depends on the common factor Y and an individual factor ε_{Xi} as follows:

$$X_i = \rho_{X_i Y} Y + \sqrt{1 - \rho_{X_i Y}{}^2} \varepsilon_{X_i}, \quad i = 1,\ 2,\ \ \ldots,\ N \tag{23.20}$$

where Y and ε_{Xi} are two independent standard Gaussian random variables.

We can get the default time $\tau_i = F_i^{-1}(\phi(X_i))$, where $\phi(\cdot)$ is the cumulative density function of a standard Gaussian variable. Then the conditional distribution of τ_i, given the common factor Y, is

$$p_t^{i|Y} = \phi\left(\frac{\phi^{-1}(F_i(t)) - \rho_{X_i Y} Y}{\sqrt{1 - \rho_{X_i Y}^2}}\right), \tag{23.21}$$

and the joint distribution function of $\tau_1, \tau_2, \cdots, \tau_n$ is as follows:

$$\begin{aligned} F(t_1, t_2, \ldots, t_n) &= \int \prod_{i=1}^{n} p_t^{i|y} f(y) dy \\ &= \int \left[\prod_{i=1}^{n} \phi\left(\frac{\phi^{-1}(F_i(t)) - \rho_{X_i Y} Y}{\sqrt{1 - \rho_{X_i Y}^2}}\right)\right] f(y) dy \end{aligned} \tag{23.22}$$

where $f(y)$ is the standard Gaussian density. The copula of default times is a Gaussian copula.

Law of Iterated Expectations

The law of iterated expectations states that $E(Y) = E(E(Y|X))$. The proof is as follows:

$$
\begin{aligned}
E(E(Y|X)) &= \int E(Y|X) f_X(x)\, dx \\
&= \int \left(\int y \ f_{X|Y}(y|x) \ \ dy \right) f_X(x)\, dx \\
&= \int \int y \ f_{X,Y}(x,y)\, dxdy \\
&= \int y \left(\int f_{X,Y}(x,y) \ \ dx \right) dy \\
&= \int y \ f(y)\, dy = E(Y)
\end{aligned}
\tag{23.23}
$$

Appendix 2: Cholesky Decomposition and Correlated Gaussian Random Numbers

Cholesky Decomposition

A symmetric positive defined real number matrix A can be decomposed as follows:

$$A = LL^T \tag{23.24}$$

where L is a lower triangular matrix with strictly positive diagonal entries and L^T is the transpose of L. This format is the Cholesky decomposition and it is unique. The entries of L are as follows:

$$L_{j,j} = \sqrt{A_{j,j} - \sum_{k=1}^{j-1} L_{j,k}^2} \tag{23.25}$$

$$L_{i,j} = \frac{1}{L_{j,j}} \left(A_{i,j} - \sum_{k=1}^{j-1} L_{i,k} L_{j,k} \right), \quad \text{for} \quad i > j \tag{23.26}$$

Correlated Gaussian Random Numbers

In financial applications, the Cholesky decomposition is usually used to create correlated Gaussian random variables. Suppose we need to generate n correlated random Gaussian numbers, X_1, X_2, $\cdots$, X_n, given a positive defined correlation coefficient matrix R. We first generate iid standard Gaussian random variables Z_1, Z_2,$\cdots$, Z_n and let

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_n \end{bmatrix}, \quad Z = \begin{bmatrix} Z_1 \\ Z_2 \\ \vdots \\ Z_n \end{bmatrix}.$$

Let $X = CZ$, where C is an $n \times n$ matrix; then $R = \mathrm{Var}(X) = E(XX^T) = CC^T$ and C is the lower triangular matrix of the Cholesky decomposition of R. Therefore, we can obtain the correlated Gaussian random variables $X_1, X_2, \cdots, X_n$ by letting $X = CZ$.

Appendix 3: Monte Carlo Simulation

Monte Carlo simulation is a computational algorithm based on random number sampling. It generates n iid random samples $X_1, X_2, \cdots, X_n$ of the random variable X and estimates $E[X]$ by $\overline{X} = \frac{X_1 + X_2 + \cdots + X_n}{n}$. $\overline{X}$ converges to $E[X]$ as $n \to \infty$, according to the law of large numbers.

Weak Law of Large Numbers (WLLN)

Let $X_1, X_2, \cdots, X_n$ be an iid sequence of random variables for which $E[X] < \infty$. Then $\frac{X_1+X_2+\cdots+X_n}{n} \xrightarrow{P} E[X]$ as $n \to \infty$. The WLLN implies that the Monte Carlo method converges to $E[X]$.

Strong Law of Large Numbers (SLLN)

Let $X_1, X_2, \cdots, X_n$ be an iid sequence of random variables for which $E[X] < \infty$. Then $\frac{X_1+X_2+\cdots+X_n}{n} \xrightarrow{a.s.} E[X]$ as $n \to \infty$. The SLLN implies that the Monte Carlo method almost surely converges to $E[X]$.

Uniform and Nonuniform Random Numbers

Usually, the computer has a uniform random number generator, which can generate a sequence of iid uniform random numbers $U_1, U_2, \cdots$ on [0,1]. Then, how do we get the nonuniform random numbers? Usually, these can be achieved by inversion. Given a distribution function $F(\cdot)$, there is a one-to-one mapping from $U(\cdot)$ to $F(\cdot)$. Since $F(\cdot)$ is nondecreasing, $F^{-1}(x) = \inf\{y : F(y) \geq x\}$. If F is continuous and strictly increasing, the inverse of the function is $F(F^{-1}(x)) = F^{-1}(F(x)) = x$. Thus, the nonuniform random numbers with distribution F can be obtained by $x = F^{-1}(u)$, where u is a uniform random number on [0,1].

References

Andersen, L., Sidenius, J., & Basu, S. (2003). All your hedges in one basket. *Risk, 16*(11), 67–72.

Bastide, D., Benhamou, E., & Ciuca, M. (2007). *A comparative analysis of basket default swaps pricing using the Stein method*. Working paper, pricing partners.

Brigo, D., & Chourdakis, K. (2009). Counterparty risk for credit default swaps: Impact of spread volatility and default correlation. *International Journal of Theoretical and Applied Finance, 12*(7), 1007–1026.

Chen, Z., & Glasserman, P. (2008). Fast pricing of basket default swaps. *Operations Research, 56*(2), 286–303.

Chiang, M. H., Yueh, M. L., & Hsieh, M. H. (2007). An efficient algorithm for basket default swap valuation. *Journal of Derivatives, 15*(2), 1–22.

Hull, J., & White, A. (2004). Valuation of a CDO and an nth to default CDS without Monte Carlo simulation. *Journal of Derivatives, 12*(2), 8–23.

Jarrow, R., & Turnbull, S. (1995). Pricing derivatives on financial securities subject to credit risk. *Journal of Finance, 1*, 53–85.

Joshi, M. S., & Kainth, D. (2004). Rapid and accurate development of prices and Greeks for Nth to default credit swaps in the Li model. *Quantitative Finance, 4*, 266–275.

Laurent, J. P., & Gregory, J. (2005). Basket default swaps, CDOs and factor copulas. *Journal of Risk, 7*(4), 103–122.

Li, D. X. (1999). The valuation of basket credit derivatives. *CreditMetrics Monitor*, 34–50.

Li, D. X. (2000). On default correlation: A copula function approach. *Journal of Fixed Income, 9*, 43–54.

Mashal, R., & Naldi, M. (2003). Pricing multi-name default swaps with counterparty risk. *Quantitative Credit Research*, 1–16.

Merton, R. C. (1974). On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance, 29*, 449–470.

Sklar, A. (1959). Fonctions de repartition a n dimensions et leurs marges. *Public Institute Statistic University of Paris, 8*, 229–231.

Stein, C. (1972). A bound for the error in the normal approximation to the distribution of a sum of dependent random variables. In *Proceedings of the sixth Berkeley symposium on mathematical statistics and probability* (pp. 583–602). Berkeley: University of California Press.

Walker, M. B. (2006). Credit default swaps with counterparty risk: A calibrated Markov model. *Journal of Credit Risk, 2*(1), 31–49.

Wu, P. C. (2010). Applying a factor copula to value basket credit linked notes with issuer default risk. *Finance Research Letters, 7*(3), 178–183.

Wu, P. C., Kao, L. J., & Lee, C. W. (2011). How issuer default risk affects basket credit linked note coupon rate. *International Journal of Information and Management Sciences, 22*(1), 59–71.

Panel Data Analysis and Bootstrapping: Application to China Mutual Funds

24

Win Lin Chou, Shou Zhong Ng, and Yating Yang

Contents

Abstract

Thompson (*Journal of Financial Economics 99*, 1–10, 2011) argues that double clustering the standard errors of parameter estimators matters the most when the number of firms and time periods are not too different. Using panel data of similar number in firms and time periods on China's mutual funds, we estimate double- and single-clustered standard errors by wild-cluster bootstrap procedure. To obtain the wild bootstrap samples in each cluster, we reuse the regressors (X) but modify the residuals by transforming the OLS residuals with weights which follow the popular two-point distribution suggested by Mammen (*Annals of Statistics 21*, 255–285, 1993) and others. We then compare them with other

W.L. Chou (✉)
Department of Economics and Finance, City University of Hong Kong, Hong Kong, China
e-mail: wlchou@e.cuhk.edu.hk

S.Z. Ng
Hong Kong Monetary Authority, Hong Kong, China
e-mail: bszng@hkma.gov.hk

Y. Yang
Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
e-mail: yatingyang.iof98g@nctu.edu.tw

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_24,

estimates in a set of asset pricing regressions. The comparison indicates that bootstrapped standard errors from double clustering outperform those from single clustering. Our findings support Thompson's argument. They also suggest that bootstrapped critical values are preferred to standard asymptotic *t*-test critical values to avoid misleading test results.

Keywords

Asset-pricing regression • Bootstrapped critical values • Cluster standard errors • Double clustering • Firm and time effects • Finance panel data • Single clustering • Wild-cluster bootstrap

24.1 Introduction

Researchers using finance panel data have increasingly realized the need to account for the residual correlation across both firms and/or time in estimating standard errors of regression parameter estimates. Ignoring such clustering can result in biased OLS standard errors. Two forms of residual dependence that are common in finance applications are time-series dependence and cross-sectional dependence. The former is called a firm effect, whereas the latter a time effect. The usual solution to account for the residual dependence is to compute clustered standard errors. The notable examples are Petersen (2009) and Thompson (2011).

Using the Monte Carlo-simulated panel data, Petersen (2009) compares the performance of many different standard error estimation methods surveyed in the literature. These methods include White's heteroskedasticity-robust standard errors, single clustering (by firm or by time), and double clustering (by both firm and time). His findings suggest that the performance of different methods depends on the forms of residual dependence. For example, in the presence of a firm effect, the clustered standard errors are unbiased and can produce correctly sized confidence intervals while those estimated by OLS, White, or Fama-MacBeth method are biased.

Much of the analysis in Petersen (2009) is based on the simulated panel data set whose data structure is certain. With the simulated panel data set, it is easier to choose among the estimation methods. This paper chooses an alternative method, namely, bootstrapping, to investigate the performance of standard errors estimated by White's OLS and single- and double-clustering methods with actually observed data. The use of the bootstrap method is motivated by Kayhan and Titman (2007) who show that bootstrapped standard errors are robust to heteroskedasticity, and serial correlation problems in panel finance data applications. Moreover, despite the wide use of the bootstrap in statistical and econometric applications, the survey finding of Petersen (2009) found that the bootstrap applications are relatively scarce in the finance literature. Hence, it may be of some interest to investigate the bootstrapping application to a set of panel finance data on Chinese mutual funds.

Table 24.1 Summary statistics for variables used in China's mutual fund regressions (Sept 2002–Aug 2006)

Variable	Mean	Std. deviation	Skewness	Kurtosis	Normality test stat. W-sq.	A-sq.
Mutual funds returns	0.0006	0.0580	0.1313	0.1489	0.1809***	0.2062***
Market excess returns	0.0025	0.0465	0.1554	−0.7801	2.4993***	18.3650***

Notes: Sample size $N = 2{,}592$. Individual mutual funds return is the dependent variable, while the market excess return is the independent variable. W-sq or W^2 = Cramer-von Mises test statistic, and A-sq or A^2 = Anderson-Darling test statistic. Both test statistics are empirical distribution function (EDF) statistics. The computing formulas for W^2 and A^2 statistics are available in Stephens (1974), p. 731
*** denotes statistical significance at the 1 % level

The bootstrap method is applied to a panel data set on the monthly returns for 54 Chinese mutual funds over the period of September 2002–August 2006. The data set is applied to a set of asset pricing regressions. Table 24.1 contains summary statistics such as sample skewness, sample excess kurtosis, and two test statistics for normality for the variables used in the asset pricing regressions. They suggest that normality does not characterize the variables. Additionally, since the time-series and/or cross-sectional independence assumption is most likely to be violated in panel data sets, ignoring these dependence could result in biased estimates of the standard errors. As evidenced in Kayhan and Titman (2007), bootstrapping is a possible alternative to handle this dependence issue.

In this paper, we are particularly interested in the performance of the bootstrapped double-clustered standard error estimates, because Thompson (2011) has argued that double clustering matters most when the number of firms and time periods are not too different. Given the panel data set we have collected which consists of 54 China mutual fund returns for 48 months with data exhibiting firm and time effects, double clustering is likely to show a significant difference. Our findings show that the bootstrapped standard errors from double clustering leads to more significant test results. We also demonstrate the importance of using bootstrapped critical values in hypothesis testing.

A number of bootstrap procedures are available in the literature. The bootstrap procedure we consider in this paper is the wild-cluster bootstrap procedure, which is an extended version of the wild bootstrap proposed by Cameron et al. (2008) in a cluster setting. This procedure has been shown by Cameron et al. (2008) to perform very well in practice, despite the fact that the pairs cluster bootstrap works well in principle. In this paper, our comparison of the finite-sample size of the bootstrapped t-statistics resulting from the pairs cluster bootstrap and wild-cluster bootstrap also indicates that the wild-cluster bootstrap performs better.

The rest of the paper is organized as follows. Section 24.2 presents the wild-cluster bootstrap procedure. Section 24.3 discusses the empirical results and the last section gives conclusions.

24.2 Wild-Cluster Bootstrap

The bootstrap we use in this paper is known as the "wild-cluster bootstrap" which is based on the nonclustered wild bootstrap proposed by Wu (1986). Proofs of the ability of the wild bootstrap to provide refinements in the linear regression model for linear regression models with heteroskedastic errors can be found in Liu (1988) and Mammen (1993). Cameron et al. (2008) extended Wu's (1986) wild bootstrap to the clustered setting.

The wild-cluster bootstrap procedure involves two stages. In the first stage, we consider the asset pricing model with G clusters (subscripted by g), and with N_g observations within each cluster, namely, $y_g = X_g'\beta + e_g$, $g = 1, \ldots, G$, β is $k \times 1$, X_g is $N_g \times k$, and y_g and e_g are $N_g \times 1$ vectors. We fit the model to the actually observed panel data set by OLS and estimate White's heteroskedasticity-robust standard errors, as well as standard errors, clustered by firm, by time, and by both. We then save residuals and denote them as $\hat{e}_g$.

The second stage is the resampling procedure which creates bootstrap samples for each cluster, $\left\{\left(\hat{y}_1^*, X_1\right), \ldots, \left(\hat{y}_G^*, X_g\right)\right\}$ where $\hat{y}_g^* = X_g'\hat{\beta} + e_g^*$. For each bootstrap sample in a cluster, the explanatory variables are reused and unchanged. The residuals e_g^* are constructed according to $e_g^* = a_g\hat{e}_g$, where the weight a_g serves as a transformation of the OLS residuals $\hat{e}_g$. A variety of constructions of weights a_g have proposed in the literature[1]. We use the two-point distribution of the weight variable a_g suggested in Mammen (1993), Brownstone and Valletta (2001), and Davidson and Flachaire (2008), namely, a_g which takes on one of the following values: (i) $\left(1-\sqrt{5}\right)/2 \approx -0.6180$ with probability $\left(1+\sqrt{5}\right)/\left(2\sqrt{5}\right) \approx 0.7236$ or (ii) $\left(1+\sqrt{5}\right)/2 \approx 1.6180$ with probability $1-\left(1+\sqrt{5}\right)/\left(2\sqrt{5}\right) \approx 0.2764$. Note that this random variable a_g has a mean zero with variance equal to one and the constraint $E(a_g^3) = 1$. We perform 1,000 replications. On each replication, a new set of e_g^* is generated and a new set of bootstrap-data is created based on $\hat{y}_G^* = X_g'\hat{\beta} + e_g^*$, and therefore a new set of parameter estimates, denoted as $\hat{\beta}^*$, is obtained.

For the 1,000 starred estimates (e_g^*), we calculate their bootstrapped standard errors using different estimation methods. The bootstrapped test statistics are calculated by dividing the 1,000 parameter estimates by the corresponding bootstrapped standard errors. The bootstrapped critical values can be obtained from the bootstrapped distribution of these test statistics. A detailed explanation of the procedure we follow is documented in Appendix 1.

[1]For example, in Cameron et al. (2008), a_g takes the value +1 with probability 0.5, or the value -1 with probability 1–0.5.

24.3 Empirical Results

24.3.1 Data and Definitions of the Variables

The Data. The sample consists of the returns on 54 publicly traded closed-end mutual funds that are gathered for 48 months from September 2002 to August 2006, a total of 2,592 observations. The mutual fund data set is purchased from the GTA Information Technology Company, Shenzhen, China. For simplicity, we divide the mutual funds investment objectives into equity growth and nongrowth funds. Our sample consists of 37 (68.5 %) growth funds and 17 (31.5 %) nongrowth funds. Although the first closed-end fund in China was sold to the public in April 1998[2], complete data for all 54 mutual funds are collected from September 2002.

The summary statistics for the variables used in China's mutual fund return regressions are displayed in Table 24.1, and the test statistics for normality for these variables suggest that the variables are non-normal.

Definitions of the variables. The following are the definitions of the variables used in the estimation procedures: $R_{i,\,t}$ = the return of a mutual fund i in excess of the risk-free rate in month t. Savings deposit rate of the People's Bank of China[3] is used as the proxy for the risk-free rate. $R_{m,\,t}$ = is the market return in excess of the risk-free rate in month t and is calculated[4] as follows:

$$R_{m,t} = 0.4 \times R_{1,t} + 0.4 \times R_{2,t} + 0.2 \times 0.0006, \tag{24.1}$$

where R_1 is the monthly return on Shanghai Stock Exchange index, R_2 the monthly return on Shenzhen Stock Exchange index, and 0.06 % is the monthly return on savings deposits.

24.3.2 Results

Table 24.2 presents the results from a set of asset pricing regressions of China mutual fund returns on its market returns. The firm and time effect in OLS residuals and data can be seen graphically in Fig. 24.1. Figure 24.1, Panel A, shows the within-firm autocorrelations in OLS residuals and independent variable, respectively, for lags 1–12. Panel B of Fig. 24.1 displays the within-month autocorrelations for residuals for lags 1–12. As the independent variable is a monthly series without cross-sectional units, we cannot calculate its within-month autocorrelations, thus no within-month plot is available for the independent variable. Figure 24.1 suggests that the residuals exhibit both firm and time effects, whereas

[2]Chen and Lin (2006), p. 384

[3]Data are taken from the website of the People's Bank of China: http://www.pbc.gov.cn/publish/zhengcehuobisi/627/index.html.

[4]The calculation follows that of Shen and Huang (2001), p. 24.

Table 24.2 Application to asset pricing modeling

		t-statistics			
		Clustered by			
		White	Firm	Time	Firm and time
Regressor	Estimate	I	II	III	IV
$R_{m,\ t}$	0.8399	46.935***	65.492***	9.017***	9.099***
1 % critical value (CV)		2.576	2.576	2.576	2.576
Intercept	−0.0016	−1.858	−2.222**	−0.326	−0.327
1 % critical value (CV)		−2.576	−2.576	−2.576	−2.576
Coefficient estimates	OLS				
R-squared	0.4539				
Regressor	Estimate	V	VI	VII	VIII
$R_{m,\ t}$	0.8401	47.215***	66.808***	50.886***	73.672***
Bootstrapped 1 % CV		1.951	2.661	2.520	11.898
Intercept	−0.0016	−1.878**	−2.262**	−2.033**	−2.659
Bootstrapped 1 % CV		−2.300	−2.717	−2.371	−12.288
Coefficient estimates	Wild-cluster bootstrap				
R-squared	0.4579				

Notes: The dependent variable is the monthly mutual fund return in excess of risk-free rate, denoted as $R_{i,t}$, and the independent variable $R_{m,t}$ is the market returns in excess of risk-free rate. Both variables are monthly observations from September 2002 to August 2006
***, and ** denote statistical significance at the 1 %, and 5 % levels, respectively

independent variable shows firm effects. Given Thompson's (2011) argument that double clustering matters most when the number of firms and time periods are not too different, our data set which has the number of mutual funds (54) similar to the number of months (48) is expected to imply that double clustering is important in our analysis.

In Table 24.2, the second column presents the OLS parameter estimates, whereas the remaining columns report their corresponding t-statistics by dividing each parameter estimate by its corresponding standard error. These t-statistics indicate all beta coefficients are statistically significant at the 1 % level, whereas the intercept is only significant in one case (under column II) when the standard error computed by single clustering by firm is used. More importantly, the *t*-statistics in Table 24.2 enable us to compare the clustered standard errors constructed from double and single clustering with the OLS White estimate. Notice that the t-statistic for beta coefficient obtained from double clustering ($\hat{SE}_{both}$) is 9.10 (column IV) which is much smaller than 46.9 calculated using the White method (column I), indicating the presence of firm and time effects. It also means that the double-clustering standard errors are much larger. A comparison of t-statistics in columns III and IV implies the $\hat{SE}_{both}$ of the beta coefficient (9.10) is similar to the standard error clustered by time which is 9.01. This means that the firm effects do not matter much. The comparison reveals that OLS White standard errors are underestimated when residuals exhibit both firm and time effects.

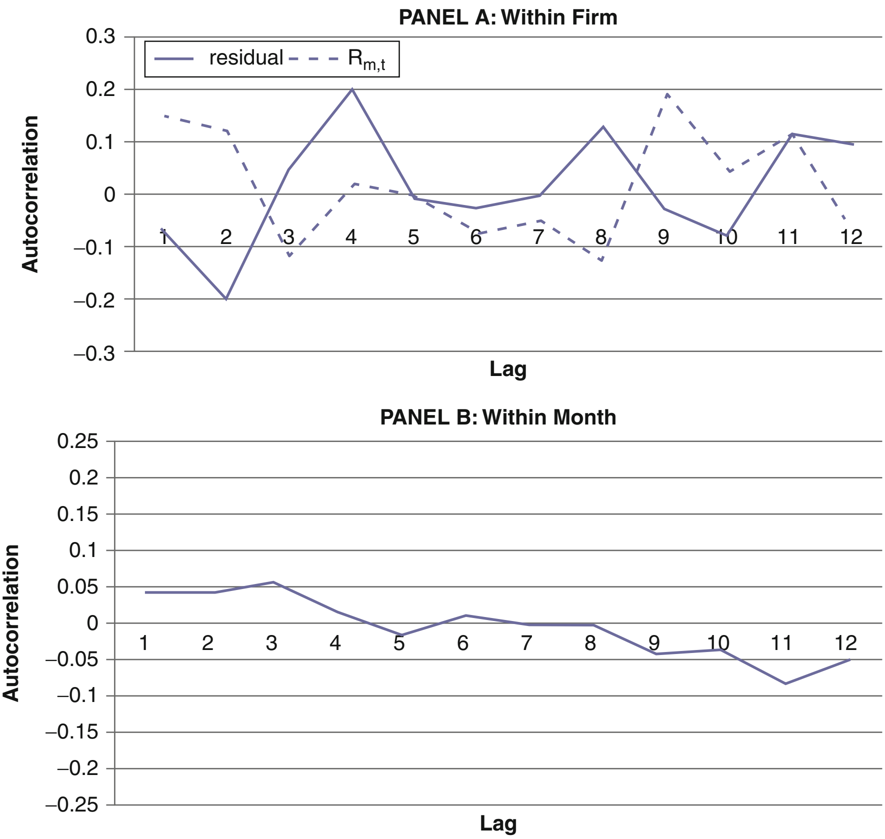


Fig. 24.1 The autocorrelations of residuals and the independent variable are plotted for 1–12 lags. The *solid lines* in *Panel A* and *B* show, respectively, within-firm and within-month autocorrelations in residuals, whereas the dashed line in *Panel A* shows the within-firm autocorrelations in the independent variable

Turning to the results obtained by the wild-cluster bootstrapping, the story changes. The bootstrapped t-statistics of the beta coefficient estimates displayed in columns V–VIII of Table 24.2 differ quite significantly from those in columns I–IV. The t-statistic is 47.2 when the bootstrapped White standard error is used and 66.8 if the bootstrapped standard error clustered by the firm is used. This means the firm effect is significant in the data. By a similar comparison, the bootstrapped t-statistic is 73.7 when the double-clustered standard error is used, meaning both the time and firm effects are strong in the residuals. The bootstrapped t-statistic is 50.9 when the bootstrapped standard error clustered by time is used implying the time effect exists in the data. These comparisons suggest both the firm and time effects matter in the computation of the bootstrapped standard errors using double as well as single clustering. Their implication is that we might follow what Kayhan and Titman (2007) have done in their study to simply compute the bootstrapped standard errors with our panel data set.

Table 24.3 Bootstrapped critical values

Percentiles	0.005	0.025	0.05	0.95	0.975	0.995
OLS White						
$R_{m,t}$	−1.842	−1.396	−1.174	1.213	1.490	1.951
Intercept	−2.300	−1.656	−1.396	1.385	1.587	2.091
Clustered by firm						
$R_{m,t}$	−2.621	−1.984	−1.698	1.766	2.007	2.661
Intercept	−2.717	−2.035	−1.700	1.778	2.014	2.672
Clustered by time						
$R_{m,t}$	−2.513	−1.795	−1.510	1.517	1.821	2.520
Intercept	−2.371	−1.956	−1.608	1.629	1.917	2.364
Clustered by firm and time						
$R_{m,t}$	−6.517	−4.376	−3.491	3.081	4.416	11.898
Intercept	−12.288	−4.053	−2.792	3.136	4.384	8.168

Note: Critical values are obtained by implementing the bootstrap procedure presented in the Appendix 1

Table 24.4 Rejection rates using wild-cluster bootstrapped critical values

	OLS White	Clustered by firm	Clustered by time	Clustered by firm and time
$R_{m,t}$	0.050	0.067	0.053	0.052
Intercept	0.042	0.034	0.037	0.036

Note: The number of replications is 1,000

The statistical significance of the bootstrapped t-statistics of the beta coefficient estimates is determined by using the bootstrapped critical values reported in Table 24.3. Compared with the bootstrapped critical values presented in Table 24.3, we notice that all bootstrapped t-statistics constructed from bootstrapped standard errors are statistically significant at the 1 % level. The intercept is now significant in three cases (columns V–VII) when the standard errors were computed by White and by single clustering. It is noteworthy that in Table 24.3 the bootstrapped critical values on beta coefficient estimates when double clustering is used are numerically larger than the corresponding asymptotic *t*-test critical values of 2.58 (1 %), 1.96 (5 %), and 1.65 (10 %), indicating that the use of the large-sample (normal approximation) critical values can lead to misleading test results when both firm and time effects exist in the residuals. On the other hand, intercept coefficient in column VIII was not significant when bootstrapped double clustering is used. Interestingly, we observe from Table 24.3 that bootstrapped critical values differ considerably depending on different standard error estimation methods.

We now examine the finite-sample size of the bootstrapped t-statistics assuming the beta coefficient takes the OLS estimate under the null hypothesis. Table 24.4 shows that for the bootstrapped t-statistics, no serious size distortion is found as

Table 24.5 Rejection rates using conventional critical values

	OLS White	Clustered by firm	Clustered by time	Clustered by firm and time
$R_{m,t}$	0.004	0.069	0.037	0.206
Intercept	0.012	0.046	0.036	0.203

Note: The number of replications is 1,000

reflected by the close to 5 % size values. However, for the tests based on the standard asymptotic *t*-test critical values, all tests suffer from serious size distortion. For example, those based on the OLS White and clustered by time are undersized, while others oversized (see Table 24.5).

24.4 Conclusion

In this paper, we examine the performance of single- and double-clustered standard errors using the wild-cluster bootstrap method. The panel data set on the Chinese mutual funds used in the analysis has similar number of firms (54) and time periods (48); we are particularly interested in the performance of the bootstrapped double-clustered standard errors. This is mainly due to the conclusion made in Thompson (2011) that double clustering the standard errors matters the most when the number of firms and time periods are not too different.

In the presence of firm and time effects, the standard OLS White standard errors are found to be underestimated when compared to standard errors computed from double clustering (columns I–IV, Table 24.2). Further, the wild-cluster bootstrapped standard errors are found to account for the firm and time effects in residuals, as evidenced in column VIII of Table 24.2. The bootstrapped t-statistic computed by OLS White method is found to be much smaller than that calculated from the double clustering, suggesting that the firm and time effects in residuals are strong.

The size values for the test statistics of the beta coefficient estimates in Table 24.4 suggest that the bootstrapped double clustering outperforms the single clustering either by firm or by time. They support Thompson's (2011) argument that double clustering the standard errors of parameter estimators matters the most when the number of firms and time periods are not too different. Size distortions reported in Table 24.5 imply that it may not be appropriate to compare the bootstrapped t-statistics with standard *t*-test critical values. These findings also suggest that to avoid obtaining misleading test results with the presence of either firm or time effects or both, the bootstrapped critical values are preferred to conventional critical values. Additionally, a comparison of the sizes displayed in Table 24.4 with those calculated using the pairs cluster bootstrap method shown in Table 24.6 also suggests the wild-cluster bootstrap approach performs better.

Table 24.6 Rejection rates using pairs cluster bootstrapped critical values

	OLS White	Clustered by firm	Clustered by time	Clustered by firm and time
$R_{m,t}$	0.043	0.040	0.040	0.040
Intercept	0.062	0.048	0.059	0.060

Note: The number of replications is 1,000

Appendix 1: Wild-Cluster Bootstrap Procedure

The following steps are used to obtain the wild-clustered bootstrapped standard errors and critical values:[5]

(i) *Define firm effects, and time effects*. The asset pricing model using individual variables $R_{i,t}$ and $R_{m,t}$ is specified as

$$R_{it} = \beta_0 + R_{m,t}\beta_1 + e_{it},\ \ i = 1, \ldots, N,\ \ t = 1, \ldots, T \tag{24.2}$$

where the variables $R_{i,t}$ and $R_{m,t}$ are, respectively, the return of mutual fund i and the market return in excess of risk-free rate in month t. β_0 and β_1 are unknown parameters. The construction of the variables is detailed in Sect. 24.3.1. e_{it} is the error term. It may be heteroskedastic but is assumed to be independent of the explanatory variable $E(e_{it} \mid R_{m,t}) = 0$.

Following Thompson (2011), we make the following assumptions on the correlations between errors, e_{it}:

(a) *Firm effects*: The errors may be correlated across time for a particular firm, that is, $E(e_{it}, e_{ik} \mid R_{m,t}, R_{m,k}) \neq 0$ for all $t \neq k$.

(b) *Time effects*: The errors may be correlated across firms within the same time period, that is,

$$E\left(e_{it}, e_{jt} \middle| R_{m,t}\right) \neq 0 \quad \text{for all} \ \ i \neq j.$$

Let G be the number of clusters, and let N_g be the number of observations within each cluster. The errors are assumed to be independent across clusters but correlated within clusters. The asset pricing model can be written as

$$\begin{aligned} R_{ig} &= \beta_0 + R_m\beta_1 + e_{ig}, \qquad & i = 1, \ldots, N_g \qquad g = 1, \ldots, G, \\ \mathbf{R}_g &= \delta\beta_0 + \mathbf{R}_{mg}\beta_1 + \mathbf{e}_g, \qquad & g = 1, \ldots, G, \end{aligned} \tag{24.3}$$

where R_{ig}, R_m, and e_{ig} are scalars; $\mathbf{R}_g$, $\mathbf{R}_{mg}$, and $\mathbf{e}_g$ are $N_g \times 1$ vectors; and δ is $N_g \times 1$ vector with all elements equal to 1.

(ii) *Fit data to model*. We fit model (Eq. 24.3) to the observed data using OLS and obtain the parameter estimates $\hat{\beta}_0$ and $\hat{\beta}_1$ together with the OLS residuals $\hat{\mathbf{e}}_g$, g = 1,. . ., G

[5]See also Cameron et al. (2008) for details

(iii) *Construct 1,000 bootstrap samples*. The bootstrap-residuals are obtained according to the following transformation relation: $e_g^* = a_g \hat{e}_g$, where a_g takes on one of the following values: (i) $\left(1-\sqrt{5}\right)/2 \approx -0.6180$ with probability $\left(1+\sqrt{5}\right)/\left(2\sqrt{5}\right) \approx 0.7236$ or (ii) $\left(1+\sqrt{5}\right)/2 \approx 1.6180$ with probability $1-\left(1+\sqrt{5}\right)/\left(2\sqrt{5}\right) \approx 0.2764$.

Hence,

(a) For each cluster $g = 1, \ldots, G$, set $\mathbf{e}_\mathbf{g}^* = 1.618^* \hat{\mathbf{e}}_\mathbf{g}$ with probability 0.2764 or $\mathbf{e}_\mathbf{g}^* = -0.618^* \hat{\mathbf{e}}_\mathbf{g}$ with probability 0.7236.

(b) Repeat (a) 1,000 times to obtain $\mathbf{e}_g^*$ and then construct the bootstrap samples $\mathbf{R}_g^*$ as follows:

$$\mathbf{R}_g^* = \delta\hat{\beta}_0 + \mathbf{R}_{\mathrm{mg}}\hat{\beta}_1 + e_g^*. \tag{24.4}$$

(iv) With each pseudo sample generated in step (iii), we estimate the parameters $\hat{\beta}_0^*$ and $\hat{\beta}_1^*$ by OLS, White standard errors ($\hat{SE}_{\mathrm{white}}$)which are OLS standard errors robust to heteroskedasticity, as well as standard errors clustered by firm ($\hat{SE}_{\mathrm{firm}}$), by time ($\hat{SE}_{\mathrm{time}}$), and by both firm and time ($\hat{SE}_{\mathrm{both}}$). The simulations are performed using GAUSS 9.0. The standard error formulas can be found in Thompson (2011).

(v) Construct bootstrapped test statistics by taking ratios of $\hat{\beta}_i^* \, (i = 0, 1)$ obtained by OLS to its corresponding $\hat{SE}_{\mathrm{white}}$, $\hat{SE}_{\mathrm{firm}}$, $\hat{SE}_{\mathrm{time}}$, and $\hat{SE}_{\mathrm{both}}$ obtained in step (iv). More specifically, the bootstrapped test statistics are expressed as follows:

$$w_i^* = \frac{\hat{\beta}_i^* - \hat{\beta}_i}{\hat{SE}\left(\hat{\beta}_i^*\right)}, i = 0, 1, \tag{24.5}$$

where $\hat{SE}\left(\hat{\beta}_i^*\right)$ can either be $\hat{SE}_{\mathrm{white}}$, $\hat{SE}_{\mathrm{firm}}$, $\hat{SE}_{\mathrm{time}}$, or $\hat{SE}_{\mathrm{both}}$.

(vi) Obtain the empirical distribution of the individual test statistics by sorting the 1,000 test statistics computed in step (v) in an ascending order. Bootstrapped critical values are then obtained from this empirical distribution at the following quantiles: 0.5 %, 2.5 %, 5 %, 95 %, 97.5 %, and 99.5 %, respectively.

References

Brownstone, D., & Valletta, R. (2001). The bootstrap and multiple imputations: Harnessing increased computing power for improved statistical tests. *Journal of Economic Perspectives, 15*, 129–141.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics, 90*, 414–427.

Chen, Z., & Lin, R. (2006). Mutual fund performance evaluation using data envelopment analysis with new risk measures. *OR Spectrum, 28*, 375–398.

Davidson, R., & Flachaire, E. (2008). The wild bootstrap, tamed at last. *Journal of Econometrics, 146*, 162–169.

Kayhan, A., & Titman, S. (2007). Firms' histories and their capital structures. *Journal of Financial Economics, 83*, 1–32.

Liu, R. Y. (1988). Bootstrap procedures under some non-IID models. *Annals of Statistics, 16*, 1696–1708.

Mammen, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models. *Annals of Statistics, 21*, 255–285.

Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *The Review of Financial Studies, 22*, 435–480.

Shen, W. T., & Huang, S. L. (2001). An empirical analysis of China's mutual funds: Performance and evaluation. *Economic Research, 9*, 22–30 (in Chinese).

Stephens, M. A. (1974). EDF statistics for goodness of fit and some comparisons. *Journal of the American Statistical Association, 69*, 730–737.

Thompson, S. B. (2011). Simple formulas for standard errors that cluster by both firm and time. *Journal of Financial Economics, 99*, 1–10.

Wu, C. F. G. (1986). Jackknife, bootstrap and other resampling methods in regression analysis. *Annals of Statistics, 14*, 1261–1295.

Market Segmentation and Pricing of Closed-End Country Funds: An Empirical Analysis

25

Dilip K. Patro

Contents

Abstract

This paper finds that for closed-end country funds, the international CAPM can be rejected for the underlying securities (NAVs) but not for the share prices. This finding indicates that country fund share prices are determined globally where as the NAVs reflect both global and local prices of risk. Cross-sectional variations in the discounts or premiums for country funds are explained by the differences

The views expressed in this paper are strictly that of the author and not of the OCC or the US department of Treasury.

D.K. Patro
RAD, Office of the Comptroller of the Currency, Washington, DC, USA
e-mail: dilip.patro@occ.treas.gov

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_25,

in the risk exposures of the share prices and the NAVs. Finally, this paper shows that the share price and NAV returns exhibit predictable variation and country fund premiums vary over time due to time-varying risk premiums. The paper employs **generalized method of moments** (**GMM**) to estimate stochastic discount factors and examines if the price of risk of closed-end country fund shares and NAVs is identical. GMM is an econometric method that was a generalization of the method of moments developed by Hansen (*Econometrica 50*, 1029–1054, 1982). Essentially GMM finds the values of the parameters so that the sample moment conditions are satisfied as closely as possible.

Keywords
Capital markets • Country funds • CAPM • Closed-end funds • Market segmentation • GMM • Net asset value • Stochastic discount factors • Time-varying risk • International asset pricing

25.1 Introduction

The purpose of this paper is to provide an empirical analysis of pricing of closed-end country equity funds in the context of rational asset-pricing models which account for the role of market segmentation and time-varying risk premiums. Specifically, the paper addresses the following issues. How are country fund share prices and net asset values (NAVs) determined? What are the implications of differential pricing of closed-end fund shares and NAVs for cross-sectional and time-series variations in the premiums of the funds? The answers to these questions contribute to the burgeoning literature on country funds.

Closed-end country equity funds are a relatively recent innovation in international capital markets. Whereas only four closed-end country equity funds traded in New York at the end of 1984, currently there are 94 closed-end country equity funds targeting over 31 countries. Past researchers have examined issues related to the benefits of diversification from holding these funds (Bailey and Lim 1992; Chang et al. 1995; Bekaert and Urias 1996) and how they should be designed and priced (Hardouvelis et al. 1993; Diwan et al. 1995; Bodurtha et al. 1995). Relatively unexplored is how the expected returns of the country fund share prices and NAVs are determined, in a framework of market segmentation and time-varying risk premiums.

Country fund share prices are determined in the USA, but their NAVs are determined in the country of origin of the fund. Models of international asset pricing (e.g., Errunza and Losq 1985) and models of country fund pricing (e.g., Errunza et al. 1998) suggest that the expected returns on country fund share prices should be determined by their covariances with the world market portfolio. However, if there are barriers to investment, such as limits on ownership, the capital markets may be segmented. In such a segmented market, the expected returns of the NAVs will be determined by their covariances with the world market portfolio as well as the home market portfolio, to the extent that the home market is not spanned by the world market.

The foundation of this paper is in such a framework of market segmentation and its implications for international capital market equilibrium. In such a market structure where the local investors have complete access to the world market but the foreign investors (e.g., the US investors) have only partial access to the local equity market in terms of the percentage of equity of a firm they can own, the prices paid by foreigners relative to local investors can be higher due to the limited supply of the local securities. To preclude arbitrage, it is assumed that the local investors cannot buy the securities at a lower price and sell it to the foreign investor at a higher price. Thus, foreign investors are willing to pay a premium due to the diversification benefits from adding that security to their purely domestic portfolio. The premium arises only if the cash flows of the local security are not spanned by the purely domestic portfolio for the foreign investor. Since closed-end country fund's share prices and NAVs are determined in two separate markets, the expected returns on the prices and NAVs can be different, leading to premiums, both positive and negative. Thus, barriers to investments may be sufficient to generate premiums.[1]

Even in the absence of institutional restrictions on foreign equity ownership or even in a purely domestic context (see, e.g., Pontiff 1997), it is still possible for observed price returns to be much more volatile than NAV returns. As shown in the academic literature, closed-end country fund share prices and NAVs are not perfect substitutes. The sources of these differences may be attributable to differences in numeraires, information, and possibly noise trading causing excess volatility in the share price returns (see Lee et al. 1991). Apart from investment restrictions, these imperfections alone may be sufficient to generate premiums. However, Bonser-Neal et al. (1990) document that relaxation of investment restrictions leads to a decrease in the share price-NAV ratio for a sample of country equity funds. Further, Bodurtha et al. (1995) document that the correlation of changes in country fund premiums and domestic fund premiums is low and insignificant, indicating that the structure of international capital markets is an important contributor in determining premiums.

Thus, the objective of this paper is to provide an analysis of what explains the expected returns on closed-end country fund share prices and NAVs in a segmented market framework. The paper utilizes both unconditional and conditional tests on mean-variance efficiency of the world market index (as proxied by the Morgan Stanley Capital International world index) and provides results on the cross-sectional and time-series variations of premiums across closed-end country equity funds. In addition, the paper employs generalized method of moments (GMM) as discussed in Appendix 1 to estimate stochastic discount factors and examines if the price of risk of closed-end country fund shares and NAVs is identical.

[1]Hence forth both premiums and discounts are referred to as premiums. Therefore, a discount is treated as a negative premium.

The sample consists of 40 closed-end country equity funds. Twenty of the funds are from developed markets and 20 funds are from emerging markets.[2] The main empirical findings of the paper are as follows. For country fund share prices, the hypothesis that the unconditional international CAPM is a valid model cannot be rejected. However, for the NAVs the international CAPM can be rejected. This finding suggests that country fund share prices and NAVs are not priced identically. The share prices reflect the global price of risk only, but the NAVs reflect both the global and the local prices of risk. It is shown that the differences in risk exposure to the world market index of share prices and NAVs can explain up to 18.7 % of the cross-sectional variations in premiums for developed market funds, but only 1.9 % of the variation for emerging market funds.

When conditioning on information is allowed, the international CAPM explains country fund share returns and NAV returns for both developed and emerging markets. However, the hypothesis that the price of risk is identical between closed-end fund shares and NAVs can be rejected for alternate stochastic discount factors for majority of the markets. This finding is consistent with market segmentation. Therefore, differential pricing of the fund shares and the underlying portfolio causes expected returns to be different and explains the existence of premiums and their variation over time. Finally, it is shown that the country fund premiums vary over time due to differential conditional risk exposures of the share prices and NAVs.

The principal contributions of this paper are as follows. Existence of premiums on closed-end funds has been a long-standing anomaly. In the domestic setting, negative premiums have been the norm, which has been attributed to taxes, management fees, illiquid stocks, or irrational factors (e.g., noise trading). Although these factors may also be important in explaining country fund premiums, unlike domestic closed-end funds, country fund share prices and NAVs are determined in two different market segments. This paper provides a rational explanation for the premiums on country funds based on differential risk exposures of the share prices and NAVs and market segmentation. Unlike the noise trading hypothesis which assumes the presence of "sentiment" or an additional factor for the share prices but not the NAVs, this paper shows that the same factor may be priced for the share prices and the NAVs, but priced differently. The differential risk exposures are shown to be among the significant determinants of cross-sectional variations in the premiums. Further, this paper examines the role of time variations in expected returns for country fund returns and attributes the time-varying country fund premiums to time-varying country fund returns.

The paper is organized as follows. Section 25.2 presents the theoretical motivation and the hypotheses. Section 25.3 presents the data and the descriptive statistics. Section 25.4 provides the empirical results for pricing of country funds. Concluding remarks are presented in the last section.

---

[2]In 1993, the World Bank defined an emerging market as a stock market in a developing country with a GNP per capita of $8,625 or less. This is the definition of an emerging market in this paper.

25.2 Theoretical Motivation

This section first focuses on the theoretical motivation for the empirical testing of the unconditional mean-variance efficiency of the world market index in the context of country funds. The following subsections present tests for cross-sectional variations in country fund premiums and the methodology for pricing country funds using stochastic discount factors. Finally, this section outlines the methodology for explaining the time variations in country fund premiums attributable to time-varying risk premium.

25.2.1 Pricing of Country Funds in the Context of International CAPM

In a global pricing environment, the world market portfolio surrogates the market portfolio. If purchasing power parity is not violated and there are no barriers to investment, the world market portfolio is mean-variance efficient (see, e.g., Solnik 1996; Stulz 1995) and expected returns are determined by the international CAPM. The international CAPM implies that the expected return on an asset is proportional to the expected return on the world market portfolio:

$$E[r]_i = \beta_i E[r_w] \tag{25.1}$$

where, for any asset i, $E[r_i]$ is the expected excess return on the asset and $E[r_w]$ is the expected excess return on the world market portfolio. The excess returns are computed in excess of the return on a risk-free asset. Following a standard practice (see Roll 1977), mean-variance efficiency of a benchmark portfolio can be ascertained by estimating a regression of the form

$$r_{i,t} = \alpha_i + \beta_{i,w} r_{w,t} + e_{i,t} \tag{25.2}$$

where $r_{i,t}$ is the excess return on a test asset and $r_{w,t}$ is the return on a benchmark portfolio.

The GMM-based methodology as outlined in Mackinlay and Richardson (1991) is employed to test mean-variance efficiency of the world market index in the context of country funds, using just a constant and the excess return on the world market index as the instruments.[3] The hypothesis that $\alpha_i = 0$, for all $i = 1, \ldots, N$, where N is the number of assets, is tested. This is the restriction implied by the international CAPM. The joint hypothesis is tested by a Wald test for country fund share prices and NAVs returns. Since country fund's share prices are determined with complete access to their markets of origin, their expected returns are expected to be determined via the

[3]Although the evidence in favor of the domestic CAPM is ambiguous, many studies such as Cumby and Glen (1990) and Chang et al. (1995) do not reject the mean-variance efficiency of the world market index. Interestingly, although Cumby and Glen (1990) do not reject mean-variance efficiency of the world market index, they reject mean-variance efficiency of the US market index.

international CAPM (see, e.g., Errunza and Losq 1985). However, if the country from which the fund originates has restrictions on foreign equity investments, the expected returns of the NAVs will be determined via their covariances with both the world market portfolio and the part of the corresponding local market portfolio which is not spanned by the world market portfolio (see Errunza et al. 1998).

$$\text{Letting } e_{i,t} = r_{i,t} - \alpha_i - \beta_{i,w} r_{w,t} \tag{25.3}$$

for the N assets, the residuals from the above equation can be stacked into a vector e_{t+1}. The model implies that $E[e_{i,t+1}|Z_t] = 0$, for $\forall$ i and $\forall$ t. Therefore, $E[e_{t+1} \otimes Z_t] = 0$ for $\forall$ t, where Z_t is a set of predetermined instruments. GMM estimation is based on minimizing the quadratic form $f'\Omega f$, where the sample counterpart of $E[e_{t+1} \otimes Z_t]$ is given by $f = \{1/T[\sum E[e_{t+1} \otimes Z_t]\}$ and Ω is an optimal weighting matrix in the sense of Hansen (1982).

25.2.2 Cross-Sectional Variations in Country Fund Premiums

If the mean-variance efficiency of the world market portfolio is not rejected, then expected returns of the test assets are proportional to their "betas" with respect to the world market portfolio. Therefore,

$$E(r_a) = \lambda_w \beta_{a,w} \text{ where } a = p \text{ or } n \tag{25.4}$$

where $E(r_p)$ and $E(r_n)$ are the expected excess returns on the share prices and NAVs, respectively, and λ_w is the risk premium on the world market index. Stulz and Wasserfallen (1995) demonstrate that the logarithm of two share prices can be written as a linear function of the differences in their expected returns.[4] Therefore,

$$\text{Prem} = \log(P/NAV) = E(r_n) - E(r_p) \tag{25.5}$$

where Prem is the premium on a fund calculated as the logarithm of the price-NAV ratio. Combining Eqs. 25.4 and 25.5 leads to the testable implication

$$\text{Prem} = \lambda_w \left(\beta_{n,w} - \beta_{p,w}\right) \tag{25.6}$$

The above equation assumes that the international CAPM is valid and world capital markets are integrated. In reality, the world markets may be segmented. Therefore, the effect of market segmentation is captured by introducing additional

[4]Stulz and Wasserfallen (1995) use the log-linear approximation of Campbell and Ammer to write the logarithm of the price of a stock as $\ln(P) = E\Sigma\eta^i[(1 - \eta)d^{t+j+1} - r^{t+j+1}] + \theta$, where η is a log-linear approximation parameter and r^{t+j+1} is the return from t+j to t+j+1. Assuming that η is the same for prices and NAVs, and the relation holds period by period, the premium can be written as above.

variables based on prior research of Errunza et al. (1998) who document that measures of access, spanning, and substitutability of the prices and NAVs have explanatory power for the cross-sectional variation of premiums. The spanning measure (SPN) is the conditional variance of the NAV return of a fund, unspanned by the US market return and the fund's share price returns, with specification as follows:

$$r_{n,t} = \alpha_i + \beta_i r_{US,t} + \beta_i r_{p,t} + e_{i,t} \tag{25.7}$$

where $e_{i,t}$ has a GARCH (1, 1) specification. The conditional volatility of $e_{i,t}$ is the measure of spanning. The measure of substitution (SUB) is the ratio of conditional volatilities of the share price and NAV returns of a fund not spanned by the US market. The specifications for the expected return equations are

$$r_{n,t} = \alpha_i + \beta_i r_{US,t} + e_{n,t} \tag{25.8}$$

$$r_{p,t} = \alpha_i + \beta_i r_{US,t} + e_{p,t} \tag{25.9}$$

where the error terms $e_{n,t}$ and $e_{p,t}$ have GARCH (1, 1) specifications.[5] The ratio of the conditional volatilities of the residuals in Eqs. 25.8 and 25.9 is used as the measure of substitutability of the share prices and NAVs.

Since it is difficult to systematically classify countries in terms of degree of access, a dummy variable is used to differentiate developed and emerging markets. The dummy variable takes value of one for developed markets. Another measure of access, which is the total purchase of securities from a country by US residents as a proportion of total global purchase, is also used as a measure of access to a particular market. It is expected that the premiums are lower for countries with easy access to their capital markets. Therefore, the coefficient on the access variable is expected to be negative. The measure of spanning is interpreted as the degree of ease with which investors could obtain substitute assets for the NAVs. As noted earlier, even if there are barriers to foreign equity ownership, availability of substitute assets can overcome the barriers. Since the spanning variable is the volatility of the residual as specified in Eq. 25.7, it is expected to be positive. The measure of imperfect substitutability of the share prices and NAVs is the ratio of the variances of the share price and NAV returns. It is expected that the premium is inversely related to this measure. Therefore, the extended cross-sectional equation is

$$\text{Prem}_i = \delta_0 + \delta_1 \text{DIF}_i + \delta_2 \text{SPN}_i + \delta_3 \text{SUB}_i + \delta_4 \text{ACC}_i + e_i \tag{25.10}$$

where:

DIF_i = the difference of the betas on the world market index for the NAVs and the share prices.

SPN_i = the conditional residual volatility from Eq. 25.7 used as measures of spanning.

---

[5]Note that unlike here, Errunza et al. (1998) use returns on industry portfolios to proxy for the US market.

SUB_i = the imperfect substitutability of the share prices and the NAVs proxied by the ratio of conditional volatilities in Eqs. 25.8 and 25.9.

ACC = a measure of access to a market proxied by a dummy variable which is 1 for developed markets [ACC(d)] or the total purchase of securities from the US residents from that country as a fraction of global security purchases [ACC(cf)].

The empirical specification in Eq. 25.10 or its variations is used in the empirical analysis for explaining cross-sectional variations in premiums. The higher the difference of risk exposures, the higher is the premium. Therefore, the sign of δ_1 is expected to be positive. However, as the level of unspanned component is higher, the level of premiums should be higher. Therefore the sign of δ_2 is expected to be positive. Also, the higher the substitutability of the share prices and the NAVs, the lower should be the premium. Therefore, δ_3 is expected to be negative. Since higher access is associated with lower premiums, δ_4 is also expected to be negative.

25.2.3 Conditional Expected Returns and Pricing of Country Funds

If investors use information to determine expected returns, expected returns could vary over time because of rational variations in risk premium. In such a scenario, the premium on a country fund can be time varying as a result of the differential pricing of the share prices and the NAVs.

Let $R_{t+1} = \frac{P_{t+1}+D_{t+1}}{P_t}$, where P_{t+1} is the price at time t+1 and D_{t+1} is the dividend at time t+1. Asset-pricing models imply that $E[M_{t+1}R_{t+1}|Z_t] = 1$, where Z_t is a subset of the investor's information set at time t (see, e.g., Ferson 1995). M is interpreted as a stochastic discount factor (SDF). The existence of the above equation derives from the law of one price. Given a particular form of an SDF, it is estimated by using GMM as outlined earlier.[6]

The SDFs examined in this paper are the SDFs implied by the international CAPM and its extension under market segmentation. The international CAPM is obtained by assuming that the SDF is a linear function of the return on the world market index ($R_{w,\,t+1}$). Specifically, the SDF is:

$$\text{International CAPM: } M = \lambda_0 + \lambda_1 R_{w,t+1} \tag{25.11}$$

A two-factor SDF, where the second factor is the return on a regional index (specifically for Asia or Latin America), is also estimated. Such a model is implied by the models of Errunza and Losq (1985).[7] Specifically, the SDFs is:

$$\text{Two-factor model: } M = \lambda_0 + \lambda_1 R_{w,t+1} + \lambda_2 R_{h,t+1} \tag{25.12}$$

[6]Ferson and Foerster (1994) show that an iterated GMM approach has superior finite sample properties. Therefore the iterated GMM approach is used in the estimations.

[7]For such conditional representation, see, for example, Ferson and Schadt (1996).

Once an SDF is estimated for a group of share prices and NAVs, the coefficients of the estimated SDFs are compared to test the hypothesis that for a given SDF, the share price and NAV are priced identically.[8]

25.2.4 Conditional Expected Returns and Time-Varying Country Fund Premiums

To examine the role of time-varying risk premiums in explaining the time variation in country fund premiums, the following procedure is used. Similar to Ferson and Schadt (1996), a conditional international CAPM in which the betas vary over time as a linear function of the lagged instrumental variables is used. The following equations are estimated via GMM:

$$r_{p,t+1} = \alpha_p + \beta_{p,w}(Z_t) r_{w,t+1} + e_{p,t+1} \tag{25.13}$$

$$r_{n,t+1} = \alpha_n + \beta_{n,w}(Z_t) r_{w,t+1} + e_{n,t+1} \tag{25.14}$$

where the excess return on the world market index is scaled by a set of instrumental variables. Using Eq. 25.6 and the above time-varying betas, the premium can be written as

$$\mathrm{Prem}_{t+1} = \gamma_0 + \gamma_1 \left[\beta_{n,w}(Z_t) - \beta_{p,w}(Z_t)\right] + e_{\mathrm{prem},t+1} \tag{25.15}$$

The empirical specification in Eq. 25.15 is used for explaining the time-varying premiums. If the time-varying betas are significantly different and their difference can explain the time variation of country fund premiums, the coefficient γ_1 is expected to be positive and significant.

25.3 Data and Descriptive Statistics

The sample includes all single-country closed-end country equity funds publicly trading in New York as of 31 August 1995. The test sample is limited to country funds with at least 100 weeks of weekly NAV observations.[9] An overview of the sample is presented in Table 25.1. The funds are classified as developed market or

[8]See Cochrane (1996) for such specifications. Cochrane (1996) calls such models as "scaled factor models." See Bansal et al. (1993) for a nonlinear specification of stochastic discount factors.

[9]Unlike industrial initial public offerings (IPOs), closed-end fund IPOs are "overpriced" (Weiss 1989). Peavy (1990) finds that new funds show significant negative returns in the aftermarket. Hanley et al. (1996) argue that closed-end funds are marketed to a poorly informed public and document the presence of flippers – who sell them in the immediate aftermarket. They also document evidence in support of price stabilization in the first few days of trading. Therefore, the first 6 months (24 weeks) of data for each fund is excluded in the empirical analysis.

Table 25.1 Closed-end country funds: sample overview

Fund name	IPO date	Symbol	No of shares (millions)	Net assets ($ millions)	Listing
Panel A: Developed market funds					
First Australia fund	16 December 1985	**AUS**	15.9	163.4	AMEX
Italy fund	27 February 1986	**ITA**	9.5	89.4	NYSE
Germany fund	18 July 1986	**GER**	13.5	173.9	NYSE
UK fund	6 August 1987	**GBR**	4.0	51.2	NYSE
Swiss Helvetia fund	19 August 1987	**SHEL**	9.2	181.8	NYSE
Spain fund	21 June 1988	**SPN**	10.0	94.2	NYSE
Austria fund	22 September 1989	**AUT**	11.7	107.9	NYSE
New Germany fund	24 January 1990	**GERN**	32.5	373.8	NYSE
Growth fund of Spain	14 February 1990	**GSPN**	17.3	164.0	NYSE
Future Germany fund	27 February 1990	**GERF**	11.9	171.5	NYSE
Japan OTC equity fund	14 March 1990	**JPNO**	11.4	115.4	NYSE
Emerging Germany fund	29 March 1990	**GERE**	14.0	130.6	NYSE
Irish investment fund	30 March 1990	**IRL**	5.0	51.4	NYSE
France growth fund	11 May 1990	**FRA**	15.3	168.4	NYSE
Singapore fund	24 July 1990	**SGP**	6.9	97.4	NYSE
China fund	7 July 1992	**CHN**	10.8	136.3	NYSE
Jardine Fleming China fund	17 July 1992	**JCHN**	9.1	114.5	NYSE
Greater China fund	17 July 1992	**GCHN**	9.6	116.2	NYSE
Japan equity fund	14 August 1992	**JPNE**	8.1	102.8	NYSE
First Israel fund	23 October 1992	**FISR**	5.0	53.7	NYSE
Total			230.7	2,657.8	
Panel B: Emerging market funds					
Mexico fund	3 June 1981	**MEX**	37.3	765.6	NYSE
Korea fund	22 August 1984	**KOR**	29.5	610.0	NYSE
Taiwan fund	16 December 1986	**TWN**	11.3	289.5	NYSE
Malaysia fund	8 May 1987	**MYS**	9.7	180.6	NYSE
Thai fund	17 February 1988	**THA**	12.2	343.8	NYSE
Brazil fund	31 March 1988	**BRA**	12.1	376.4	NYSE
India growth fund	12 August 1988	**INDG**	5.0	146.7	NYSE

(*continued*)

Table 25.1 (continued)

Fund name	IPO date	Symbol	No of shares (millions)	Net assets ($ millions)	Listing
ROC Taiwan fund	12 May 1989	**RTWN**	27.9	365.7	NYSE
Chile fund	26 September 1989	**CHL**	7.0	367.0	NYSE
Portugal fund	1 November 1989	**PRT**	5.3	75.9	NYSE
First Philippine fund	8 November 1989	**FPHI**	9.0	212.0	NYSE
Turkish investment fund	5 December 1989	**TUR**	7.0	33.5	NYSE
Indonesia fund	1 March 1990	**INDO**	4.6	42.3	NYSE
Jakarta growth fund	10 April 1990	**JAKG**	5.0	43.2	NYSE
Thai capital fund	22 May 1990	**THAC**	6.2	124.8	NYSE
Mexico equity and income fund	14 Aug 1990	**MEXE**	8.6	103.0	NYSE
Emerging Mexico fund	2 October 1990	**EMEX**	9.0	117.6	NYSE
Argentina fund	10 October 1991	**ARG**	9.2	108.1	NYSE
Korea investment fund	14 February 1992	**KORI**	6.0	87.4	NYSE
Brazilian equity fund	9 April 1992	**BRAE**	4.6	91.4	NYSE
Total			226.5	4,484.5	

This table presents the sample of closed-end country funds. The funds are classified as from developed markets or emerging markets using the World Bank's definition of emerging markets. The China funds are classified as developed market funds because the majority of their investments are in Hong Kong. The net assets reported is as of 31 December 1994. The listing of the funds is denoted as NYSE, New York Stock Exchange and AMEX, American Stock Exchange. The Germany fund, Korea fund, and the Thai capital fund also trade on the Osaka Stock Exchange. The future Germany fund is now called Central European equity fund

emerging market funds using the World Bank's definition of an emerging market. The three China funds are classified as developed market funds because the majority of their assets are in Hong Kong. The weekly (Friday closing) prices, NAVs, and corresponding dividends for each fund are obtained from *Barrons* and *Bloomberg*. [10] The returns are adjusted for stock splits and dividends. The 7-day Eurodollar deposit rate, provided by the Federal Reserve, is used as the risk-free benchmark.

[10]For some of the funds, such as the India growth fund, the prices and net asset values are as of Wednesday closing. This may lead to nonsynchronous prices and NAVs. However, as Bodurtha et al. (1995) and Hardouvelis et al. (1993) show, the effects of nonsynchronous trading are not pervasive and do not affect the analysis.

Table 25.2 Closed-end country funds: descriptive statistics

Panel A: Developed market funds							
	Price returns (%)		NAV returns (%)		Premium (%)		
Fund	Mean	Std	Mean	Std	Mean	Std	Volatility ratio
AUS	0.092	3.051	0.106	2.438	−10.182	5.172	1.565*
ITA	0.009	4.126	−0.029	2.943	−5.985	9.311	1.965*
GER	0.106	3.943	0.130	2.372	−1.120	11.316	2.762*
GBR	0.140	3.363	0.148	2.285	−11.873	5.728	2.165*
SHEL	0.330	4.482	0.259	2.017	−3.497	4.989	4.938*
SPN	−0.007	4.128	0.033	2.505	−0.076	11.923	2.716*
AUT	−0.005	3.874	−0.048	2.334	−9.860	7.709	2.754*
GERN	0.113	3.787	0.113	2.146	−15.003	6.020	3.113*
GSPN	0.239	5.159	0.143	2.675	−14.522	8.655	3.717*
GERF	0.194	3.288	0.179	2.392	−14.246	5.942	1.889*
JPNO	0.171	5.004	0.014	3.474	6.522	10.494	2.074*
GERE	0.051	3.548	0.014	2.200	−14.750	6.200	2.601*
IRL	0.261	3.220	0.189	2.226	−15.473	5.816	2.091*
FRA	0.133	3.456	0.090	2.133	−13.266	8.020	2.625*
SGP	0.261	4.238	0.134	2.599	−1.136	10.606	2.658*
Average	0.138	2.284	0.092	1.541	−8.310	4.775	1.481*
Panel B: Emerging market funds							
	Price returns (%)		NAV returns (%)		Premium (%)		
Symbol	Mean	Std	Mean	Std	Mean	Std	Volatility ratio
MEX	0.841	12.670	0.897	14.695	−5.783	8.646	0.743
KOR	0.277	4.778	0.295	3.428	18.625	11.146	1.943*
TWN	0.177	5.636	0.123	3.562	8.477	12.806	2.503*
MYS	0.308	4.581	0.251	3.233	−3.636	7.599	2.007*
THA	0.232	4.359	0.358	3.458	−3.264	10.251	1.588*
BRA	0.786	7.048	0.831	6.607	1.464	8.481	1.137*
INDG	0.297	5.518	0.130	4.144	0.833	15.397	1.773*
RTWN	0.235	5.026	0.030	2.998	0.443	8.892	2.809*
CHL	0.354	5.629	0.315	4.454	−8.589	8.363	1.597*
PRT	0.209	4.639	0.129	2.284	−6.313	7.311	4.125*
FPHI	0.489	4.490	0.359	2.589	−21.127	5.345	3.005*
TUR	0.112	6.078	0.130	7.164	13.052	16.077	0.720
INDO	0.199	5.387	0.000	2.638	15.190	10.105	4.168*
JAKG	0.249	4.987	0.045	2.308	5.221	7.947	4.666*
THAC	0.449	4.921	0.377	3.510	−8.516	6.409	1.964*
Average	0.347	2.478	0.284	1.939	0.447	4.093	1.277*

This table presents the descriptive statistics of the sample of closed-end country funds. The sample covers the period January 1991–August 1995 (244 weekly observations). All the returns are weekly returns in US dollars. The premium on a fund is calculated as the logarithm of the price-net asset value ratio. The "volatility ratio" is the ratio of variance of price returns over the variance of NAV returns for that fund. An asterisk (*) denotes that the variance of price returns is significantly higher than the variance of NAV returns, from an F-test at the 5 % level of significance

Table 25.2 reports the descriptive statistics for the sample of funds. The IFC global indices are available weekly from December 1988. Therefore, for emerging market funds which were launched before December 1988, the analysis begins in December 1988. To enable comparison across funds, a common time frame of January 1991–August 1995 is chosen. Since, the data for the first 6 months are not used in the analysis, only 30 funds listed on or before July 1990 are considered. The mean premium on the index of developed market funds is −8.31 %, whereas the mean premium on the emerging market funds is 0.44 %. A *t*-test of the difference of means indicates that the mean premium of the index of the emerging market funds is significantly higher than the mean premiums on the index of developed market funds at the 1 % level of significance.

Figure 25.1 plots the time series of the premiums for the index of developed and emerging market funds. The figure clearly indicates that the premiums on emerging market funds are usually positive and higher than the premiums on developed market funds, highlighting the effects of market segmentation. Table 25.2 also presents the ratio of the variance of the excess price returns over the variance of the excess NAV returns. For 27 funds out of the 30 funds, the ratios are significantly higher than one.[11]

A set of predetermined instrumental variables similar to instruments used in studies of predictability of equity returns for developed markets (Fama and French 1989; Ferson and Harvey 1993) and emerging markets (Bekaert 1995) are used in the empirical analysis when testing conditional asset-pricing models. The instruments are the dividend yield on the Standard and Poor's 500 Index, calculated as the last quarter's dividend annualized and divided by the current market price (DIVY); the spread between 90-day Eurodollar deposit rate and the yield on 90-day US treasury bill (TED); and the premium on an index of equally weighted country funds (FFD). Only three instruments are used in order to be parsimonious representation of the conditional asset-pricing models.

TED is calculated using data from the Federal Reserve of Chicago and DIVY is obtained from *Barrons*. FFD is constructed using the sample of 30 funds listed in Table 25.2. Table 25.3 reports the sample characteristics of these instruments. As Fama and French (1989) show, the dividend yields and interest rates track the business cycle, and expected returns vary as a function of these instruments. Table 25.4 reports the descriptive statistics for the returns on the market indices and the correlations of returns on the market indices of various countries with the returns on the US index. The developed market indices are from Morgan Stanley Capital International and the emerging market indices are from the International Finance Corporation. All the developed market indices have significantly higher correlations with the US market index, compared to the emerging markets, which implies that there are potential diversification benefits from investing in emerging markets.

[11]Assuming that the price returns and NAV returns are from two normal populations, the test statistic which is the ratio of the variances has an F distribution (if the two variances are estimated using the same sample size). The null hypothesis that the variance of price returns is greater than the variance of NAV returns is tested using this statistic at the 5 % level of significance.

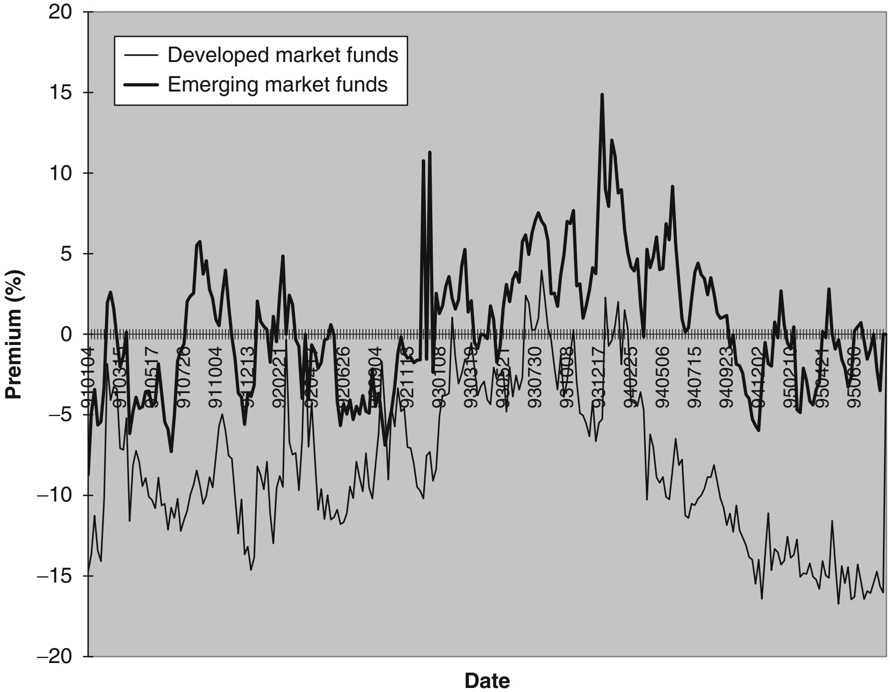


Fig. 25.1 Country fund premiums

Table 25.3 Summary statistics for the instrumental variables

Panel A: Means and standard deviations (%)			
	Mean		Std. dev
TED	0.34		0.18
DIVY	2.91		0.23
FFD	−3.87		3.93
Panel B: Correlations of the instrumental variables			
	TED	DIVY	FFD
TED	1.00		
DIVY	0.50	1.00	
FFD	−0.55	−0.25	1.00

The statistics is based on weekly data from January 1991–August 1995 (244 weekly observations). The spread between 90-day Eurodollar deposits and 90-day US treasury yields (TED), the dividend yield on the Standard and Poor's 500 index (DIVY), and an equally weighted index of the premiums for the sample of funds in Table 25.2 (FFD)

Table 25.4 Market indices: descriptive statistics

Country	Mean (%)	Std. (%)	Correlation with USA
Panel A: Developed markets			
Austria	−0.002	2.675	0.178
Hong Kong	0.485	3.366	0.174
Australia	0.227	2.212	0.292
Israel	−0.000	3.351	0.130
France	0.167	2.272	0.334
Germany	0.180	2.416	0.238
Spain	0.090	2.627	0.329
Ireland	0.209	2.675	0.299
Italy	0.085	3.558	0.146
Japan	0.117	3.006	0.180
Singapore	0.336	2.018	0.231
Switzerland	0.375	2.186	0.285
USA	0.235	1.371	1.000
UK	0.143	2.186	0.335
Europe	0.169	1.808	0.377
Panel B: Emerging markets			
Argentina	0.841	6.562	0.136
Brazil	1.145	7.577	0.221
Chile	0.687	3.180	0.125
India	0.259	4.534	−0.021
Indonesia	0.114	3.302	−0.031
Korea	0.130	3.432	0.040
Malaysia	0.410	2.802	0.163
Mexico	0.351	4.303	0.195
Philippines	0.636	3.618	0.141
Portugal	0.147	2.560	0.155
Taiwan	0.132	4.483	0.153
Thailand	0.520	3.609	0.189
Turkey	0.121	7.660	0.061
Asia	0.211	2.137	0.188
Latin America	0.554	3.314	0.640
World	0.172	1.410	0.277

This table presents the descriptive statistics for the returns on market indices of developed and emerging markets for the time period January 1991–August 1995. The table also presents the correlation of the returns, with the returns on the US market. All the developed market indices are weekly indices from Morgan Stanley Capital International. The emerging market indices are from the International Finance Corporation

25.4 Empirical Results: Pricing of Country Funds

This section presents the empirical results for pricing of country funds. The results from the unconditional tests indicate that, in general, developed market closed-end fund returns have significant risk exposures to the world market index, while emerging market closed-end fund returns have significant risk exposures to both the world market index and the corresponding local market index. Second, the hypothesis of unconditional mean-variance efficiency of the world market index cannot be rejected using the share price returns of either developed or emerging market funds and NAV returns of some of the developed market funds. However, the hypothesis of unconditional mean variance of the world market index can be rejected for emerging market NAVs and some developed market NAVs. This finding indicates that while the share prices reflect the global price of risk, the NAVs may reflect both the global and the respective local prices of risk. Tests of predictability using global instruments indicate that country fund share price and NAVs exhibit significant predictable variation. When conditional asset-pricing restrictions are examined using alternate stochastic discount factors, the results indicate that the share prices and NAVs of the closed-end country funds are not priced identically for some developed and all the Asian market funds. This finding is consistent with market segmentation. Finally, it is shown that the time-varying premiums for country funds are attributable to time-varying risk premiums. The detailed results are discussed below.

25.4.1 Unconditional Risk Exposures of Country Funds

To ascertain the risk exposures of the share price and NAV returns of the sample of country funds, the following econometric specification is employed:

$$r_{p,t} = \alpha_p + \beta_{p,w} r_{w,t} + \beta_{p,h} r_{h,t} + e_{p,t} \tag{25.16}$$

$$r_{n,t} = \alpha_n + \beta_{n,w} r_{w,t} + \beta_{n,h} r_{h,t} + e_{n,t} \tag{25.17}$$

where, for any country fund "i" (the subscript "i" has been dropped for convenience):

$r_{p,t}$ = the excess total return on the share price (including dividends) of a country fund between t−1 and t.

$r_{n,t}$ = the excess return on the NAV of a country fund between t−1 and t.

$r_{w,t}$ = the excess return on the Morgan Stanley Capital International (MSCI) world market index.

$r_{h,t}$ = the excess return on the MSCI or International Finance Corporation (IFC) global index, corresponding to the country of origin of the fund.

The coefficients $\beta_{p,w}$, $\beta_{p,h}$ are the risk exposures on the world and home market portfolios for the price returns, and $\beta_{n,w}$,$\beta_{n,h}$ are the risk exposures on the world and home market portfolios for the NAV returns. Two hypotheses are tested. The hypotheses are share price returns have significant risk exposures to the world market index and the NAV returns have significant risk exposures to both the world market index and the corresponding local market index. These hypotheses are implied by the international CAPM and its extension under market segmentation (see Errunza and Losq 1985 and Diwan et al. 1995). If assets are priced in a global environment, the country fund share price as well as NAV returns should have significant risk exposure to the world market index only. However, when there are barriers to investment, the local market factor may be a source of systematic risk for the NAVs.

The risk exposures on the prices and NAVs could be different when the prices and NAVs are imperfect substitutes. By jointly estimating Eqs. 25.16 and 25.17, the hypothesis that the risk exposures on the world market and the home market indices are identical for the price and NAV returns is tested via a Wald test which is χ^2 distributed, with degrees of freedom equal to the number of restrictions.[12] Since the local indices are significantly correlated with the world index, for ease of interpretation, they are made orthogonal to each other by regressing the local index returns on a constant and the world index return and using the residuals as the local index return. Therefore, the risk exposure on the local or regional index is the marginal exposure in the presence of the world market index.

The results from the regressions to estimate risk exposures are reported in Table 25.5. Panel A of Table 25.5 presents the results for developed market funds and panel B presents the results for the emerging market funds. The results presented in Table 25.5 indicate that 15 of the 15 developed market funds and 12 of the 15 emerging market funds price returns have significant exposure to the world index at the 5 % level of significance. Also, 12 of the developed market funds and 14 of the emerging market funds price returns have significant exposure to their corresponding local market index, in the presence of the orthogonal world market index. For the NAVs, all 15 of the developed funds and ten of the emerging market funds returns have significant exposure to the world market index. Moreover, 14 of the developed and 13 of the emerging market NAV returns have significant exposure to their local market index, in the presence of the world market index. The adjusted R-squares across all the regressions vary from a low of zero to a high of 92 %.[13]

[12]For a set of linear restrictions, the Wald test statistic is given by $W = [R\beta\text{-}r]'[R\,\mathrm{Var}(\beta)\,R'][R\beta\text{-}r]$, where β is the vector of estimated parameters and $R\beta = r$ is the set of linear restrictions.

[13]Risk exposures of the price and NAV returns were also estimated using regional indices in place of the local market indices for the funds from Europe, Latin America, and Asia. These results indicate that, out of the 12 European funds, only two funds' price returns have significant risk exposure to the MSCI Europe index in the presence of the MSCI world index. Also, out of the seven Latin American funds, all seven price returns and six NAV returns have significant exposure to the Latin American index. Also, out of the 11 Asian funds, nine price returns and eight NAV returns have significant risk exposures to the Asian index.

Table 25.5 Risk exposures of country funds

Panel A: Developed market funds										
	Coefficients on price returns (t-stats)				Coefficients on NAV returns (t-stats)				Chi-squares for Wald tests [p-values]	
Fund	α_p	$\beta_{p,h}$	$\beta_{p,w}$	Adj. R^2	α_n	$\beta_{n,h}$	$\beta_{n,w}$	Adj. R^2	χ^2_{local}	χ^2_{world}
AUS	−0.00	0.47	0.71	0.22	−0.00	0.92	0.39	0.67	6.61	24.00
	(−0.66)	(5.03)*	(5.99)*		(−1.26)	(26.79)*	(8.05)*		[0.01]	[0.00]
ITA	−0.00	0.44	0.60	0.17	−0.00	0.76	0.70	0.87	0.21	11.36
	(−0.73)	(5.18)*	(3.15)*		(−3.06)*	(33.69)*	(14.30)*		[0.64]	[0.00]
GER	−0.00	0.49	1.05	0.19	0.00	0.91	0.95	0.89	0.28	11.68
	(−0.15)	(3.89)*	(5.38)*		(0.95)	(35.56)*	(30.87)*		[0.59]	[0.00]
GBR	0.00	0.32	0.87	0.17	0.00	0.92	0.74	0.73	0.64	18.77
	(0.11)	(2.78)*	(6.33)*		(1.42)	(21.17)*	(11.44)*		[0.42]	[0.00]
SHEL	0.00	0.08	0.63	0.03	0.00	0.11	0.32	0.05	1.09	0.03
	(1.07)	(0.62)	(2.24)*		(1.41)	(1.11)	(3.07)*		[0.29]	[0.84]
SPN	−0.00	0.34	0.89	0.11	−0.00	0.80	0.82	0.68	0.06	19.63
	(−0.80)	(3.85)*	(3.27)*		(−1.79)	(20.66)*	(21.50)*		[0.80]	[0.00]
AUT	−0.00	0.30	0.97	0.14	−0.00	0.72	0.77	0.71	1.21	19.06
	(−0.79)	(3.46)*	(6.57)*		(−1.37)	(17.57)*	(9.59)*		[0.26]	[0.00]
GERN	−0.00	0.50	1.10	0.23	−0.00	0.79	0.85	0.87	1.73	4.70
	(−0.37)	(4.12)*	(6.27)*		(−1.02)	(27.79)*	(25.90)*		[0.18]	[0.03]
GSPN	0.00	0.35	0.77	0.06	0.00	0.75	0.89	0.59	0.15	22.75
	(0.44)	(3.85)*	(3.22)*		(0.11)	(16.15)*	(10.63)*		[0.69]	[0.00]
GERF	0.00	0.45	0.91	0.23	0.00	0.93	0.89	0.92	0.01	23.94
	(0.29)	(4.24)*	(6.00)*		(0.90)	(39.49)*	(30.39)*		[0.88]	[0.00]
JPNO	−0.00	0.23	1.95	0.31	−0.00	0.74	1.16	0.41	6.86	10.25
	(−0.31)	(1.82)	(6.35)*		(−0.71)	(7.71)*	(8.18)*		[0.00]	[0.00]
GERE	−0.00	0.51	0.80	0.18	−0.00	0.84	0.76	0.85	0.04	12.03
	(−0.90)	(5.87)*	(4.34)*		(−3.46)*	(34.82)*	(23.53)*		[0.83]	[0.00]

IRL	0.00	0.08	0.81	0.13	0.00	0.75	0.65	0.78	1.04	48.12
	(0.75)	(0.95)	(5.73)*		(0.33)	(21.43)*	(12.18)*		[0.30]	[0.00]
FRA	−0.00	0.56	0.79	0.21	−0.00	0.84	0.68	0.79	0.63	6.13
	(−0.20)	(5.46)*	(6.02)*		(−1.47)	(26.34)*	(17.95)*		[0.42]	[0.01]
SGP	0.00	0.63	0.93	0.15	−0.00	0.65	0.43	0.25	11.27	0.01
	(0.26)	(4.18)*	(5.65)*		(−0.14)	(6.35)*	(4.61)*		[0.00]	[0.90]
Panel B: Emerging market funds										
	Coefficients on price returns (t-stats)				Coefficients on NAV returns (t-stats)				Chi-squares for Wald tests [p-values]	
Symbol	α_p	$\beta_{p,h}$	$\beta_{p,w}$	Adj. R^2	α_n	$\beta_{n,h}$	$\beta_{n,w}$	Adj. R^2	χ^2_{local}	χ^2_{world}
MEX	0.00	0.42	0.84	0.02	0.00	0.39	0.96	0.01	0.07	0.16
	(1.36)	(2.86)*	(2.60)*		(1.12)	(2.01)*	(2.35)*		[0.78]	[0.68]
KOR	0.00	0.20	−0.09	0.01	0.00	0.00	−0.01	−0.00	3.09	0.14
	(0.71)	(1.77)	(−0.58)		(0.73)	(0.05)	(−0.09)		[0.07]	[0.70]
TWN	−0.00	0.49	0.95	0.19	−0.00	0.38	0.30	0.22	1.49	8.57
	(−0.26)	(4.85)*	(4.34)*		(−0.33)	(6.91)*	(2.26)*		[0.22]	[0.00]
MYS	0.00	0.62	1.01	0.20	0.00	0.92	0.74	0.63	8.35	0.83
	(0.60)	(5.98)*	(3.74)*		(0.89)	(23.92)*	(7.15)*		[0.00]	[0.36]
THA	0.00	0.56	1.02	0.27	0.00	0.87	0.72	0.82	18.04	3.07
	(0.26)	(7.80)*	(5.71)*		(2.27)*	(23.26)*	(13.28)*		[0.00]	[0.07]
BRA	0.00	0.50	0.90	0.31	0.00	0.64	1.07	0.58	7.38	0.29
	(1.89)	(10.84)*	(2.98)*		(2.41)*	(17.65)*	(5.64)*		[0.00]	[0.59]
INDG	0.00	0.34	−0.06	0.07	0.00	0.35	0.06	0.14	0.00	0.20
	(0.77)	(2.15)*	(−0.27)		(0.26)	(5.04)*	(0.45)*		[0.98]	[0.64]
RTWN	−0.00	0.56	1.01	0.30	−0.00	0.53	0.42	0.62	0.12	10.87
	(−0.10)	(6.47)*	(5.87)*		(−1.58)	(12.03)*	(6.17)*		[0.71]	[0.00]

(*continued*)

Table 25.5 (continued)

Panel B: Emerging market funds

	Coefficients on price returns (t-stats)				Coefficients on NAV returns (t-stats)				Chi-squares for Wald tests [p-values]	
Symbol	α_p	$\beta_{p,h}$	$\beta_{p,w}$	Adj. R^2	α_n	$\beta_{n,h}$	$\beta_{n,w}$	Adj. R^2	χ^2_{local}	χ^2_{world}
CHL	0.00	0.77	0.46	0.21	0.00	0.89	0.04	0.40	1.24	5.13
	(0.90)	(7.57)*	(2.54)*		(1.19)	(25.34)*	(0.71)		[0.26]	[0.02]
PRT	0.00	0.31	0.75	0.07	−0.00	0.73	0.58	0.70	7.67	0.76
	(0.18)	(2.40)*	(4.29)*		(−1.12)	(13.25)*	(9.52)*		[0.00]	[0.38]
FPHI	0.00	0.64	0.64	0.30	0.00	0.56	0.24	0.60	1.77	6.07
	(0.68)	(8.72)*	(3.94)*		(0.55)	(9.56)*	(2.49)*		[0.18]	[0.01]
TUR	0.00	0.03	0.34	0.00	0.00	0.02	0.08	−0.00	0.00	0.92
	(0.05)	(0.60)	(1.28)		(0.05)	(0.39)	(0.24)		[0.96]	[0.33]
INDO	0.00	0.48	0.91	0.14	−0.00	0.64	0.06	0.66	2.74	15.83
	(0.10)	(4.94)*	(4.62)*		(−1.11)	(12.85)*	(0.92)		[0.09]	[0.00]
JAKG	0.00	0.54	1.07	0.23	−0.00	0.57	0.04	0.68	0.06	28.64
	(0.26)	(4.35)*	(6.26)*		(−0.83)	(13.64)*	(0.73)		[0.79]	[0.00]
THAC	0.00	0.62	1.26	0.29	0.00	0.89	0.73	0.84	7.80	4.78
	(1.23)	(7.12)*	(5.10)*		(1.83)	(19.05)*	(15.04)*		[0.00]	[0.02]

This table presents the results from GMM estimation of risk exposures of the price ($r_{p,\ t}$) and the NAV ($r_{n.t}$) returns of country fund share and NAV excess returns on the MSCI world index ($r_{w,\ t}$)and the excess return on the corresponding local market index ($r_{h,\ t}$). For developed market funds, the time series covers the time from January 1991–August 1995. The price and the NAV risk exposures are estimated simultaneously using an exactly identified system of equations. The t-stats robust to heteroskedasticity and serial correlation (six Newey-West lags) are presented below the coefficients. An asterisk (*) denotes significance at the 5 % level of significance. The last two columns present the Chi-squares from a Wald test for the equality of the coefficients on the world market and local market indices for the price and NAV returns. The models estimated are of the form

$r_{a,t} = \alpha_a + \beta_{a,h} r_{h,.t} + \beta_{a,w} r_{w,t} + e_{a,t}$ where a = p and n

These results confirm the hypothesis that returns on prices and NAVs are generated as a linear function of the returns on the world market and possibly the corresponding local market index. The fact that 27 out of the 30 funds have significant risk exposure to the world market index indicates that the share prices may be determined via the global price of risk. The higher adjusted R-squares for the NAVs indicate that the world market index and the corresponding local market index have higher explanatory power for the time-series variations of the returns. The intercepts in the univariate regressions are not significantly different from zero, which suggests that the return on the world market index and an orthogonal return on the local market index are mean-variance efficient for the price and NAV returns.

When Wald tests are performed to test the hypothesis that the risk exposures on the prices and NAVs are identical, the results presented in the last two columns of Table 25.5 indicate that for 21 (14 from developed) out of the 30 funds, the null hypothesis of the same betas on the world index is rejected at the 10 % level of significance. Similarly, for eight funds the null hypothesis of the same betas on the local index is rejected at the 5 % level. This is a very important finding since it indicates that the systematic risks of the shares and NAVs are different. The fact that closed-end funds may have different risk exposures for the share and NAVs has been documented for domestic funds by Gruber (1996). Also different risk exposures for restricted and unrestricted securities have been documented by Bailey and Jagtiani (1994) for Thai securities and by Hietala (1989) for Finnish securities.

These results indicate that country fund share prices and NAVs may have differential risk exposures. The different risk exposures will result in different expected returns for the share prices and NAVs, which is one of the sources of the premiums. This section clearly shows that country fund share price and NAV returns have significant risk exposure to the world market index. The different risk exposures will result in different expected returns for the share prices and NAVs, which is one of the sources of the premiums. The issue of whether the country funds are priced in equilibrium via the international CAPM is the focus of next section.

25.4.2 Pricing of Country Funds in the Context of International CAPM

The results of the mean-variance efficiency of the MSCI world index, using Eq. 25.2, are presented in Table 25.6. The table reports the χ^2 and the p-values from a Wald test for the hypothesis that a set of intercepts are jointly zero.[14] This hypothesis is the restriction implied by unconditional international CAPM. Failure to reject this hypothesis would imply that the world market index is mean-variance efficient and expected returns are proportional to the expected returns on the world

[14]Other studies such as Cumby and Glen (1990) fail to reject the mean-variance efficiency of the MSCI world market portfolio for national indices of developed markets. Also, Chang et al. (1995) fail to reject the mean-variance efficiency of the MSCI world market index for a sample of developed as well as emerging market indices.

market index. The hypothesis is tested using GMM for an exactly identified system as shown in section IA. To ensure a longer time series, the tests are conducted using funds listed before December 1990.

The null hypothesis that the intercepts are jointly zero cannot be rejected for the share prices of the 15 developed market funds as well as the 15 emerging market funds. But, the null hypothesis is rejected at the 5 % level for the NAVs of both developed and emerging market funds. The tests are also conducted on subsets of funds to check the robustness of the results. Since the classification of developed and emerging markets is based on national output, it may not always capture whether a country's capital market is well developed. The subset consists of funds with zero intercepts for the NAV returns, on an individual basis. For this subset of the developed market funds consisting of 11 funds, though not reported here, the mean-variance efficiency of the world index cannot be rejected for both the share prices and NAVs. For emerging market funds NAVs, even for subsets of funds, the MSCI world market index is not mean-variance efficient. These results indicate that the world market index is an appropriate benchmark for the share prices but not necessarily for the NAVs.

These results have important implications for capital market integration. Since for the NAVs, the mean-variance efficiency of the world market index is rejected, it implies that the share prices and NAVs are priced differently. This differential pricing is sufficient to generate premiums on the share prices. Also, the fact that mean-variance efficiency of the world market index cannot be rejected for share prices of both developed and emerging markets is consistent with the theoretical prediction of Diwan et al. (1995). Since the country fund share prices are priced with complete access to that market, in equilibrium, their expected returns should be proportional to their covariances with the world market portfolio. If the NAVs are priced with incomplete access, the world market portfolio is not mean-variance efficient with respect to those returns. The results in this section are consistent with this notion. As the earlier section shows, the risk exposures of the share prices and NAVs may differ. If the international CAPM is a valid model for the share prices and NAVs, as outlined in section IB, this differential risk exposure can explain cross-sectional variations in premiums. The next section analyzes the effect of the differential risk exposures of the share prices and NAVs on the country fund premiums.

25.4.3 Cross-Sectional Variations in Country Fund Premiums

This section presents the results for tests of the hypothesis that cross-sectional variation in country fund premiums is positively related to the differences in the risk exposures of the share prices and NAVs. The theoretical motivation for this was presented in section IC. Equation 25.10 is estimated for each week during the last 75 weeks of the sample and the parameter estimates are averaged. The results indicate that cross-sectional variation in country fund premiums is explained by the

Table 25.6 Pricing of country funds in the context of international CAPM

Test assets	Chi-squares for price returns [p-value]	Chi-squares for NAV returns [p-value]
Developed market country funds (15 funds)	15.71 [0.40]	44.48 [0.00]
Emerging market country funds (15 funds)	10.64 [0.77]	25.47 [0.04]

This table presents the Chi-squares for the null hypothesis that the intercepts are jointly zero for a set of assets when their excess returns are regressed on the excess return on the MSCI world index. The regression estimated is

$r_{i,t} = \alpha_i + \beta_i r_{w,t} + e_{i,t},\ i = 1.... N$

The null hypothesis H_o: $\alpha_i = 0$ for all $i = 1...N$ is tested by a Wald test using GMM estimates from an exactly identified system. The p-values are based on standard errors robust to heteroskedasticity and serial correlation (13 Newey-West lags). The sample covers January 1991–August 1995. The test assets are excess price/NAV returns of country funds

The composition of the test assets are as follows – the developed market funds are AUT, AUS, FRA, GER, GERE, GERF, GERN, GSPN, IRL, ITA, JPNO, SGP, SPN, SHEL, and GBR. The emerging market funds are BRA, CHL, FPHI, INDG, INDO, JAKG, KOR, MYS, MEX, PRT, RTWN, TWN, THAC, THA, and TUR

differential risk exposures of a closed-end country fund share price and NAV returns. The estimation proceeds are as follows: after the betas are estimated using 75 weeks of data prior to the last 75 weeks of the sample, the difference in the betas on the world index is used as an explanatory variable in regressions with the premiums as the dependent variable.

The results are reported in Table 25.7. The results indicate that, for the full sample, the difference in risk exposure is significant and the adjusted R-square is 9.5 %. For the developed market funds, the results in panel B indicate that the differences in the betas on the world index have significant explanatory power for the cross-section of premiums for the developed market funds, with an adjusted R-square of 18.7 %. For the emerging market funds, however, as reported in panel C, the differences in risk exposures are not significant, and the adjusted R-square is only 1.9 %. This result is not surprising, given the result of the previous section that the world market index is not mean-variance efficient for the emerging market NAVs. When a dummy variable which takes value of one for the developed markets is added, it is highly negatively significant, indicating that the premiums for developed market funds are significantly lower than the emerging market fund premiums. The differential risk exposure is however significant only at the 10 % level, when the dummy variable is added.

As an extended specification, measures of spanning, integration, and substitution (based on prior research of Errunza et al. 1998) are used as additional explanatory variables. When measures of spanning, substitution, and access are used as additional explanatory variables, the adjusted R-squares go up to 39.5 % for the full sample. The measure of spanning has a positive sign as expected. The measure of spanning is also significant across the developed market funds. However, contrary

Table 25.7 Cross-sectional variations in country fund premiums

CONST	DIF	SPN	SUB	ACC(d)	ACC (CF)	Adj. R^2
Panel A: Full sample						
−0.05	0.08					0.095
(1.60)	(2.85)*					
−0.15	0.06	63.77	0.03	−0.02		0.395
(0.07)	(2.00)*	(2.04)*	(1.55)	(1.40)		
−0.17	0.07	69.48	0.03		−0.00	0.383
(2.79)*	(2.00)*	(2.28)*	(1.60)		(0.02)	
−0.16		72.59	0.03			0.300
(2.73)*		(2.26)*	(1.73)			
−0.00	0.05			−0.08		0.184
(0.18)	(2.03)*			(2.85)*		
Panel B: Developed market funds						
−0.09	0.11					0.187
(2.51)*	(4.33)*					
−0.29	0.12	207.76	0.05			0.551
(4.23)*	(4.03)*	(2.05)*	(2.27)*			
−0.29	0.12	210.39	0.08		0.00	0.565
(4.28)*	(3.81)*	(2.06)*	(2.12)*		(0.00)	
−0.28		234.11	0.04			0.370
(4.07)*		(2.12)*	(2.14)*			
Panel C: Emerging market funds						
0.00	0.01					0.019
(0.25)	(0.40)					
−0.10	0.02	53.72	0.02			0.395
(1.21)	(0.60)	(1.75)**	(0.86)			
−0.01	0.02	49.56	0.01		0.03	0.383
(0.95)	(0.51)	(1.59)	(0.95)		(0.15)	
−0.09		50.91	0.01			0.301
(1.11)		(1.64)	(0.86)			

Results from an OLS regression of premiums on country funds on DIF (the difference between the betas on the world market index, for the price returns and NAV returns). The coefficient ACC (CF) is a measure of access proxied by the total purchase of securities by US residents (divided by the global purchase) from the country of origin of the fund. The coefficient ACC(d) is a dummy variable that takes value one for developed markets and zero for emerging markets. The coefficients below are the averages from the regression of the premiums on the independent variables for each week for the last 75 weeks of the sample (Mar 94–Aug 95). The t-statistics are presented in the parenthesis * and ** denote significance at the 5 % and 10 % levels of significance, respectively. The spanning measure (SPN) is the conditional variance of the NAV return of a fund, unspanned by the US market return and the fund's share price returns, with specification as follows:

$r_{n,t} = \alpha_i + \beta_i r_{US,t} + + \beta_i r_{p,t} + e_{i,t}$

where $e_{i,t}$ has a GARCH(1, 1) specification. The measure of substitution (SUB) is the ratio of conditional volatilities of the share price and NAV returns of a fund not spanned by the US market. The specification are

$r_{n,t} = \alpha_i + \beta_i r_{US,t} + e_{n,t}, \ r_{p,t} = \alpha_i + \beta_i r_{US,t} + e_{p,t}$

The error terms $e_{n,t}$ and $e_{p,t}$ have GARCH (1, 1) specifications as follows (h_t is variance of the error term):

$h_t = \theta_0 + \theta_1 e^2_{i,t-1} + \theta_2 h_{t-1}$

The equation for the cross-sectional regressions is

$Prem_i = \delta_0 + \delta_1 DIF_i + \delta_2 SPN_i + \delta_3 SUB_i + \delta_4 ACC_i + e_i$

to expectation, the measure of spanning is not significant across emerging market funds. This could be due to very low variability of the variables within emerging markets. The measure of substitutability, which is the ratio of conditional variances of the price returns and NAV returns, is not significant. The measure of substitutability persists to be insignificant for the subsets of emerging market funds. The measure of access proxied by purchase of securities by US residents from that market is not significant.

These results indicate that differential risk exposures and measures of access and spanning are significant determinants of cross-sectional variation in closed-end country fund premiums. Also, the measures of segmentation explain premiums better across developed and emerging markets compared to only within emerging markets. The phenomenon that differential risk exposures can explain cross-sectional variations in premiums on unrestricted securities relative to securities restricted to only local investors has been previously documented for Thai equities by Bailey and Jagtiani (1994) and for Swiss equities by Stulz and Wasserfallen (1995). Therefore, both market segmentation and other factors which make the closed-end country fund shares and the underlying securities imperfect substitutes – which is the source of the differences in the risk exposures and the excess volatility of the share prices – account for the cross-sectional variations in the premiums. The significance of the difference in risk exposures indicates that the greater the difference in risk exposures across a sample of country funds, the higher the premium. Second, the premium is higher for country funds originating from countries whose capital markets are less integrated with the US capital market.

Although an important finding, the differential risk exposures cannot explain the time variation in premiums. The academic literature has documented that the country fund premiums vary over time. Figure 25.1 illustrates the time variation of country fund premiums. Section IIIB reported the results for pricing of country funds in the context of the unconditional international CAPM which indicated that country fund share prices are priced consistent with the international CAPM, whereas the NAVs may or may not be priced via the international CAPM. Unconditional mean-variance efficiency of a benchmark portfolio implies conditional mean-variance efficiency, but not vice versa (Ferson 1995). If expected returns conditional on an information set are also different for the share prices and the NAVs, it could explain not only the premiums but also their variation over time, which is further explored in section IIIF. To analyze the time variability of expected returns, the predictability of the funds share price and NAV returns is examined in the next section.

25.4.4 Predictability of Closed-End Country Fund Returns

To assess the predictability of price and NAV returns, 4-week cumulative returns (the sum of returns over a 4-week period beginning the next week) are used, since most studies of predictability have used a monthly horizon. Predictability of the returns would imply that expected returns vary over time due to rational variation in risk premiums. To test the hypothesis of predictability, the returns for both share prices and NAVs are regressed on a set of global and fund-specific instruments, to ascertain the predictive power of these instruments. The global instruments are TED and DIVY. The fund-specific instrument is the lagged premium on an equally weighted index of all the funds in the sample (FFD). The lagged premium was found to have predictive power for price returns by Bodurtha et al. (1995). Also, Errunza et al. (1998) find that when time series of premiums is regressed on a global fund premium, it is highly significant. However, unlike Errunza et al. (1998) here the lagged premium is used. Significance of this variable would indicate that investors use information about past premiums to form expectations about future prices. The regression is of the form

$$\sum_{t=t+1}^{t=t+4} r_{a,t} = \eta_0 + \eta_1 FFD_t + \eta_2 TED_t + \eta_3 DIVY_t + e_{a,t} \tag{25.18}$$

where $r_{a,t}$ is the excess return on the price or NAV of an equally weighted portfolio of developed or emerging market funds.

The results from the regressions are presented in Table 25.8. Wald tests of no predictability, which test the hypothesis that a set of coefficients are jointly zero, indicate that the fund factor predicts price returns for 4 funds only and it also predicts 15 funds' NAV returns. The predictive power of the lagged premiums for price returns has been attributed to noise trading (see Bodurtha et al. 1995). The fact that it predicts NAV returns may indicate that noise trading is not an important determinant of expected returns of share prices of closed-end funds. When global instruments are used, the Wald test of no predictability can be rejected at the 5 % level for only five funds' price returns and 15 funds' NAV returns. Overall, using all the instruments, the null hypothesis of no predictability can be rejected for six share prices and 20 NAVs. The poor predictability of the share prices is puzzling, since the share prices are determined in the USA and the instruments used have been shown to predict returns in the USA.

In summary, closed-end country fund price and NAV returns exhibit considerable predictability. The existence of predictable variation in returns is interpreted as evidence of time-varying expected returns. The differences between the predictability of prices and NAVs also imply that they are not priced identically, which is further examined in the next section using time-varying expected returns.

25.4.5 Conditional Expected Returns and Pricing of Country Funds

Table 25.9 reports the results of estimating the SDFs specified in Eqs. 25.11 and 25.12. The models in panels A and B correspond to the international CAPM and a two-factor model, with the world market return and the regional returns for Latin America or Asia as the factors. For both the models,

Table 25.8 Predictability of country fund returns

Panel A: Developed market funds						
	Chi-squares [p-values]			Chi-squares [p-values]		
Fund	Fund factor	Global instruments	All instruments	Fund factor	Global instruments	All instruments
AUS	0.11	1.93	1.95	8.23	3.44	9.24
	[0.73]	[0.38]	[0.58]	[0.00]	[0.17]	[0.02]
ITA	0.71	1.30	1.36	5.86	7.27	9.19
	[0.39]	[0.51]	[0.71]	[0.01]	[0.02]	[0.02]
GER	4.91	2.51	8.67	5.18	1.98	8.04
	[0.02]	[0.28]	[0.03]	[0.02]	[0.37]	[0.04]
GBR	0.07	1.23	2.58	3.33	5.64	5.98
	[0.78]	[0.53]	[0.45]	[0.06]	[0.05]	[0.11]
SHEL	1.81	1.43	2.51	0.05	1.86	1.86
	[0.17]	[0.48]	[0.47]	[0.81]	[0.39]	[0.60]
SPN	4.96	6.19	7.06	10.01	16.95	18.88
	[0.02]	[0.04]	[0.06]	[0.00]	[0.00]	[0.00]
AUT	0.02	0.01	0.03	1.86	3.25	6.03
	[0.86]	[0.99]	[0.99]	[0.17]	[0.19]	[0.10]
GERN	1.67	0.10	2.62	3.33	2.13	6.26
	[0.19]	[0.95]	[0.45]	[0.06]	[0.34]	[0.09]
GSPN	1.36	2.53	2.61	4.58	11.22	11.29
	[0.24]	[0.28]	[0.45]	[0.03]	[0.00]	[0.01]
GERF	0.20	0.29	1.42	3.58	2.91	5.87
	[0.64]	[0.86]	[0.69]	[0.05]	[0.23]	[0.11]
JPNO	7.12	3.18	8.73	13.96	5.20	19.60
	[0.00]	[0.20]	[0.03]	[0.00]	[0.07]	[0.00]
GERE	0.14	0.03	0.44	4.69	1.32	8.12
	[0.70]	[0.98]	[0.93]	[0.03]	[0.51]	[0.04]
IRL	0.00	1.93	2.15	1.80	14.56	14.58
	[0.95]	[0.38]	[0.54]	[0.17]	[0.00]	[0.00]
FRA	3.56	1.52	5.02	0.00	0.23	0.25
	[0.05]	[0.46]	[0.17]	[0.94]	[0.89]	[0.96]
SGP	0.00	2.55	2.55	11.05	1.37	11.28
	[0.96]	[0.27]	[0.46]	[0.00]	[0.50]	[0.01]
Panel B: Emerging market funds						
	Chi-squares [p-values]			Chi-squares [p-values]		
Symbol	Fund factor	Global instruments	All instruments	Fund factor	Global instruments	All instruments
MEX	0.85	5.65	5.82	0.53	6.62	7.24
	[0.35]	[0.05]	[0.12]	[0.46]	[0.03]	[0.06]
KOR	0.74	1.04	2.00	4.20	8.87	13.46
	[0.38]	[0.59]	[0.57]	[0.04]	[0.01]	[0.00]
TWN	0.06	3.15	3.67	19.77	16.71	25.93
	[0.79]	[0.20]	[0.29]	[0.00]	[0.00]	[0.00]

(*continued*)

Table 25.8 (continued)

Panel B: Emerging market funds						
	Chi-squares [p-values]			Chi-squares [p-values]		
Symbol	Fund factor	Global instruments	All instruments	Fund factor	Global instruments	All instruments
MYS	0.20	1.89	2.83	1.91	2.21	3.47
	[0.64]	[0.38]	[0.41]	[0.16]	[0.33]	[0.32]
THA	0.03	1.67	1.85	4.07	4.92	7.24
	[0.85]	[0.43]	[0.60]	[0.04]	[0.08]	[0.06]
BRA	0.03	6.34	6.99	0.32	4.90	5.39
	[0.84]	[0.04]	[0.07]	[0.56]	[0.08]	[0.14]
INDG	0.47	2.03	2.03	7.50	3.79	24.21
	[0.49]	[0.36]	[0.56]	[0.00]	[0.15]	[0.00]
RTWN	1.72	15.45	15.58	19.35	15.42	24.72
	[0.18]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
CHL	0.05	4.94	6.97	1.78	7.32	7.34
	[0.80]	[0.08]	[0.07]	[0.18]	[0.02]	[0.06]
PRT	1.17	0.45	2.41	6.56	0.01	9.51
	[0.27]	[0.79]	[0.49]	[0.01]	[0.99]	[0.02]
FPHI	0.00	3.30	3.43	1.03	2.91	5.61
	[0.97]	[0.19]	[0.32]	[0.31]	[0.23]	[0.13]
TUR	1.37	1.11	3.45	0.35	0.34	1.42
	[0.24]	[0.57]	[0.32]	[0.55]	[0.83]	[0.70]
INDO	2.06	1.09	2.26	2.70	7.81	7.90
	[0.15]	[0.57]	[0.51]	[0.09]	[0.02]	[0.04]
JAKG	0.28	0.87	1.35	1.75	19.36	19.93
	[0.59]	[0.64]	[0.71]	[0.18]	[0.00]	[0.00]
THAC	0.03	4.22	4.79	1.98	2.33	3.53
	[0.84]	[0.12]	[0.18]	[0.15]	[0.31]	[0.31]

Table presents the Chi-squares from a Wald test to ascertain importance of the fund factor and global instruments in predicting the 4-week ahead cumulative returns of the prices and NAVs. The fund factor FFD is the lagged premium on an equally weighted index of all country funds in the sample. The global instruments include the lagged spread on 90-day Eurodollar deposits and 90-day US treasury yields (TED) and the lagged dividend yield on the S&P 500 index (DIVY). The estimates are obtained via GMM using an exactly identified set of moment conditions. The table below reports Chi-squares for a Wald test that a set of one or more instruments is zero. The p-values are robust to heteroskedasticity. The model estimated is as follows, where

$$\sum_{t=t+1}^{t=t+4} r_{a,t} = \eta_0 + \eta_1 FFD_t + \eta_2 TED_t + \eta_3 DIVY_t + e_{a,t}$$

the conditional restrictions are tested. As noted earlier, if expected returns vary over time due to rational variation in risk premiums, conditional expected returns should be used. The conditional restrictions are tested by using a set of lagged instrumental variables TED and DIVY as predictors of excess returns. The tests use sets of assets – two sets of four developed market funds, a set of three Latin American funds, and two sets of three Asian funds. To ensure an adequate time series, the funds selected are the ones listed prior to December 1989. Also, using too many assets would result in too many moment conditions.[15]

The funds selected are listed in Table 25.9. After estimating a model using GMM, its goodness-of-fit is ascertained by the J-statistic. The J-statistic is a test of the overidentifying restrictions of the model. The p-values given below the J-statistic indicate that none of the models are rejected for any test assets, at conventional significance levels. This finding is consistent with the idea that for a group of assets, there may be a number of SDFs that satisfy the pricing relation. Also, failure to reject the conditional international CAPM even for the NAVs indicates that while the unconditional international CAPM is not a valid model for emerging market funds and some developed market funds, the conditional international CAPM may be a valid model. This finding is consistent with the finding of Buckberg (1995) who fails to reject the conditional international CAPM for a set of emerging market indices. It must be noted that, unlike traditional conditional models, using stochastic discount factors implies a nonlinear relation between asset returns and the returns on the market index. Failure to reject such specifications may indicate that nonlinear specifications perform better in explaining equity returns.

Table 25.9 also reports the results for tests of hypothesis that the coefficients of the SDFs of the price returns and the NAV returns are identical. The p-values for the Wald tests indicate that the null hypothesis can be rejected at all conventional levels for a subset of the developed market funds and all the Asian funds. This is a striking result, since it indicates that, although the same factors may be priced for the share prices and NAVs, the factors are priced differently. In traditional asset-pricing framework, it is equivalent to saying that the risk premiums are different. This result is consistent with the previous finding of Harvey (1991) who reports that for a sample of industrial markets, although the conditional international CAPM cannot be rejected, the price of risk varies across countries. Also, De Santis (1995) finds that stochastic discount factors that can price developed market indices cannot price emerging market indices. The difference in the price of risk is one of the sources of the premiums for the country funds. When the two-factor model is used, again the tests reject the hypothesis that the coefficients are identical for the Asian funds. However, the hypothesis of identical coefficients for the price and NAV returns cannot be rejected for a set of developed market funds and the Latin American funds.

The above results imply that a subset of the developed markets and the Asian market are segmented from the US market. However, a subset of the developed markets and the Latin American markets are integrated with the US market. The results for Latin America are not affected, when tests are conducted excluding the time period surrounding the Mexican currency crisis. Therefore, differential risk exposures are more important in explaining the premiums for the Latin American markets and the developed markets that are integrated with the world market. The results in this section clearly show that while some markets may be integrated

[15]Cochrane (1996) shows that iterated GMM estimates behave badly when there are too many moment conditions (37 in his case).

Table 25.9 GMM estimation of stochastic discount factors

	Price returns			NAV returns						
Test assets	λ_0	λ_1	λ_2	λ_0	λ_1	λ_2	DF	J [p-value]	J1 [p-value]	J2 [p-value]
Developed market funds			**International CAPM**							
AUT, GBR, GER, ITA	2.14 (4.02)*	−1.14 (−2.15)*		2.19 (5.51)*	−1.29 (−3.11)*		20	11.98 [0.91]	0.53 [0.76]	0.07 [0.79]
SPN, SHEL, AUS, AUT	2.81 (−4.46)*	−1.81 (−2.88)*		0.81 (1.64)	0.18 (0.37)		20	11.23 [0.94]	8.38 [0.01]	7.72 [0.00]
Asian funds										
KOR, MYS, TWN	4.27 (4.08)*	−3.27 (−3.13)*		2.95 (3.56)*	−1.95 (−2.36)*		14	8.11 [0.88]	5.25 [0.07]	6.38 [0.01]
THA, INDG, FPHI	−0.99 (−0.74)	1.99 (1.50)		−3.42 (−2.20)*	4.42 (2.84)*		14	9.28 [0.81]	11.21 [0.00]	10.57 [0.00]
Latin American funds										
BRA, CHL, MEX	4.46 (2.45)*	−3.46 (−1.90)*		3.82 (2.37)*	−2.82 (−1.75)		14	9.80 [0.77]	0.42 [0.80]	0.36 [0.54]

Two−factor model										
Asian funds										
KOR, MYS, TWN	−2.35 (−2.92)*	3.54 (4.06)*	−1.16 (−1.48)	−3.10 (−3.63)*	4.10 (4.80)*	−1.19 (−1.74)	18	11.51 [0.87]	3.12 [0.26]	2.94 [0.08]
THA, INDG, FPHI	3.00 (5.93)*	−2.00 (−3.95)*	−1.35 (−1.63)	1.88 (3.86)*	−0.88 (−1.81)	−1.50 (−1.84)	18	10.66 [0.90]	8.16 [0.04]	7.65 [0.00]
Latin American funds										
BRA, CHL, MEX	0.70 (0.57)	0.29 (0.24)	−1.09 (−3.27)*	0.83 (0.98)	0.16 (0.19)	−0.76 (−3.97)*	18	12.64 [0.81]	2.27 [0.51]	0.02 [0.88]

Estimation of stochastic discount factors, based on the equation $E[M(t,1)\ R(t,1)|Z_t] = 1$ via GMM. The sample covers January 1990–August 1995. The instrument set Z_t includes a constant, the lagged spread on 90-day Eurodollar deposits and 90-day US Treasury yields (LTED), and the dividend yield on Standard and Poor's 500 index (LDIVY). The estimations with the returns on regional indices (R_h) also use the lagged regional return (for Asia or Latin America) as an instrument

The stochastic discount factors estimated are

ICAPM: $M = \lambda_0 + \lambda_1 R_w$

Two-factor: $M = \lambda_0 + \lambda_1 R_w + \lambda_2 R_h$

DF is the degrees of freedom for Hansen's test of the overidentifying restrictions (the J-statistic), which has an χ^2 distribution. The t-stats for the coefficients are given in parenthesis (an asterisk indicates significance at the 5 % level) and the p-values for the J-stat is given in brackets. The last two columns report Chi-squares and p-values for tests of hypotheses that the coefficients of the price returns and NAV returns are identical. J1 is for a joint test for all the coefficients and J2 is a test for the coefficients on the world market index with the world capital market, others are not. If capital markets are segmented, that is sufficient to generate the premiums on the share prices. However, if markets are integrated, differential risk exposures may explain the premiums. This finding is different from the existing literature on country fund pricing, which has attributed the existence of premiums on irrational factors such as noise trading. The preceding analysis clearly shows that differential pricing of the same factors or differential risk exposures on the same factor may lead to premiums on the share prices. The next section provides an analysis of the effect of time-varying expected returns on country fund premiums.

25.4.6 Conditional Expected Returns and Time-Varying Country Fund Premiums

Table 25.9 reports the results of regressing the premiums on the country funds on the differences in conditional risk exposures of share price returns and NAV returns estimated using Eqs. 25.13, 25.14, and 25.15. The conditional risk exposures are estimated as a linear function of lagged dividend yields and the term premium. The null hypothesis is that the coefficient of the difference in the conditional betas is significant. The results indicate that the difference in conditional betas is highly significant in explaining the time-varying premiums on country funds. For 24 out of the 30 funds, the coefficient is significant. Also, the adjusted R-squares are high, especially for the developed market funds. For many of the funds, however, the coefficient is negative, implying that using the world market index alone is not sufficient to explain the return-generating process and the returns on the NAVs may reflect the local price of risk.

The adjusted R^2 values are higher for the developed market funds. This result is consistent with earlier results which show that the world market index is not an appropriate benchmark for emerging market NAVs. Also, majority of the emerging market funds are from Asia. The results in Table 25.9 indicated that these markets are segmented from the world market. Therefore, different risk exposures to the world market index do not explain much of the variation in the premiums for the emerging markets.

This is an important finding since the existing literature on closed-end funds has attributed the existence of premiums to irrational factors, such as noise trading. The preceding analysis shows clearly that segmented capital markets in which risk premiums vary over time are sufficient to generate two different prices for the same set of cash flows. Also, the differences in prices may vary over time as expected returns vary over time due to rational variations in expected returns. If the price of risk is different across markets, the same security will have different expected returns. The difference in these expected returns results in the premium on the share price. If the expected returns vary over time because of rational variation in risk premiums, the premiums will also vary over time as a function of the differential expected returns (Table 25.10).

25.5 Conclusions

Closed-end country funds are becoming an increasingly attractive source of capital in international equity markets. This paper provides an empirical analysis of the pricing of these funds in the context of rational asset-pricing models. The paper finds that differential risk exposures, market segmentation, and time-varying risk premiums play important roles in the differential pricing of the share prices and NAVs. Based on the unconditional international CAPM, the existence of premiums can be attributed to different risk exposures, as is the case with developed market funds and differential pricing of the shares and NAVs, as is the case with emerging market funds and some developed market funds.

The paper also analyzes the pricing of country funds when conditioning information is allowed. The results indicate that, for alternate stochastic discount factors, for majority of the funds, the pricing of country fund shares and NAVs is consistent with the conditional international CAPM. However, tests of the estimated stochastic discount factors indicate that, for a subset of the developed market funds and all the Asian funds, closed-end country fund share prices and NAVs are priced differently. This result indicates that the international capital markets are not fully integrated. Finally, this paper shows that the premiums on country funds vary over time because of time variation in expected returns.

The findings in this paper have several possible extensions. The focus of the paper has been on rational explanations for country fund premiums based on differential risk exposures and market segmentation effects. It will be interesting to extend this analysis and examine the effect of other factors such as taxes, numeraires, liquidity, and bid-ask spreads.

Appendix 1: Generalized Method of Moments (GMM)

GMM is an econometric method that was a generalization of the method of moments developed by Hansen (1982). The moment conditions are derived from the model. Suppose Y_t is a multivariate independently and identically distributed (i.i.d) random variable. The econometric model specifies the relationship between Z_t and the true parameters of the model (θ_0). To use GMM there must exist a function $g(Z_t, \theta_0)$ so that

$$m(\theta_0) \equiv E[g(Z_t, \theta_0)] = 0 \tag{25.19}$$

In GMM, the theoretical expectations are replaced by sample analogs:

$$f(\theta, Z_t) = 1/T \sum g(Z_t, \theta). \tag{25.20}$$

The law of large numbers ensures that the RHS of above equation is the same as

$$E[f(Z_t, \theta_0)]. \tag{25.21}$$

Table 25.10 Time-varying closed-end country fund premiums

Panel A: Developed market funds			
Fund	γ_0	γ_1	Adj. R^2
AUS	−0.07 (−10.13)*	0.11 (5.04)*	0.08
ITA	−0.07 (−13.05)*	0.10 (5.86)*	0.19
GER	−0.01 (−1.55)*	−0.00 (−0.29)	−0.00
GBR	−0.11 (−34.40)*	0.08 (8.53)*	0.22
SHEL	−0.06 (−13.38)*	−0.10 (−7.12)*	0.13
SPN	−0.00 (−0.14)	−0.18 (−4.89)*	0.19
AUT	−0.03 (−3.47)*	0.21 (7.47)*	0.13
GERN	−0.14 (−37.73)*	0.04 (3.45)*	0.13
GSPN	−0.15 (−35.72)*	0.13 (7.13)*	0.22
GERF	−0.14 (−40.05)*	0.03 (2.47)*	0.01
JPNO	0.09 (9.44)*	0.04 (3.79)*	0.05
GERE	−0.15 (−39.20)*	0.04 (4.54)*	0.06
IRL	−0.16 (−44.37)*	−0.13 (−5.72)*	0.12
FRA	−0.15 (−24.47)*	−0.24 (−5.22)*	0.21
SGP	−0.37 (−2.03)*	−0.72 (−1.96)	0.02
Panel B: Emerging market funds			
Fund	γ_0	γ_1	Adj. R^2
MEX	−0.07 (−16.93)*	0.08 (6.66)*	0.24
KOR	0.21 (26.45)*	−0.15 (−8.42)*	0.20
TWN	−0.07 (−1.01)	−0.26 (−2.28)*	0.01
MYS	−0.03 (−6.67)*	0.02 (2.25)*	0.02
THA	0.00 (0.04)	0.08 (0.43)	−0.00
BRA	0.01 (2.44)*	0.05 (4.81)*	0.09
INDG	0.00 (0.61)	0.01 (0.83)	−0.00
RTWN	−0.60 (−0.76)	0.00 (1.46)	0.01
CHL	−0.34 (−6.76)*	−0.62 (−5.14)*	0.20
PRT	−0.06 (−13.60)*	−0.03 (−2.47)*	0.02
FPHI	−0.20 (−25.86)*	0.01 (0.83)	−0.00
TUR	0.15 (11.34)*	0.10 (2.16)*	0.01
INDO	0.23 (7.85)*	0.10 (2.97)*	0.07
JAKG	0.38 (8.91)*	0.34 (7.76)*	0.27
THAC	−0.22 (−4.97)*	−0.24 (−3.11)*	0.03

Results from an OLS regression of premiums on country funds on the difference between the time-varying betas on the world market index. The t-statistics robust to heteroskedasticity are presented in the parenthesis. An asterisk (*) denotes significance at the 5 % level of significance. The regression estimated is of the form

$$\mathrm{Prem}_{t+1} = \gamma_0 + \gamma_1[\beta_{n,w}(Z_t) - \beta_{p,w}(Z_t)] + e_{prem,t+1}$$

The time-varying betas are estimated using the conditional international CAPM for an index of developed and emerging market funds price and NAV returns, using the equations

$$r_{a,t+1} = \alpha_a + \beta_{a,w} r_{w,t+1} + \sum_i \gamma_{a,i}\left(r_{w,t+1} * z_t\right) + e_{a,t+1} \quad a = p \text{ and } n$$

where $r_{p,t+1}$ is the excess return on the share price and $r_{n,t+1}$ is the excess NAV return and Prem $_{t+1}$ is the premium for the time period January 1991–August 1995. $r_{w,t+1}$ is the excess return on the world market index and Z_t is a set of instrumental variables DIVY and TED

The sample GMM estimator of the parameters may be written as (see Hansen 1982)

$$\Theta = \arg\min \left[1/T \sum g(Z_t, \theta)\right]' W_T 1/T \sum g(Z_t, \theta)\Big] \tag{25.22}$$

So essentially GMM finds the values of the parameters so that the sample moment conditions are satisfied as closely as possible. In our case for the regression model,

$$y_t = X_t{}'\beta + \varepsilon_t \tag{25.23}$$

The moment conditions include

$$E[(y_t - X_t{}'\beta)x_t] = E[\varepsilon_t x_t] = 0 \text{ for all t} \tag{25.24}$$

So the sample moment condition is

$$1/T \sum (y_t - X_t{}'\beta)x_t$$

and we want to select β so that this is as close to zero as possible. If we select β as $(X'X)^{-1}(X'y)$, which is the OLS estimator, the moment condition is exactly satisfied. Thus, the GMM estimator reduces to the OLS estimator and this is what we estimate. For our case the instruments used are the same as the independent variables. If, however, there are more moment conditions than the parameters, the GMM estimator above weighs them. These are discussed in detail in Greene (2008, Chap. 15). The GMM estimator has the asymptotic variance

$$\left(X'Z\,(Z'\Omega Z)^{-1}Z'X\right)^{-1} \tag{25.25}$$

The White robust covariance matrix may be used for Ω as discussed in appendix C when heteroskedasticity is present. Using this approach, we estimate GMM with White heteroskedasticity consistent t-stats.

References

Bailey, W., & Jagtiani, J. (1994). Foreign ownership restrictions and stock prices in the Thai Capital Market. *Journal of Financial Economics, 36*, 57–87.

Bailey, W., & Lim, J. (1992). Evaluating the diversification benefits of the new country funds. *Journal of Portfolio Management, 18*, 74–80.

Bansal, R., Hsieh, D. A., & Viswanathan, S. (1993). A new approach to international arbitrage pricing. *Journal of Finance, 48*, 1719–1747.

Bekaert, G. (1995). Market integration and investment barriers in equity markets. *World Bank Economic Review, 9*, 75–107.

Bekaert, G., & Urias, M. (1996). Diversification, integration and emerging market closed-end funds. *Journal of Finance, 51*, 835–869.

Bodurtha, J. N., Jr., Kim, D.-S., & Lee, C. M. C. (1995). Closed-end country funds and U.S. market sentiment. *Review of Financial Studies, 8*, 881–919.

Bonser-Neal, C., Brauer, G., Neal, R., & Wheatly, S. (1990). International investment restrictions and closed-end country fund prices. *Journal of Finance, 45*, 523–547.

Chang, E., Eun, C., & Kolodny, R. (1995). International diversification through closed-end country funds. *Journal of Banking and Finance, 19*, 1237–1263.

Cochrane, J. H. (1996). A cross-sectional test of an investment-based asset pricing model. *Journal of Political Economy, 104*(3), 572–621.

Cumby, R. E., & Glen, J. D. (1990). Evaluating the performance of international mutual funds. *Journal of Finance, 45*(2), 497–521.

De Santis, G. (1995). Volatility bounds for stochastic discount factors: Tests and implications from international financial markets. Working paper, University of Southern California, Jan.

Diwan, I., Vihang, E., & Senbet, L.W. (1995). Diversification benefits of country funds. In M. Howell (Ed.), *Investing in emerging markets*. Washington, DC: Euromoney/World Bank.

Errunza, V., & Losq, E. (1985). International asset pricing under mild segmentation: Theory and test. *Journal of Finance, 40*, 105–124.

Errunza, V., Senbet, L. W., & Hogan, K. (1998). The pricing of country funds from emerging markets: Theory and evidence. *International Journal of Theoretical and Applied Finance, 1*, 111–143.

Fama, E. F., & French, K. R. (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics, 25*, 23–49.

Ferson, W. (1995). Theory and empirical testing of asset pricing models. In R. Jarrow, V. Maksimovic, & W. T. Ziemba (Eds.), *Handbooks in operation research and management science* (Vol. 9). Amsterdam: North-Holland.

Ferson, W., & Foerster, S. (1994). Finite sample properties of generalized methods of moments in tests of conditional asset pricing models. *The Journal of Financial Economics, 36*, 29–55

Ferson, W. E., & Harvey, C. R. (1993). The risk and predictability of international equity returns. *Review of Financial Studies, 6*(3), 527–566.

Ferson, W. E., & Schadt, R. (1996). Measuring fund strategy and performance in changing economic conditions. *Journal of Finance, 51*(2), 425–461.

Greene, W. (2008). *Econometric analysis*. Washington, DC: Pearson.

Gruber, M. J. (1996). Presidential address: Another puzzle: The growth in actively managed mutual funds. *Journal of Finance, 51*, 783–810.

Hanley, K. W., Lee, C., & Seguin, P. (1996). The marketing of closed-end fund IPOs. *Journal of Financial Intermediation, 5*, 127–159.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica, 50*, 1029–1054.

Hardouvellis, G., Wizman, T., & La Porta, R. (1993). What moves the discount on country equity funds? Working paper 4571, NBER.

Harvey, C. (1991). The world price of covariance risk. *Journal of Finance, 46*, 111–159.

Hietala, P. T. (1989). Asset pricing in partially segmented markets: Evidence from the finnish market. *Journal of Finance, 44*(3), 697–718.

Lee, C. M. C., Shleifer, A., & Thaler, R. (1991). Investor sentiment and the closed-end fund puzzle. *Journal of Finance, 46*(1), 75–109.

Mackinlay, C., & Richardson, M. (1991). Using generalized method of moments to test mean-variance efficiency. *Journal of Finance, 46*(2), 511–527.

Peavy, J. W., III. (1990). Returns on initial public offerings of closed-end funds. *Review of Financial Studies, 3*, 695–708.

Pontiff, J. (1997). Excess volatility and closed-end funds. *American Economic Review, 87*(1), 155–169.

Roll, R. (1977). A critique of the asset pricing theory's tests. *Journal of Financial Economics, 4*, 129–176.

Solnik, B. (1996). *International investments* (3rd ed.). New York: Addison-Wesley.

Stulz, R. (1995). International portfolio choice and asset pricing: An integrative survey. In R. Jarrow, V. Maksimovic, & W. T. Ziemba (Eds.), *Handbooks in operation research and management science* (Vol. 9). Amsterdam: North-Holland.

Stulz, R., & Wasserfallen, W. (1995). Foreign equity investment restrictions, capital flight, and shareholder wealth maximization: Theory and evidence. *Review of Financial Studies, 8*, 1019–1057.

Weiss, K. (1989). The post-offering price performance of closed-end funds. *Financial Management, 18*, 57–67.

A Comparison of Portfolios Using Different Risk Measurements

26

Jing Rung Yu, Yu Chuan Hsu, and Si Rou Lim

Contents

Abstract

In order to find out which risk measurement is the best indicator of efficiency in a portfolio, this study considers three different risk measurements: the mean-variance model, the mean absolute deviation model, and the downside risk model. Meanwhile short selling is also taken into account since it is an important

J.R. Yu (✉) • Y.C. Hsu • S.R. Lim
National Chi Nan University, Nantou, Taiwan
e-mail: jennifer@ncnu.edu.tw; s96213020@ncnu.edu.tw; s96213021@ncnu.edu.tw

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_26,

strategy that can bring a portfolio much closer to the efficient frontier by improving a portfolio's risk-return trade-off. Therefore, six portfolio rebalancing models, including the MV model, MAD model, and the downside risk model, with/without short selling, are compared to determine which is the most efficient. All models simultaneously consider the criteria of return and risk measurement. Meanwhile, when short selling is allowed, models also consider minimizing the proportion of short selling. Therefore, multiple objective programming is employed to transform multiple objectives into a single objective in order to obtain a compromising solution. An example is used to perform simulation, and the results indicate that the MAD model, incorporated with a short selling model, has the highest market value and lowest risk.

Keywords

Portfolio selection • Risk measurement • Short selling • MV model • MAD model • Downside risk model • Multiple objective programming • Rebalancing model • Value-at-risk • Conditional value-at-risk

26.1 Introduction

Determining how to maximize the profit and minimize the risk of a portfolio is an important issue in portfolio selection. The mean-variance (MV) model of portfolio selection is based on the assumptions that investors are risk averse and the return of assets is normally distributed (Markowitz 1952). This model is regarded as the basis of modern portfolio theory (Deng et al. 2005).

However, the MV model is limited in that it only leads to optimal decisions if the investor's utility functions are quadratic or if investment returns are jointly elliptically distributed (Grootveld and Hallerbach 1999; Papahristoulou and Dotzauer 2004). Thus, numerous researches have focused on risk, return, and diversification in the development of investment strategies. Also, a large number of researches have been proposed to improve the performance of investment portfolios (Deng et al. 2000; Yu and Lee 2011).

Konno and Yamazaki (1991) proposed a linear mean absolute deviation (MAD) portfolio optimization model. The MAD model replaces the variance of objective function in the MV model with the mean absolute deviation. The major advantage of the MAD model is that the estimation of the covariance matrix of asset returns is not needed. Also, it is much easier to solve large-scale problems with linear programming than with quadratic approaches (Simaan 1997).

Portfolio selection under shortfall constraints originated from Roy's (1952) safety-first theory. Economists have found that investors care about downside losses more than they care about upside gains. Therefore, Markowitz (1959) suggested using semi-variance as a measure of risk, instead of variance, because semi-variance measures downside losses rather than upside gains. The use of downside risk (DSR) measures is proposed due to the problems encountered in using a conventional mean-variance analysis approach in the presence of non-normality in the emerging market data. Unlike the mean-variance framework, the downside risk measure does not assume that the return distributions of assets are normal. In addition, the increasing emphasis of investors on limiting losses might make the downside risk measure more intuitively appealing (Stevenson 2001; Ang et al. 2006). The downside risk measure can help investors make proper decisions when returns are non-normally distributed, especially for emerging market data or for an international portfolio selection (Vercher et al. 2007).

Another type of shortfall constraint is value-at-risk (VaR), which is a percentile-based metric system for risk measurement purposes (Jorion 1996). It is defined as the maximum loss that a portfolio can suffer at a given level of confidence and at a given horizon (Fusai and Luciano 2001). However, VaR is too weak to handle the situation when losses are not "normally" distributed, as loss distribution tends to exhibit "fat tail" or empirical discreteness. Conditional value-at-risk (CVaR) is an alternative measure that quantifies losses that might be encountered in the tail of loss distribution (Rockafellar and Uryasev 2002; Topaloglou et al. 2002). A confidence level is required when employing the measurements of VaR and CVaR. The allocation of a portfolio varies with the varying confidence level. Therefore, neither measure is included for comparison in this chapter.

Generally, short selling has good potential to improve a portfolio's risk-return trade-off (White 1990; Kwan 1997) and is considered by most investors to obtain interest arbitrage; however, it comes with high risk (Angel et al. 2003). Since high risks should be avoided, the role of short selling is minimized in this chapter. Instead, three kinds of portfolio models with and without short selling are compared.

Section 26.2 introduces the mean-variance, mean absolute, and downside risk models. In Sect. 26.3, the rebalancing models with/without short selling are proposed. In Sect. 26.4, the performances of three different risk measurements with/without short selling are compared by using the historical data of 45 stocks listed in the TSE50 index. Finally, Sect. 26.5 presents the conclusions along with suggestions for future research.

26.2 Portfolio Selection Models

In this section, the mean-variance, the mean absolute deviation, and the downside risk models are introduced separately. First, the notations are defined as follows:
n is the number of available securities.
w_i is the investment portion in securities i for $i = 1, \ldots, n$.
r_i is the return on securities i.
μ is the expected portfolio return.
${\sigma_i}^2$ is the variance of the return on securities i.
${\sigma_{ij}}^2$ is the covariance between the returns of securities i and j.
r_{it} is the return on securities i in period t for $t = 1, \ldots, T$, which is assumed to be available through historical data.
R_i is equal to $\frac{1}{T}\sum_{t=1}^{T} r_{it}$.
d_t is the deviation between the return and the average return.

26.2.1 The Mean-Variance Model (MV)

The MV model uses the variance of the return as the measure of risk and formulates the portfolio optimization problem as the following quadratic programming problem (Markowitz 1952):

$$\begin{aligned} &\text{Min } \sigma_p = \sum_{i=1}^{n} w_i^2 \sigma_i^2 + \sum_{i=1}^{n} \sum_{j=1(i \neq j)}^{n} \sigma_{ij} w_i w_j \\ &\text{s.t. } \sum_{i=1}^{n} r_i w_i \geq \mu, \end{aligned} \tag{26.1}$$

$$\sum_{i=1}^{n} w_i = 1, \tag{26.2}$$

$$w_i \geq 0, \tag{26.3}$$

for $i = 1, \ldots, n$.

Constraint (26.1) expresses the requirements μ of a portfolio return, and constraint (26.2) is the budget constraint. The model is known to be valid if an investor is risk averse in the sense that he prefers less standard deviation of the portfolio rather than more. Since $w_i \geq 0$, a short sale is not allowed here.

26.2.2 The Mean Absolute Deviation Model (MAD)

The mean-variance model is weak in constructing a large-scale portfolio due to the computational difficulty associated with solving a large-scale quadratic programming problem with a dense covariance matrix. The MAD model (Konno and Yamazaki 1991) replaces the variance in the objective function of the MV model with the mean absolute deviation as follows:

$$\text{Min } \frac{1}{T} \sum_{t=1}^{T} \left| \sum_{i=1}^{n} (r_{it} - R_i) w_i \right|$$

s.t. Constraints (26.1) $\sim$ (26.3).

Because of the absolute deviation, the MAD model can be linearized as following (Chang 2005):

$$\text{Min } \frac{1}{T} \sum_{t=1}^{T} d_t$$

s.t. Constraints (26.1) $\sim$ (26.3),

$$d_t + \sum_{i=1}^{n} (r_{it} - R_i)w_i \geq 0,\ t = 1, \ldots, \mathrm{T}, \tag{26.4}$$

$$d_t - \sum_{i=1}^{n} (r_{it} - R_i)w_i \geq 0,\ \ t = 1, \ldots, \mathrm{T}. \tag{26.5}$$

If the return is lower than the average return, constraint (26.4) is a binding constraint which means $d_t = -\sum_{i=1}^{n} (r_{it} - R_i)w_i$ for $t = 1, \ldots, T$. Otherwise, constraint (26.5) is a binding constraint which means $d_t = \sum_{i=1}^{n} (r_{it} - R_i)w_i$ for $t = 1, \ldots, T$. For more details on the reformulation, please refer to Appendix 1.

Apparently, the MAD model does not require the covariance matrix of asset returns, and consequently its estimation is not needed. Large-scale problems can be solved faster and more efficiently because the MAD model has a linear rather than quadratic nature.

26.2.3 The Downside Risk Model

Vercher et al. (2007) consider the equivalent formulation of the portfolio selection problem (Speranza 1993) and reformulate the following linear optimization model by considering downside risk measurement:

$$\mathrm{Min}\ \frac{1}{T}\sum_{t=1}^{T} d_t^-$$

s.t. Constraints (26.1) $\sim$ (26.3),

$$d_t^- + \sum_{i=1}^{n} (r_{it} - R_i)w_i \geq 0,\ t = 1, \ldots, T \tag{26.6}$$

where $d_t = d_t^+ + d_t^-,\ d_t^+,\ d_t^- \geq 0$.

Downside risk measurement focuses on returns falling below some critical level (Grootveld and Hallerbach 1999). Differing from the MAD model, the downside risk model ignores constraint (26.5). If the return is lower than the average return, constraint (26.6) is a binding constraint. Please refer to Appendix 2 for more details.

26.3 The Proposed Model

Multiple period portfolio selection models with rebalancing mechanisms have become attractive in the financial field in order to get desired returns in situations that are subject to future changes (Yu et al. 2010). To reflect a changing situation in the models, the rebalancing mechanism is adopted for multiple periods (Yu and Lee 2011).

Six rebalancing models are introduced. These are MV, MAD, DSR, MV_S, MAD_S, and DSR_S. The first three models lack short selling, and the other three have short selling. The model notations are denoted as follows: $w_{i,0}^{+}$ is the weight of security i held in the previous period, $w_{i,0}^{-}$ is the weight of securities i sold short in the previous period, w_i^{+} is the total weight of securities i bought after rebalancing, and w_i^{-} is the total weight of securities i sold short after rebalancing.

With each rebalancing, l_i^{+} is the weight of securities i bought in this period, l_i^{-} is the weight of securities i sold in this period, s_i^{+} is the weight of securities i sold short in this period, s_i^{-} is the weight of securities i repurchased in this period, u_i is the binary variable that indicates whether the securities i are selected for buying, v_i is the binary variable that indicates whether the securities i are selected for selling short, and k is the initial margin requirement for short selling.

26.3.1 The MV Model

The conventional MV model can be regarded as a bi-objective model without short selling, whose objective functions are the maximization of portfolio return and minimization of portfolio risk, as measured by the portfolio variance:

$$\text{Max} \sum_{i=1}^{n} R_i w_i$$

$$\text{Min}\, \sigma_p$$

s.t. Constraint (26.2),

$$w_i = w_{i,0}^{+} + \ell_i^{+} - \ell_i^{-}, \tag{26.7}$$

$$0.05u_i \le w_i \le 0.2u_i, \tag{26.8}$$

for $i = 1, 2, \ldots, n$.

Constraint (26.7) is the rebalancing constraint; it shows the current weight for the ith security according to the previous period. Constraint (26.8) is the required range of weights for each security in buying. For simplexity, the upper and lower bounds of each weight are set 0.2 and 0.05, respectively.

26.3.2 The MAD Model

The objectives of the MAD model include the maximization of return and minimization of the mean absolute deviation, which is transformed into a linear deviation as follows:

$$\text{Max} \sum_{i=1}^{n} R_i w_i$$

$$\text{Min}\ \frac{1}{T}\sum_{t=1}^{T} d_t$$

s.t. Constraints (26.2) ~ (26.5), (26.7), (26.8).

26.3.3 The DSR Model

The DSR model considers the objectives of maximizing the return and minimizing the downside risk, which is transformed into a linear risk as follows:

$$\text{Max}\sum_{i=1}^{n} R_i w_i$$

$$\text{Min}\frac{1}{T}\sum_{t=1}^{T} d_t^-$$

s.t. Constraints (26.2), (26.3), (26.6), (26.7), (26.8).

When short selling is allowed, the above three models are reformulated as follows (Yu and Lee 2011):

26.3.4 The MV_S Model

$$\text{Max}\sum_{i=1}^{n} R_i\left(w_i^+ - w_i^-\right)$$

$$\text{Min}\,\sigma_p$$

$$\text{Min}\sum_{i=1}^{n} w_i^-$$

$$\text{s.t.}\ \sum_{i=1}^{n}\left(w_i^+ + k w_i^-\right) = 1, \tag{26.9}$$

$$w_i^+ = w_{i,0}^+ + \ell_i^+ - \ell_i^-, \tag{26.10}$$

$$w_i^- = w_{i,0}^- + s_i^+ - s_i^-, \tag{26.11}$$

$$0.05u_i \le w_i^+ \le 0.2u_i, \tag{26.12}$$

$$0.05v_i \leq w_i^- \leq 0.2v_i, \tag{26.13}$$

$$u_i + v_i = y_i, \tag{26.14}$$

for $i = 1, 2, \ldots, n$.

Unlike the MV model, the MV_S model has an extra objective, namely, minimizing the short selling proportion. The total budget is including the cost of buying and short selling as shown in constraint (26.9). k is the initial margin requirement for short selling. Constraint (26.10) indicates the current weight for the ith securities based on the previous period. Constraint (26.11) indicates the current short selling proportion of the ith security adjusted by the previous period. Constraints (26.12) and (26.13) limit the upper and lower bounds, respectively, of the long and short selling proportion for the ith security.

26.3.5 The MAD_S Model

Based on the MAD model (Konno and Yamazaki 1991), the MAD_S model replaces the objective function of minimizing the variance in the MV_S model (Yu and Lee 2011) with the objective of minimizing the absolute deviation of average return, as follows:

$$\text{Min} \sum_{i=1}^{n} \left(\left| \left(w_i^+ - w_i^-\right) r_{it} - \left(w_i^+ - w_i^-\right) R_i \right| \right)$$

The objective of the mean absolute deviation can be transformed into a linear problem:

$$\text{Min} \frac{1}{T} \sum_{t=1}^{T} d_t$$

$$\text{s.t. } d_t + \sum_{i=1}^{n} \left(w_i^+ - w_i^-\right) r_{it} - \left(w_i^+ - w_i^-\right) R_i \geq 0, \tag{26.15}$$

$$d_t - \sum_{i=1}^{n} \left(w_i^+ - w_i^-\right) r_{it} - \left(w_i^+ - w_i^-\right) R_i \geq 0, \tag{26.16}$$

for $t = 1, \ldots, T$.

The following is the MAD_S model:

$$\text{Max} \sum_{i=1}^{n} R_i \left(w_i^+ - w_i^-\right)$$

$$\text{Min}\ \frac{1}{T}\sum_{t=1}^{T} d_t$$

$$\text{Min}\sum_{i=1}^{n} w_i^-$$

s.t. Constraints (26.9) $\sim$ (26.16).

26.3.6 The DSR_S Model

The DSR_S model focuses on the deviation when the return falls below the average return, as follows:

$$\text{Max}\sum_{i=1}^{n} R_i\left(w_i^+ - w_i^-\right)$$

$$\text{Min}\ \frac{1}{T}\sum_{t=1}^{T} d_t^-$$

$$\text{Min}\sum_{i=1}^{n} w_i^-$$

s.t. Constraints (26.9) $\sim$ (26.14),

$$d_t^- + \sum_{i=1}^{n}\left(w_i^+ - w_i^-\right)r_{it} - \left(w_i^+ - w_i^-\right)R_i \geq 0, \tag{26.17}$$

for $t = 1, \ldots, T$.

Apparently, all six models have multiple objectives. Therefore, multiple objective programming (Zimmermann 1978, and Lee and Li 1993) is adopted to transform the multiple objectives into a single objective. For more details, please refer to Appendix 3. Taking the MAD_S model as an example, we can reformulate the multiple objectives as follows:

$$\begin{aligned}&\text{Max}\ \lambda\\ &\text{s.t.}\ \ \lambda \leq \frac{(r^* - r_l)}{\left(r_g - r_l\right)},\end{aligned} \tag{26.18}$$

$$\lambda \leq \frac{(\sigma^* - \sigma_l)}{\left(\sigma_g - \sigma_l\right)}, \tag{26.19}$$

$$\lambda \leq \frac{\left(w^{*} - w_{l}^{-}\right)}{\left(w_{g}^{-} - w_{l}^{-}\right)}, \tag{26.20}$$

$$r^{*} = \sum_{i=1}^{n} R_{i}\left(w_{i}^{+} - w_{i}^{-}\right), \tag{26.21}$$

$$\sigma^{*} = \frac{1}{T}\sum_{t=1}^{T} d_{t}^{-}, \tag{26.22}$$

$$w^{-*} = \sum_{i=1}^{n} w_{i}^{-}, \tag{26.23}$$

Constraints (26.9) $\sim$ (26.16).

For multiple objective programming (Lee and Li 1993), r^* is the return of the portfolio, r_l is the anti-ideal return of the portfolio, r_g is the ideal return of the portfolio that maximizes the objective, σ^* is the inherent risk of the portfolio, σ_l is the anti-ideal risk of the portfolio, σ_g is the ideal risk of the portfolio, w^{-*} is the short selling proportion of the portfolio, w_l is the anti-ideal short selling proportion of the portfolio, and w_g^- is the ideal short selling proportion of the portfolio. The constraints (26.18–26.20) are the achievements for maximizing the return, minimizing the absolute deviation and objectives minimizing the short selling problem of the corresponding portfolio, which are less than or equal to the whole achievement (λ). The whole achievement (λ) should be maximized.

In the same way, the other five multiple objective models can be reformulated in turn as a single-objective model. The details of the transformation are introduced in Appendix 3.

26.4 Experimental Results

Forty-five stocks listed on the Taiwan Stock Exchange were adopted and used to compare the six models discussed above in order to determine which one is the best. The benchmark is the Taiwan 50 Index (TSE50).

The exchange codes of the 45 stocks are listed in Table 26.1. The duration of the analyzed data is from November 1, 2006, to November 24, 2009. The historical data of the first 60 transaction days are used to build the initial models. For the monthly updates, 20 transaction days are set as a sliding window. This study assumes a budget of \$1 million NTD is invested in the Taiwan Stock Market. For investments based on the weights generated by the initial models, the first transaction day is January 25, 2007, and there are 34 rebalancing times in total. The models are executed on an Intel Pentium Dual CPU E2200 2.20GHz and 2G RAM computer, with Lingo11.0, an optimizing software.

From Table 26.2, it is apparent that the MAD and DSR models are more efficient than the MV model because they both use linear transformation for problem solving. This is much faster and more efficient when handling a large-scale

Table 26.1 The exchange codes of 45 stocks

No	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Code	1101	1102	1216	1301	1303	1326	1402	1722	2002	2105	2308	2311	2317	2324	2325
No	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
Code	2330	2347	2353	2354	2357	2382	2409	2454	2498	2603	2801	2880	2881	2882	2883
No	31	32	33	34	35	36	37	38	39	40	41	42	43	44	45
Code	2885	2886	2888	2890	2891	2892	2912	3009	3231	3474	3481	5854	6505	8046	9904

Table 26.2 The running time of six portfolios

Models without short selling	**MV**	**MAD**	**DSR**
	00:09:06	00:01:01	00:01:01
Models with short selling	**MV_S**	**MAD_S**	**DSR_S**
	00:10:01	00:01:15	00:01:14

(hh:mm:ss)

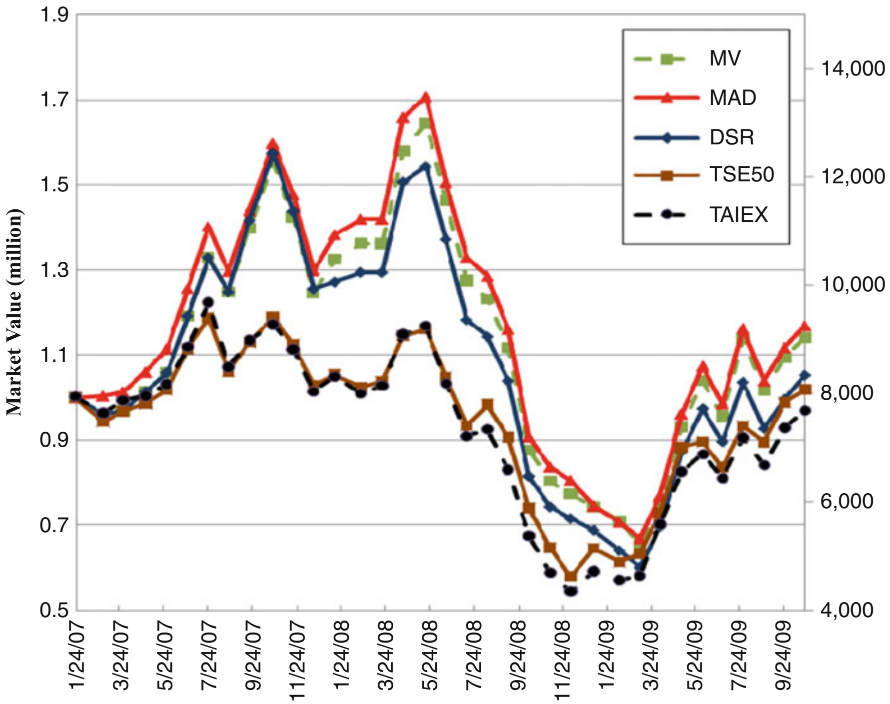


Fig. 26.1 The market value of three risk measurements without short selling, the TSE50, and TAIEX

problem. As one can see in Table 26.2, the MV model takes 9 min and 6 sec to compute the result. However, the MAD and DSR models only take 1 min or slightly more to solve the same data. Moreover, it is not necessary to calculate the covariance matrix to set up the MAD and DSR models; this makes it very easy to update the models when new data are added (Konno and Yamazaki 1991).

Figures 26.1, 26.2 and 26.3 show the comparisons of the MV, MAD, and DSR models, respectively. These are the models without short selling. Figures 26.4, 26.5 and 26.6 show the comparisons of the MV_S, MAD_S, and DSR_S models, respectively. These are the models with short selling. The TSE50 is used as the benchmark to these models. As shown in Fig. 26.1, the market value of the MAD

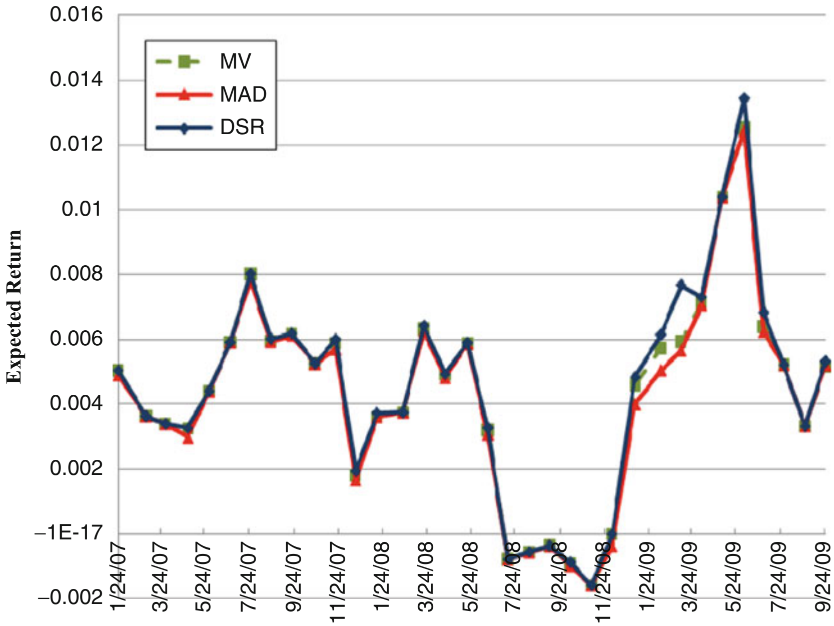


Fig. 26.2 The expected return of three risk measurements without short selling

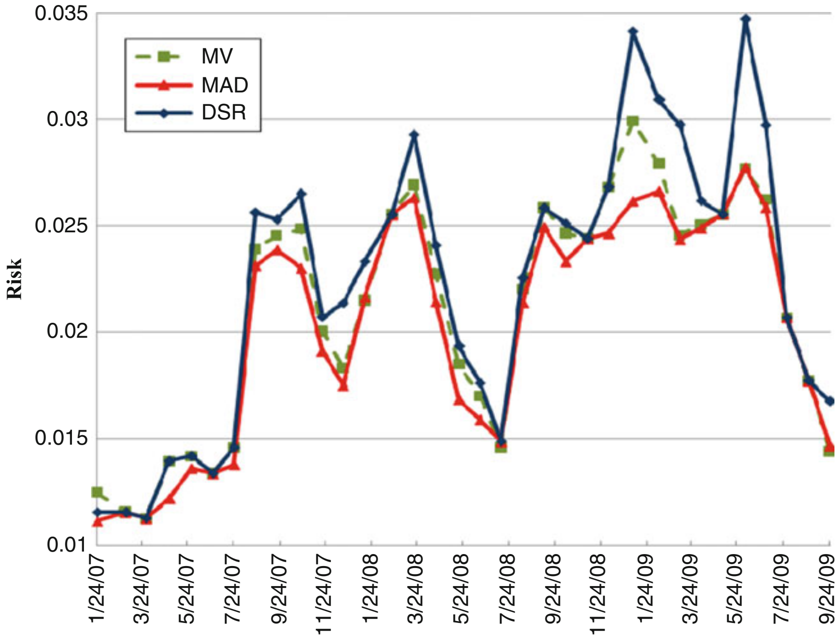


Fig. 26.3 The risk of three risk measurements without short selling

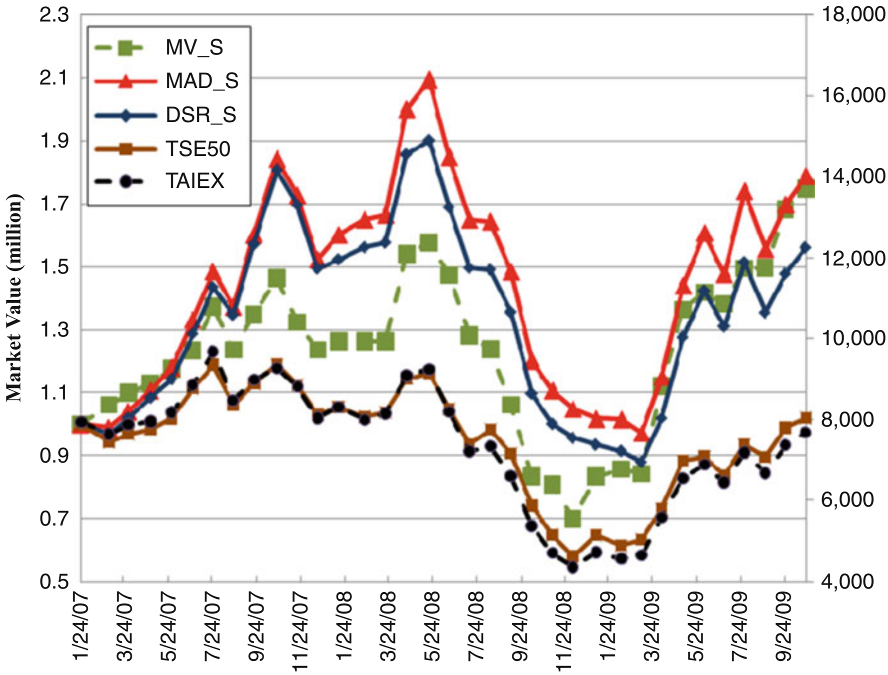


Fig. 26.4 The market value of three risk measurements with short selling, the TSE50, and TAIEX

model is always greater than that of the other models and the benchmark. Figures 26.2 and 26.3 display the expected return and risk of the portfolios constructed with three risk measurements without short selling. Figure 26.2 shows that the expected returns of these three portfolios are almost the same. However, Fig. 26.3 shows that the MAD model has the lowest risk under a similar expected return to that of the others. Especially, in January 24, 2009, the risk in the MAD model was much lower than in the other two models.

Downside risk measurement is applied when investors only take the negative return into consideration and focus on the loss of investment. In other words, their concern is with the real loss, not with the positive deviation of the average return in portfolios. Therefore, in Fig. 26.3, the risk generated by the DSR model is always higher than other measurements.

In Fig. 26.4, the market value of the MV_S, MAD_S, and DSR_S models is compared. Since these models take short selling into consideration, the portfolio selection is more flexible, and the risk is much lower, as Fig. 26.6 shows. Under the same expected return, the MAD_S model has the lowest risk among the three models, using the mean absolute deviation risk measure. Figure 26.4 shows that it also has the highest market value. Even though the market value of each model is increased after short selling is allowed, the market value of the MAD_S model is always higher than the other models and the benchmark. Evidently, the MAD model is suggested for use as the best risk measurement tool with or without short sell.

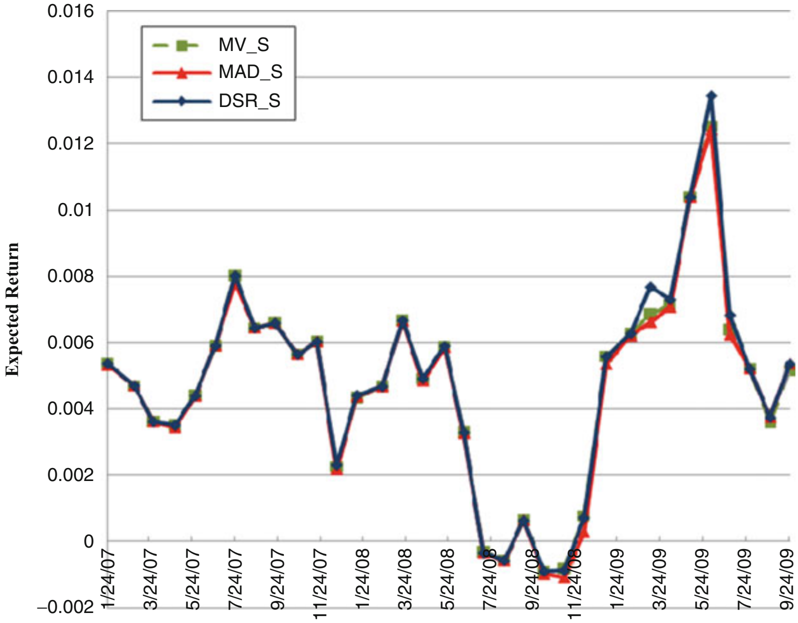


Fig. 26.5 The expected return of three risk measurements with short selling

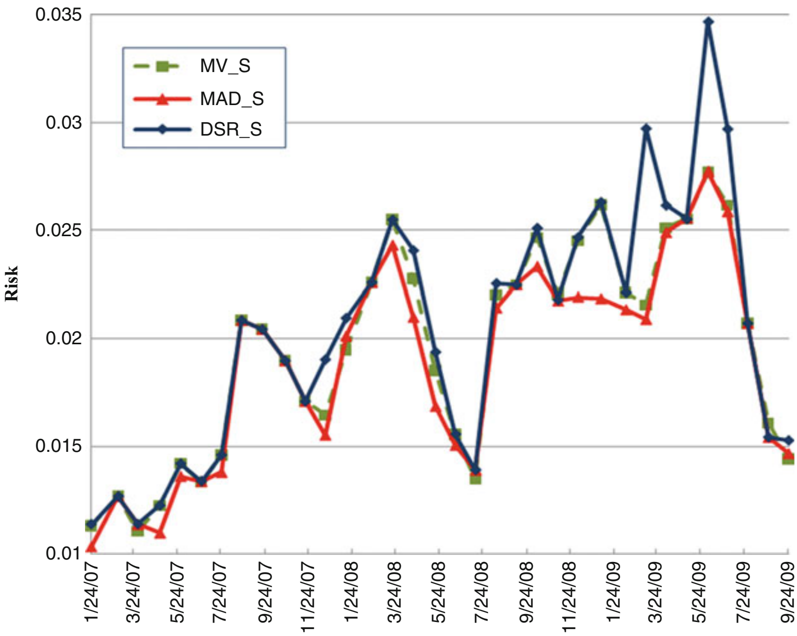


Fig. 26.6 The risk of three risk measurements with short selling

26.5 Conclusion

This chapter compares six different rebalancing models, with/without short selling, in order to determine which is more flexible for portfolio selection. One of the advantages of the MAD and DSR models is that they can be linearized; thus, they are faster and more efficient than the MV model, especially with large-scale problems.

The experimental results indicate that the MAD and MAD_S models are efficient in handling data and show higher market value than the other models; moreover, they have lower risks in situations with/without short selling.

However, there remain important risk measurements, such as VaR and CVaR, for future research to investigate. Thus, future studies may focus on developing rebalancing models with the measures of VaR and CVaR. Since the rebalancing period is fixed, the dynamic rebalancing mechanism is required for the first change environment. In addition, a portfolio selection model able to predict future returns is required.

Appendix 1

The Linearization of the MAD Model

Konno and Yamazaki (1991) assume that r_{it} is the realization of random variable r_i during period t ($t = 1, \ldots, T$), and $R_i = \frac{1}{T}\sum_{i=1}^{n} r_{it}$.

The MAD model is as follows:

$$\begin{aligned} &\text{Max } \frac{1}{T}\sum_{t=1}^{T}\left|\sum_{i=1}^{n}(r_{it}-R_i)w_i\right| \\ &\text{s.t.} \sum_{i=1}^{n} R_i x_i \geq \mu, \\ &\sum_{i=1}^{n} w_i = 1, \\ &w_i \geq 0, \end{aligned}$$

for $i = 1, \ldots, n$.

$$\text{Let}\left|\sum_{i=1}^{n}(r_{it}-R_i)w_i\right| = d_t = {d_t}^{+} + {d_t}^{-},$$

$${d_t}^{+}\text{and } {d_t}^{-} \geq 0,$$

$$\text{then}\sum_{i=1}^{n}(r_{it}-R_i)w_i = {d_t}^{+} - {d_t}^{-},$$

$$d_t - d_t^- = \sum_{i=1}^{n} (r_{it} - R_i) w_i + d_t^-,$$

$$d_t^- = \frac{1}{2}\left(d_t - \sum_{i=1}^{n} (r_{it} - R_i) w_i\right) \geq 0.$$

$$2d_t^- = d_t - \sum_{i=1}^{n} (r_{it} - R_i) w_i.$$

Similarity, perform the same process,

$$d_t^+ = \sum_{i=1}^{n} (r_{it} - R_i) w_i + d_t^-,$$

$$d_t - d_t^+ = -\sum_{i=1}^{n} (r_{it} - R_i) w_i + d_t^+,$$

$$2d_t^+ = -\sum_{i=1}^{n} (r_{it} - R_i) w_i + d_t,$$

$$d_t^+ = \frac{1}{2}\left(\sum_{i=1}^{n} (r_{it} - R_i) w_i + d_t\right) \geq 0,$$

$$d_t^+, d_t^- \geq 0.$$

Two constraints are added because d_t^+, $d_t^- \geq 0$.
Then the model can be transformed into a linear model as follows:

$$\text{Min}\ \frac{1}{T}\sum_{t=1}^{T}\left|\sum_{i=1}^{n} (r_{it} - R_i) w_i\right| = \frac{1}{T}\sum_{t=1}^{T} d_t^+ + d_t^-$$

$$d_t^+ = \frac{1}{2}\left(\sum_{i=1}^{n} (r_{it} - R_i) w_i + d_t\right) \geq 0,$$

$$\Rightarrow \sum_{i=1}^{n} (r_{it} - R_i) w_i + d_t \geq 0.$$

$$d_t^{-} = \frac{1}{2}\left(-\sum_{i=1}^{n}(r_{it} - R_i)w_i + d_t\right) \geq 0,$$

$$\Rightarrow -\sum_{i=1}^{n}(r_{it} - R_i)w_i + d_t \geq 0.$$

where $d_t = d_t^{+} + d_t^{-}$.

Therefore, the MAD model can be linearized as the following linear model:

$$\begin{aligned}
&\text{Min } \frac{1}{T}\sum_{t=1}^{T} d_t \\
&\text{s.t. } d_t + \sum_{i=1}^{n}(r_{it} - R_i)w_i \geq 0, t = 1, \ldots, T, \\
&d_t - \sum_{i=1}^{n}(r_{it} - R_i)w_i \geq 0, t = 1, \ldots, T, \\
&\sum_{i=1}^{n} R_i x_i \geq \mu, \\
&\sum_{i=1}^{n} w_i = 1, \\
&w_i \geq 0,
\end{aligned}$$

for $i = 1, \ldots, n$.

Appendix 2

The Linearization of the DSR Model

The use of variance as a measure of risk makes no distinction between gains and losses. The following mean semi-absolute deviation risk measurement proposed by Speranza (1993) is used to find the portfolios with minimum semi-variance:

$$\text{Min } \frac{\frac{1}{T}\sum_{t=1}^{T}\left|\sum_{i=1}^{n}(r_{it} - R_i)w_i\right| - \sum_{i=1}^{n}(r_{it} - R_i)w_i}{2}$$

$$\text{s.t.} \sum_{i=1}^{n} R_i x_i \geq \mu,$$

$$\sum_{i=1}^{n} w_i = 1,$$

$$w_i \geq 0,$$

for $i = 1, \ldots n$.

Because of the absolute deviation, $\left|\sum_{i=1}^{n} (r_{it} - R_i) w_i\right|$ of the DSR model can be linearized in the same manner as the MAD model:

$$\sum_{i=1}^{n} (r_{it} - R_i) w_i = d_t{}^{+} - d_t{}^{-}$$

$$\frac{\left|\sum_{i=1}^{n} (r_{it} - R_i) w_i\right| - \sum_{i=1}^{n} (r_{it} - R_i) w_i}{2} = \frac{d_t{}^{+} + d_t{}^{-} - d_t{}^{+} + d_t{}^{-}}{2} = d_t{}^{-}$$

$$d_t{}^{+} = \frac{1}{2}\left(\sum_{i=1}^{n} (r_{it} - R_i) w_i + d_t\right) \geq 0,$$

$$d_t{}^{-} \geq 0.$$

Then the DSR model is reformulated as follows:

$$\text{Min} \frac{1}{T} \sum_{t=1}^{T} d_t^{-}$$

$$\text{s.t.} \sum_{i=1}^{n} R_i x_i \geq \mu,$$

$$\sum_{i=1}^{n} w_i = 1,$$

$$w_i \geq 0,$$

$$d_t^{-} + \sum_{i=1}^{n} (r_{it} - R_i) w_i \geq 0, t = 1, \ldots, T,$$

$$d_t{}^{-} \geq 0, t = 1, \ldots, T,$$

for $i = 1, \ldots, n$.

Appendix 3

Multiple Objective Programming

The aforementioned multiple objective models in Sect. 26.3 are solved by fuzzy multiple objective programming (Zimmermann 1978; Lee and Li 1993) in order to transform the multiple objective model into a single-objective model. Fuzzy multiple objective programming based on the concept of fuzzy set uses a min operator to calculate the membership function value of the aspiration level, λ, for all of the objectives.

The following is a multiple objective programming problem (Lee and Li 1993):

$$\mathrm{Max}\, Z = [Z_1, Z_2, \ldots, Z_l]^T = [c_1 x, c_2 x, \ldots, c_l x]^T$$

$$\mathrm{Min}\, W = [W_1, W_2, \ldots, W_l]^T = [q_1 x, q_2 x, \ldots, q_l x]^T$$

$$\text{s.t. } Ax^* b,$$

$$x \geq 0,$$

where C_k, $k = 1, 2, \ldots, l$, c_s, $s = 1, 2, \ldots, r$, and x are n-dimensional vectors; b is an m-dimensional vector; A is an $m \times n$ matrix; and * denotes the operators $\leq$, $=$, or $\geq$ The program aimed to achieve its maximization of the achievement level for each objective while also considering a trade-off among the conflicting objectives or criteria. The ideal and anti-ideal solutions must be obtained in advance. This ideal solution and anti- ideal solutions are given by the decision maker, respectively, as follows:

$$I^+ = \left(Z_1^*, Z_2^*, \ldots, Z_l^*; W_1^*, W_2^*, \ldots, W_l^*\right),$$

$$I^- = \left(Z_1^-, Z_2^-, \ldots, Z_r^-; W_1^-, W_2^-, \ldots, W_r^-\right).$$

The membership (achievement) functions for the objectives are defined as follows:

$$\mu_k(Z_K) = \frac{Z_K(x) - Z_k^-}{Z_k^* - Z_k^-}, k = 1, 2, \ldots, l,$$

$$\mu_s(Z_s) = \frac{W_s^- - W_s(x)}{W_s^- - W_s^*}, s = 1, 2, \ldots, r.$$

Then the "min" operator is used; the multiple objective programming is formulated as follows:

$$
\begin{aligned}
&\text{Max } \lambda \\
&\text{s.t. } \lambda \leq \left(Z_K(x) - Z_k^-\right)/\left(Z_k^* - Z_k^-\right), k = 1, 2, \ldots, l, \\
&\qquad \lambda \leq \left(W_s^- - W_s(x)\right)/\left(W_s^- - W_s^*\right), s = 1, 2, \ldots, r, \\
&\qquad x \in X,
\end{aligned}
$$

where λ is defined as $\lambda = \min_i \mu(x) = \min_{k,s} \left(\mu_k(Z), \mu_s(W_s)\right)$.

References

Ang, A., Chen, J., & Xing, Y. (2006). Downside risk. *The Review of Financial Studies, 19*, 1191–1239.

Angel, J. J., Christophe, S. E., & Ferri, M. G. (2003). A close look at short selling on Nasdaq. *Financial Analysts Journal, 59*, 66–74.

Chang, C. T. (2005). A modified goal programming approach for the mean-absolute deviation portfolio optimization model. *Applied Mathematics and Computation, 171*, 567–572.

Deng, X. T., Wang, S. Y., & Xia, Y. S. (2000). Criteria models and strategies in portfolio selection. *Advanced Modeling and Optimization, 2*, 79–104.

Deng, X. T., Li, Z. F., & Wang, S. Y. (2005). A minimax portfolio selection strategy with equilibrium. *European Journal of Operational Research, 166*, 278–292.

Fusai, G., & Luciano, E. (2001). Dynamic value at risk under optimal and suboptimal portfolio policies. *European Journal of Operational Research, 135*, 249–269.

Grootveld, H., & Hallerbach, W. (1999). Variance vs downside risk: Is there really that much difference? *European Journal of Operational Research, 114*, 304–319.

Jorion, P. H. (1996). *Value at risk: A new benchmark for measuring derivative risk*. New York: Irwin Professional Publishers.

Konno, H., & Yamazaki, H. (1991). Mean-absolute deviation portfolio optimization model and its applications to Tokyo stock market. *Management Science, 37*, 519–531.

Kwan, C. C. Y. (1997). Portfolio selection under institutional procedures for short selling: Normative and market equilibrium considerations. *Journal of Banking & Finance, 21*, 369–391.

Lee, E. S., & Li, R. J. (1993). Fuzzy multiple objective programming and compromise programming with Pareto optimum. *Fuzzy Sets and Systems, 53*, 275–288.

Markowitz, H. M. (1952). Portfolio selection. *The Journal of Finance, 7*, 77–91.

Markowitz, H. M. (1959). *Portfolio selection, efficient diversification of investments*. New York: Wiley.

Papahristoulou, C., & Dotzauer, E. (2004). Optimal portfolios using linear programming models. *Journal of the Operational Research Society, 55*, 1169–1177.

Rockafellar, R. T., & Uryasev, S. (2002). Conditional value-at-risk for general loss distribution. *Journal of Banking and Finance, 26*, 1443–1471.

Roy, A. D. (1952). Safety first and the holding of assets. *Econometrica, 20*, 431–449.

Simaan, Y. (1997). Estimation risk in portfolio selection: The mean variance model versus the mean absolute deviation model. *Management Science, 43*, 1437–1446.

Speranza, M. G. (1993). Linear programming models for portfolio optimization. *Finance, 14*, 107–123.

Stevenson, S. (2001). Emerging markets, downside risk and the asset allocation decision. *Emerging Markets Review, 2*, 50–66.

Topaloglou, N., Vladimirou, H., & Zenios, S. A. (2002). CVaR models with selective hedging for international asset allocation. *Journal of Banking and Finance, 26*, 1535–1561.

Vercher, E., Bermudez, J. D., & Segura, J. V. (2007). Fuzzy portfolio optimization under downside risk measures. *Fuzzy Sets and Systems, 158*, 769–782.

White, J. A. (1990). More institutional investors selling short: But tactic is part of a wider strategy. *Wall Street Journal, 7*, 493–501.

Yu, J. R., & Lee, W. Y. (2011). Portfolio rebalancing model using multiple criteria. *European Journal of Operational Research, 209*, 166–175.

Yu, M., Takahashi, S., Inoue, H., & Wang, S. (2010). Dynamic portfolio optimization with risk control for absolute deviation model. *Journal of Operational Research, 201*, 349–364.

Zimmermann, H. J. (1978). Fuzzy programming and linear programming with several objective functions. *Fuzzy Sets and Systems, 1*, 44–45.

Using Alternative Models and a Combining Technique in Credit Rating Forecasting: An Empirical Study

27

Cheng-Few Lee, Kehluh Wang, Yating Yang, and Chan-Chien Lien

Contents

Abstract

Credit rating forecasting has long time been very important for bond classification and loan analysis. In particular, under the Basel II environment, regulators

C.-F. Lee
Department of Finance and Economics, Rutgers Business School, Rutgers, The State University of New Jersey, Piscataway, NJ, USA

Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
e-mail: cflee@business.rutgers.edu

K. Wang • Y. Yang (✉)
Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
e-mail: lkwang@mail.nctu.edu.tw; yatingyang.iof98g@nctu.edu.tw

C.-C. Lien
Treasury Division, E.SUN Commercial Bank, Taipei, Taiwan

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_27,

in Taiwan have requested the banks to estimate the default probability of the loan based on its credit classification. A proper forecasting procedure for credit rating of the loan is crucially important in abiding the rule.

Credit rating is an ordinal scale from which the credit category of a firm can be ranked from high to low, but the scale of the difference between them is unknown. To model the ordinal outcomes, this study first constitutes an attempt utilizing the ordered logit and the ordered probit models, respectively. Then, we use ordered logit combining method to weigh different techniques' probability measures as described in Kamstra and Kennedy (*International Journal of Forecasting 14*, 83–93, 1998) to form the combining model.

The samples consist of firms in the TSE and the OTC market and are divided into three industries for analysis. We consider financial variables, *market* variables, as well as macroeconomic variables and estimate their parameters for out-of-sample tests. By means of cumulative accuracy profile, the receiver operating characteristics, and McFadden R^2, we measure the goodness-of-fit and the accuracy of each prediction model. The performance evaluations are conducted to compare the forecasting results, and we find that combining technique does improve the predictive power.

Keywords

Bankruptcy prediction • Combining forecast • Credit rating • Credit risk • Credit risk index • Forecasting models • Logit regression • Ordered logit • Ordered probit • Probability density function

27.1 Introduction

This study explores the credit rating forecasting techniques for firms in Taiwan. We employ the ordered logit and the ordered probit models for rating classification and then a combining procedure to integrate both. We then examine empirically the performance of these alternative methods, in particular, whether the combining forecasting performs better than any individual method.

Credit rating forecasting has long time been very important for bond classification and loan analysis. In particular, under the Basel II environment, regulators in Taiwan have requested the banks to estimate the default probability of the loan based on its credit classification. A proper forecasting procedure for credit rating of the loan is crucially important in abiding the rule.

Different forecasting models and estimation procedures have various underlying assumptions and computational complexities. They have been used extensively by researchers in the literature. Review papers like Hand and Henley (1997), Altman and Sounders (1997), and Crouhy et al. (2000) have traced the developments of the credit classification and bankruptcy prediction models over the last two decades.

Since Beaver's (1966) pioneered work, there have been considerable researches on the subject of the credit risk. Many of them (Altman 1968; Pinches and Mingo 1973; Altman and Katz 1976; Altman et al. 1977; Pompe and Bilderbeek 2005) use the multivariate discriminant analysis (MDA) which assumes normality for the explanatory variables of the default class. Zmijewski (1984) utilizes the probit model, and Ohlson (1980) applies the logit model in which discrete or continuous data can be fitted.

Kaplan and Urwitz (1979), Ederington (1985), Lawrence and Arshadi (1995), and Blume et al. (1998) show that it is a consistent structure considering credit rating as ordinal scale instead of interval scale. That is, the different values of the dependent variables as different classes represent an ordinal, but not necessarily a linear scale. For instance, higher ratings are less risky than lower ratings, but we don't have a quantitative measure indicating how much less risky they are.

Kaplan and Urwitz (1979) conduct an extensive examination of alternative prediction models including N-chotomous probit analysis which can explain the ordinal nature of bond ratings. To test the prediction accuracy of various statistical models, Ederington (1985) compares the linear regression, discriminant analysis, ordered probit, and unordered logit under the same condition. He concludes that the ordered probit can have the best prediction ability and the linear regression is the worst.

In a survey paper on forecasting methods, Mahmoud (1984) concludes that combining forecasts can improve accuracy. Granger (1989) summarizes the usefulness of combining forecasts. Clemen (1989) observes that combining forecasts increase accuracy, whether the forecasts are subjective, statistical, econometric, or by extrapolation. Kamstra and Kennedy (1998) integrate two approaches with logit-based forecast-combining method which is applicable to dichotomous, polychotomous, or ordered-polychotomous contexts.

In this paper, we apply the ordered logit and the ordered probit models in credit rating classification for listed firms in Taiwan and then combine two rating models with a logit regression technique. The performance of each model is then evaluated and we find that combining technique does improve the predictive power.

27.2 Methodology

27.2.1 Ordered Probit Model

Credit rating is an ordinal scale from which the credit category of a firm can be ranked from high to low, but the scale of the difference between them is unknown. To model the ordinal outcomes, let the underlying response function be

$$Y^* = X\beta + \varepsilon \tag{27.1}$$

where Y^* is the latent variable, X is a set of explanatory variables, and ε is the residual. Y^* is not observed, but from which we can classify the category j:

$$Y_i = j \qquad \text{if } \tau_{j-1} < Y_i^* \leq \tau_j \; (i = 1, 2, \ldots, n; \quad j = 1, 2, \ldots, J). \tag{27.2}$$

Maximum likelihood estimation can be used to estimate the parameters given a specific form of the residual distribution.

For the ordered probit model, ε is normally distributed with mean 0 and variance 1. The probability density function is

$$\phi(\varepsilon) = \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{\varepsilon^2}{2}\right) \tag{27.3}$$

and the cumulative density function is

$$\Phi(\varepsilon) = \int_{-\infty}^{\varepsilon} \frac{1}{\sqrt{2\pi}} \exp\left(-\frac{t^2}{2}\right) dt. \tag{27.4}$$

27.2.2 Ordered Logit Model

For the ordered logit model, ε has a logistic distribution with mean 0 and variance $\pi^2/3$. The probability density function is

$$\lambda(\varepsilon) = \frac{\exp(\varepsilon)}{[1 + \exp(\varepsilon)]^2} \tag{27.5}$$

and the cumulative density function is

$$\Lambda(\varepsilon) = \frac{\exp(\varepsilon)}{1 + \exp(\varepsilon)}. \tag{27.6}$$

27.2.3 Combining Method

To combine the ordered logit and the ordered probit models for credit forecasting, the logit regression method as described in Kamstra and Kennedy (1998) is applied. We first assume that firm's credit classification is determined by an index θ. Suppose there are J rating classes, ordered from 1 to J. If θ exceeds the threshold value τ_j, $j = 1, \ldots, j - 1$, credit classification changes from j rating to $j + 1$ rating. The probability of company i being in rating j is given by the integral of a standard logit from $\tau_{j-1} - \theta_i$ to $\tau_j - \theta_i$.

Each forecasting method is considered as producing $J - 1$ measures $\omega_{ji} = \tau_j - \theta_i$, $j = 1, \ldots, j - 1$ for each firm. These measures can be estimated as

$$\omega_{ji} = \ln\left[\frac{P_{1i} + \cdots + P_{ji}}{1 - P_{1i} - \cdots - P_{ji}}\right] \tag{27.7}$$

Table 27.1 Sample numbers across the industry

Industry	Non-bankruptcy observations	Bankruptcy observations	Total
Panel A: *In-sample (2000.Q1 ~ 2004.Q4)*			
Traditional	3,993	509	4,502
Manufacturing	2,450	411	2,861
Electronics	8,854	191	9,045
Panel B: *Out-of-sample (2005.Q1 ~ 2005.Q3)*			
Traditional	629	63	692
Manufacturing	432	38	470
Electronics	1,990	72	2,062

where P_{ji} is a probability estimate for firm i in rating j. The combining method proposed by Kamstra and Kennedy (1998) consists of finding, via MLE, an appropriate weighted average of ω's in ordered logit and ordered probit techniques.[1]

To validate the model, we use cumulative accuracy profile (CAP) and its summary statistics, the accuracy ratio (AR). A concept similar to the CAP is the receiver operating characteristic (ROC) and its summary statistics, the area under the ROC curve (AUC). In addition, we also employ the McFadden's R^2 (pseudo R^2) to evaluate the performance of the credit rating model. McFadden's R^2 is defined as $1-$(unrestricted log-likelihood function/restricted log-likelihood function).

27.3 Empirical Results

Data are collected from the Taiwan Economic Journal (TEJ) database for the period between the first quarter in 2000 and the third quarter in 2005, with the last three quarters used for out-of-sample tests. The sample consists of firms traded in the Taiwan Security Exchange (TSE) and the OTC market.

The credit rating of the sample firms is determined by the Taiwan Corporate Credit Risk Index (the TCRI). Among ten credit ratings, 1–4 represent the investment grade levels, 5–6 represent the low-risk levels, and 7–9 represent the high-risk or speculative grade levels. The final rating, 10, represents the bankruptcy level.

Table 27.1 exhibits the descriptive statistics for the samples which are divided into three industry categories. Panel A contains the in-sample observations, while panel B shows the out-of-sample observations. There are 509 bankruptcy cases in the traditional industry, 411 in the manufacturing sector, and 191 in the electronics industry for in-sample data. For out-of-sample data, there are 63 in the traditional, 38 in the manufacturing, and 72 in the electronics industries, respectively. Table 27.2 displays the frequency distributions of the credit ratings for in-samples in these three industries.

[1]See Kamstra and Kennedy (1998) for the detail description.

Table 27.2 Frequency distributions of the credit ratings

Ratings	Traditional	Manufacturing	Electronics
1	10	3	156
2	47	71	259
3	151	62	252
4	380	294	1,066
5	921	321	2,343
6	1,044	559	2,736
7	645	465	1,272
8	459	338	490
9	336	337	280
10	509	411	191
Subtotal	4,502	2,861	9,045

Note: Level 10 represents the bankruptcy class

Bonfim (2009) finds that not only the firms' financial situation has a central role in explaining default probabilities, but also macroeconomic conditions are very important when assessing default probabilities over time. Based on previous studies in the literature, 62 explanatory variables including financial ratios, market conditions, and macroeconomic factors are considered. We use the hybrid stepwise method to find the best predictors in the ordered probit and ordered logit models. The combining technique using logit regression is then applied.

27.3.1 Model Estimates

27.3.1.1 Ordered Logit Model

Table 27.3 illustrates the in-sample estimation results under the ordered logit model for each industry. From the likelihood ratio, score ratio, and Wald ratio, with significant level at 1 %, we can determine the goodness-of-fit for each model.

For the traditional industry, the coefficients of Fixed Asset to Long Term Funds Ratio (Fixed Asset to Equity and Long Term Liability Ratio), Interest Expense to Sales Ratio, and Debt Ratio are positive. It shows that firms with higher ratios will get worse credit ratings as well as higher default probabilities. On the other hand, the coefficients of Accounts Receivable Turnover Ratio, Net Operating Profit Margin, Return on Total Assets (Ex-Tax, Interest Expense), Depreciation to Sales Ratio, Free Cash Flow to Total Debt Ratio, Capital Spending to Gross Fixed Assets Ratio, Retained Earning to Total Assets Ratio, and Ln (Total Assets/GNP price-level index) are negative so that firms tend to have good credit qualities as well as lower default probabilities when these ratios become higher. All these signs meet our expectation.

For the manufacturing industry, the coefficient of the dummy variable for the Negative Net Income for the last 2 years is positive, so the losses worsen the credit rating. On the other hand, the coefficients of Equity to Total Asset Ratio, Total Assets Turnover Ratio, Return on Total Assets (Ex-Tax, Interest Expense),

Table 27.3 Regression results estimated by the ordered logit. This table represents the regression results estimated by the ordered logit model. Panel A shows the 11 explanatory variables fitted in the traditional industry. Panel B shows the nine explanatory variables fitted in the manufacturing industry. Panel C shows the ten explanatory variables fitted in the electronics industry

		Parameters	
Explanatory variables		Estimates	Standard errors
Panel A: Traditional			
X7	Fixed Asset to Long Term Funds Ratio	0.968***	(0.098)
X12	Accounts Receivable Turnover Ratio	−0.054***	(0.0087)
X19	Net Operating Profit Margin	−5.135***	(0.546)
X27	Return on Total Assets (Ex-Tax, Interest Expense)	−5.473***	(0.934)
X29	Depreciation to Sales Ratio	−6.662***	(0.537)
X30	Interest Expense to Sales Ratio	22.057***	(1.862)
X35	Free Cash Flow to Total Debt Ratio	−0.971***	(0.094)
X41	Capital Spending to Gross Fixed Assets Ratio	−0.746***	(0.167)
X46	Debt Ratio	4.781***	(0.305)
X47	Retained Earning to Total Assets Ratio	−5.872***	(0.336)
X50	Ln (Total Assets/GNP price-level index)	−6.991***	(1.024)
Panel B: *Manufacturing*			
X2	Equity to Total Asset Ratio	−7.851***	(0.317)
X15	Total Assets Turnover Ratio	−0.714***	(0.172)
X27	Return on Total Assets (Ex-Tax, Interest Expense)	−7.520***	(0.932)
X40	Accumulative Depreciation to Gross Fixed Assets	−1.804***	(0.199)
X47	Retained Earning to Total Assets Ratio	−3.138***	(0.340)
X50	Ln (Total Assets/GNP price-level index)	−5.663***	(1.213)
X52	1:If Net Income was Negative for the Last Two Years 0: Otherwise	1.274***	(0.108)
X54	Ln (Age of the firm)	−0.486***	(0.102)
X60	Ln (Net Sales)	−1.063***	(0.065)
Panel C: *Electronics*			
X2	Equity to Total Asset Ratio	−3.239***	(0.207)
X27	Return on Total Assets (Ex-Tax, Interest Expense)	−6.929***	(0.349)
X30	Interest Expense to Sales Ratio	25.101***	(1.901)
X35	Free Cash Flow to Total Debt Ratio	−0.464***	(0.031)
X40	Accumulative Depreciation to Gross Fixed Assets	−0.781***	(0.127)
X44	Cash Reinvestment Ratio	−0.880***	(0.167)
X45	Working Capital to Total Assets Ratio	−3.906***	(0.179)
X47	Retained Earning to Total Assets Ratio	−3.210***	(0.174)
X49	Market Value of Equity/Total Liability	−0.048***	(0.004)
X50	Ln (Total Assets/GNP price-level index)	−8.797***	(0.704)

***Represents significantly different from zero at 1 % level
**Represents significantly different from zero at 5 % level
*Represents significantly different from zero at 10 % level

Accumulative Depreciation to Gross Fixed Assets, Retained Earning to Total Assets Ratio, Ln (Total Assets/GNP price-level index), Ln (Age of the firm), and Ln (Net Sales) are all negative. The ratios improve the credit standings.

For the electronics industry, the coefficient of Interest Expense to Sales Ratio is positive, and the coefficients of Equity to Total Asset Ratio, Return on Total Assets (Ex-Tax, Interest Expense), Free Cash Flow to Total Debt Ratio, Accumulative Depreciation to Gross Fixed Assets, Cash Reinvestment Ratio, Working Capital to Total Assets Ratio, Retained Earning to Total Assets Ratio, Market Value of Equity/Total Liability, and Ln (Total Assets/GNP price-level index) are negative.

In general, the matured companies like those in the traditional and the manufacturing industries should focus mainly on their capabilities in operation and in liquidity. Also, the market factors are important for the manufacturing firms. For high-growth industry like the electronics, we should pay more attention to their market factors and the liquidity ratios. The common explanatory variables among three industries seem related to the operating returns, the retained earnings, and the asset size.

27.3.1.2 Ordered Probit Model

Table 27.4 shows the in-sample estimation results using the ordered probit model for each industry.

For the traditional industry, the coefficients of Fixed Asset to Long Term Funds Ratio, Interest Expense to Sales Ratio, and Debt Ratio are positive in the ordered probit model, which are similar to the results from the ordered logit model. On the other hand, the coefficients of Accounts Receivable Turnover Ratio, Net Operating Profit Margin, Return on Total Assets (Ex-Tax, Interest Expense), Depreciation to Sales Ratio, Operating Cash Flow to Total Liability Ratio, Capital Spending to Gross Fixed Assets Ratio, and Retained Earning to Total Assets Ratio are negative, also showing no big difference with the results from the ordered logit model.

For the manufacturing industry, the coefficients of Accounts Payable Turnover Ratio and the dummy variable for the negative Net Income for the last 2 years are positive. On the other hand, the coefficients of Equity to Total Asset Ratio, Quick Ratio, Total Assets Turnover Ratio, Return on Total Assets (Ex-Tax, Interest Expense), Accumulative Depreciation to Gross Fixed Assets, Cash Flow Ratio, Retained Earning to Total Assets Ratio, and Ln (Age of the firm) are negative. The results are also similar to those of the ordered logit model, only that the ordered logit seems focusing more on the size factors (sales, asset), while the ordered probit concerns more on the liquidity (quick ratio, cash flow ratio, and payables) of the firm.

For the electronics industry, the coefficients of Interest Expense to Sales Ratio is positive and the coefficients of Free Cash Flow to Total Debt Ratio, Working Capital to Total Assets Ratio, Retained Earning to Total Assets Ratio, Market Value of Equity/Total Liability, LN (Total Assets/GNP price-level index), and LN (Age) are negative.

Table 27.5 shows the threshold values estimated by the two models. Threshold values represent the cutting points for neighboring ratings.

Table 27.4 Regression results estimated by the ordered probit. This table represents the regression results estimated by the ordered probit model. Panel A shows the ten explanatory variables fitted in the traditional industry. Panel B shows the ten explanatory variables fitted in the manufacturing industry. Panel C shows the nine explanatory variables fitted in the electronics industry

		Parameters	
Explanatory variables		Estimates	Standard errors
Panel A: Traditional			
X7	Fixed Asset to Long Term Funds Ratio	0.442***	(0.058)
X12	Accounts Receivable Turnover Ratio	−0.027***	(0.005)
X19	Net Operating Profit Margin	−2.539***	(0.311)
X27	Return on Total Assets (Ex-Tax, Interest Expense)	−3.335***	(0.534)
X29	Depreciation to Sales Ratio	−3.204***	(0.287)
X30	Interest Expense to Sales Ratio	9.881***	(1.040)
X34	Operating Cash Flow to Total Liability Ratio	−0.514***	(0.121)
X41	Capital Spending to Gross Fixed Assets Ratio	−0.459***	(0.096)
X46	Debt Ratio	3.032***	(0.188)
X47	Retained Earning to Total Assets Ratio	−3.317***	(0.192)
Panel B: *Manufacturing*			
X2	Equity to Total Asset Ratio	−4.648***	(0.183)
X10	Quick Ratio	−0.036***	(0.009)
X11	Accounts Payable Turnover Ratio	0.021***	(0.004)
X15	Total Assets Turnover Ratio	−0.742***	(0.072)
X27	Return on Total Assets (Ex-Tax, Interest Expense)	−4.175***	(0.522)
X40	Accumulative Depreciation to Gross Fixed Assets	−0.989***	(0.112)
X43	Cash Flow Ratio	−0.190***	(0.045)
X47	Retained Earning to Total Assets Ratio	−1.804***	(0.188)
X52	1:If Net Income was Negative for the Last Two Years 0: Otherwise	0.753***	(0.061)
X54	Ln (Age of the firm)	−0.289***	(0.056)
Panel C: *Electronics*			
X2	Equity to Total Asset Ratio	−1.851***	(0.117)
X27	Return on Total Assets (Ex-Tax, Interest Expense)	−3.826***	(0.197)
X30	Interest Expense to Sales Ratio	11.187***	(1.052)
X35	Free Cash Flow to Total Debt Ratio	−0.233***	(0.017)
X45	Working Capital to Total Assets Ratio	−2.101***	(0.101)
X47	Retained Earning to Total Assets Ratio	−1.715***	(0.098)
X49	Market Value of Equity/Total Liability	−0.027***	(0.002)
X50	Ln (Total Assets/GNP price-level index)	−4.954***	(0.399)
X54	Ln (Age of the firm)	−0.247***	(0.028)

***Represents significantly different from zero at 1 % level
**Represents significantly different from zero at 5 % level
*Represents significantly different from zero at 10 % level

Table 27.5 Threshold values estimated by the ordinal analysis. This table shows the threshold values estimated by the ordered logit and the ordered probit models. There are nine threshold parameters given ten credit ratings

Threshold parameter	Ordered logit model			Ordered probit model		
	Traditional	Manufacturing	Electronics	Traditional	Manufacturing	Electronics
τ1	−23.361***	−28.610***	−31.451***	−12.272***	−16.891***	−17.306***
	(0.633)	(0.987)	(0.478)	(0.332)	(0.447)	(0.252)
τ2	−21.236***	−24.879***	−29.759***	−11.231***	−15.042***	−16.410***
	(0.545)	(0.766)	(0.459)	(0.294)	(0.360)	(0.243)
τ3	−19.441***	−23.848***	−28.855**	−10.316***	−14.446***	−15.917***
	(0.519)	(0.749)	(0.451)	(0.282)	(0.352)	(0.240)
τ4	−17.639***	−21.657***	−26.846***	−9.339***	−13.175***	−14.797**
	(0.502)	(0.717)	(0.436)	(0.275)	(0.337)	(0.233)
τ5	−15.389***	−20.333***	−24.451***	−8.080***	−12.411***	−13.441***
	(0.484)	(0.703)	(0.421)	(0.268)	(0.331)	(0.227)
τ6	−13.317***	−18.309***	−21.777***	−6.918***	−11.251***	−11.937***
	(0.473)	(0.683)	(0.407)	(0.263)	(0.322)	(0.221)
τ7	−11.584***	−16.461***	−19.715***	−5.972***	−10.207***	−10.809***
	(0.465)	(0.670)	(0.400)	(0.261)	(0.314)	(0.218)
τ8	−9.825***	−14.738***	−18.019***	−5.049***	−9.246***	−9.929***
	(0.459)	(0.662)	(0.397)	(0.260)	(0.308)	(0.217)
τ9	−8.231***	−12.166***	−16.010***	−4.256***	−7.860***	−9.005***
	(0.457)	(0.655)	(0.400)	(0.260)	(0.301)	(0.219)

***Represents significantly different from zero at 1 % level
**Represents significantly different from zero at 5 % level
*Represents significantly different from zero at 10 % level

27.3.2 Credit Rating Forecasting

Tables 27.6 and 27.7 illustrate the prediction results of the ordered logit and the ordered probit models. Following Blume et al. (1998), we define the most probable rating as the actual rating or its immediate adjacent ratings. The ratio of the number of the predicted ratings as the most probable ratings to the total number of the ratings being predicted can assess the goodness-of-fit for the model. For out-of-sample firms, the predictive power of the ordered logit model for each industry is 86.85 %, 81.06 %, and 86.37 %, respectively; and the predictive power of the ordered probit model for each industry is 86.42 %, 80.21 %, and 84.87 %, respectively. The results from two models are quite similar.

27.3.3 Estimation Results Using the Combining Method

Table 27.8 depicts the regression results using the Kamstra-Kennedy combining forecasting technique. The coefficients are the logit estimates on ω's. These ω values are all positive and strongly significant for each industry.

Table 27.6 Out-of-sample predictions by the ordered logit model

Panel A: *Traditional*

		Predicted rating									
		1	2	3	4	5	6	7	8	9	10
Actual rating	1	0	0	3	0	0	0	0	0	0	0
	2	0	0	0	6	5	1	0	0	0	0
	3	0	0	4	8	12	0	0	0	0	0
	4	0	0	4	26	34	7	0	0	0	0
	5	0	0	0	14	101	41	2	0	0	0
	6	0	0	0	2	53	104	29	5	0	0
	7	0	0	0	0	5	24	26	11	0	0
	8	0	0	0	0	0	21	26	13	2	3
	9	0	0	0	0	0	1	4	15	7	10
	10	0	0	0	0	0	3	2	9	15	34
Prediction ratio	86.85%										

Panel B: *Manufacturing*

		Predicted rating									
		1	2	3	4	5	6	7	8	9	10
Actual rating	1	0	0	0	0	0	0	0	0	0	0
	2	0	3	0	4	3	2	0	0	0	0
	3	0	2	2	4	2	5	0	0	0	0
	4	0	0	0	7	21	24	0	1	0	0
	5	0	0	0	5	29	51	5	3	0	0
	6	0	0	0	8	20	68	18	3	0	0
	7	0	0	1	1	0	18	38	9	3	1
	8	0	0	0	0	0	4	16	5	9	1
	9	0	0	0	0	0	0	9	12	10	5
	10	0	0	0	0	0	0	4	5	8	21
Prediction ratio	81.06%										

Panel C: *Electronics*

		Predicted rating									
		1	2	3	4	5	6	7	8	9	10
Actual rating	1	0	1	2	0	15	0	0	0	0	0
	2	0	0	0	5	22	2	0	0	0	0
	3	0	0	0	1	36	11	2	0	0	0
	4	0	2	4	13	144	42	0	0	0	0
	5	0	1	1	9	279	180	15	1	0	0
	6	0	0	0	11	293	303	39	2	0	0
	7	0	0	0	0	16	206	79	9	0	1
	8	0	0	0	0	2	50	80	26	10	2
	9	0	0	0	0	0	6	11	21	22	13
	10	1	0	0	0	3	2	5	12	15	34
Prediction ratio	86.37%										

Table 27.7 Out-of-sample predictions by the ordered probit model

Panel A: Traditional											
	Predicted rating										
		1	2	3	4	5	6	7	8	9	10
Actual rating	1	0	2	1	0	0	0	0	0	0	0
	2	0	0	0	4	6	2	0	0	0	0
	3	0	0	3	5	16	0	0	0	0	0
	4	0	0	6	22	36	7	0	0	0	0
	5	0	0	0	10	97	48	3	0	0	0
	6	0	0	0	2	45	106	36	4	0	0
	7	0	0	0	0	4	28	21	13	0	0
	8	0	0	0	0	1	19	26	14	2	3
	9	0	0	0	0	1	0	5	10	9	12
	10	0	0	0	0	0	2	3	11	15	32
Prediction ratio	86.42%										
Panel B: *Manufacturing*											
	Predicted rating										
		1	2	3	4	5	6	7	8	9	10
Actual rating	1	0	0	0	0	0	0	0	0	0	0
	2	1	2	0	4	5	0	0	0	0	0
	3	1	1	2	4	2	5	0	0	0	0
	4	0	0	0	9	20	23	1	0	0	0
	5	0	0	0	6	28	51	5	3	0	0
	6	0	0	0	7	22	71	15	2	0	0
	7	0	0	0	1	1	23	34	8	3	1
	8	0	0	0	0	0	6	15	4	9	1
	9	0	0	0	0	0	0	12	8	11	5
	10	0	0	0	0	0	1	4	5	8	20
Prediction ratio	80.21%										
Panel C: *Electronics*											
	Predicted rating										
		1	2	3	4	5	6	7	8	9	10
Actual rating	1	0	0	0	3	15	0	0	0	0	0
	2	0	0	0	1	26	2	0	0	0	0
	3	0	0	0	0	40	8	2	0	0	0
	4	0	0	0	5	155	45	0	0	0	0
	5	0	0	0	7	295	164	20	0	0	0
	6	0	0	0	8	313	286	40	1	0	0
	7	0	0	0	0	11	207	87	5	0	1
	8	0	0	0	0	2	54	96	14	2	2
	9	0	0	0	0	0	5	26	30	6	6
	10	0	0	0	0	3	4	12	21	8	24
Prediction ratio	84.87%										

Table 27.8 Regression results estimated by the combining forecast

Threshold parameter	Combining forecasting model					
	Traditional		Manufacturing		Electronics	
$\tau1$	−13.1875	(0.3033)***	−13.6217	(0.5475)***	−11.6241	(0.1591)***
$\tau2$	−11.3484	(0.1854)***	−9.9864	(0.1918)***	−10.5484	(0.1483)***
$\tau3$	−9.8997	(0.1524)***	−9.1712	(0.1704)***	−9.973	(0.1456)***
$\tau4$	−8.3781	(0.1370)***	−7.5337	(0.1441)***	−8.7037	(0.143)***
$\tau5$	−6.4390	(0.1257)***	−6.4459	(0.1315)***	−7.0973	(0.1414)***
$\tau6$	−4.5627	(0.1137)***	−4.8259	(0.1142)***	−5.0093	(0.1364)***
$\tau7$	−3.1037	(0.1016)***	−3.3724	(0.1003)***	−3.1906	(0.1261)***
$\tau8$	−1.5619	(0.0885)***	−2.0849	(0.0913)***	−1.5682	(0.1145)***
$\tau9$	−0.0816	(0.0861)	−0.1127	(0.0929)	0.3735	(0.1169)**
ω_{Logit}	1.1363	(0.0571)***	1.2321	(0.029)***	0.6616	(0.0528)***
ω_{Probit}	−0.0825	(0.0330)*	−0.1437	(0.0085)***	0.1867	(0.0272)***

Numbers in parentheses represent the standard errors
***Represents significantly different from zero at 1 % level
**Represents significantly different from zero at 5 % level
*Represents significantly different from zero at 10 % level

Table 27.9 shows the prediction results of the combining forecasting model. For out-of-sample test, the predictive power of the combining model for each industry is 89.88 %, 82.77 %, and 88.02 %, respectively, which are higher than those of the ordered logit or ordered probit models by 2–4 %.

27.3.4 Performance Evaluation

To evaluate the performance of each model, Fig. 27.1 illustrates the ROC curves estimated by the three models, respectively.

From these ROC curves we can distinguish the performance of each rating model. Furthermore, we can compare the AUC and AR calculated from the ROC and CAP (See Table 27.10).

For the traditional industry, the AUCs from the ordered logit, the ordered probit, and the combining model are 95.32 %, 95.15 %, and 95.32 %, respectively. For the manufacturing industry, they are 94.73 %, 93.66 %, and 95.51 %, respectively. And for the electronics industry, they are 92.43 %, 92.30 %, and 94.07 %, respectively. These results apparently show that the combining forecasting model performs better than any individual one.

27.4 Conclusion

This study constitutes an attempt to explore the credit rating forecasting techniques. The samples consist of firms in the TSE and the OTC market and

Table 27.9 Out-of-sample credit rating prediction by the combining forecast

Panel A: Traditional

		Predicted rating									
		1	2	3	4	5	6	7	8	9	10
Actual rating	1	0	0	2	1	0	0	0	0	0	0
	2	0	0	0	6	6	0	0	0	0	0
	3	0	0	4	7	13	0	0	0	0	0
	4	0	0	2	28	34	7	0	0	0	0
	5	0	0	0	12	100	44	2	0	0	0
	6	0	0	0	2	48	113	26	4	0	0
	7	0	0	0	0	5	27	23	11	0	0
	8	0	0	0	0	0	23	25	12	2	3
	9	0	0	0	0	0	1	4	14	8	10
	10	0	0	0	0	0	3	2	9	16	33
Prediction ratio	89.88%										

Panel B: *Manufacturing*

		Predicted rating									
		1	2	3	4	5	6	7	8	9	10
Actual rating	1	0	0	0	0	0	0	0	0	0	0
	2	0	2	1	3	6	0	0	0	0	0
	3	0	1	2	5	2	5	0	0	0	0
	4	0	0	0	6	24	22	0	1	0	0
	5	0	0	0	5	30	50	5	3	0	0
	6	0	0	0	6	23	70	17	1	0	0
	7	0	0	1	1	0	27	32	7	2	1
	8	0	0	0	0	1	3	16	6	8	1
	9	0	0	0	0	0	1	8	12	11	4
	10	0	0	0	0	0	0	4	4	7	23
Prediction ratio	82.77%										

Panel C: *Electronics*

		Predicted rating									
		1	2	3	4	5	6	7	8	9	10
Actual rating	1	0	3	0	0	15	0	0	0	0	0
	2	0	0	0	4	18	7	0	0	0	0
	3	0	0	0	1	34	13	2	0	0	0
	4	0	4	4	13	128	56	0	0	0	0
	5	0	1	1	11	256	193	23	1	0	0
	6	0	0	0	12	248	338	47	3	0	0
	7	0	0	0	0	14	190	91	15	0	1
	8	0	0	0	0	1	44	82	32	10	1
	9	0	0	0	0	0	5	10	23	24	11
	10	0	0	0	0	3	1	6	11	17	34
Prediction ratio	88.02%										

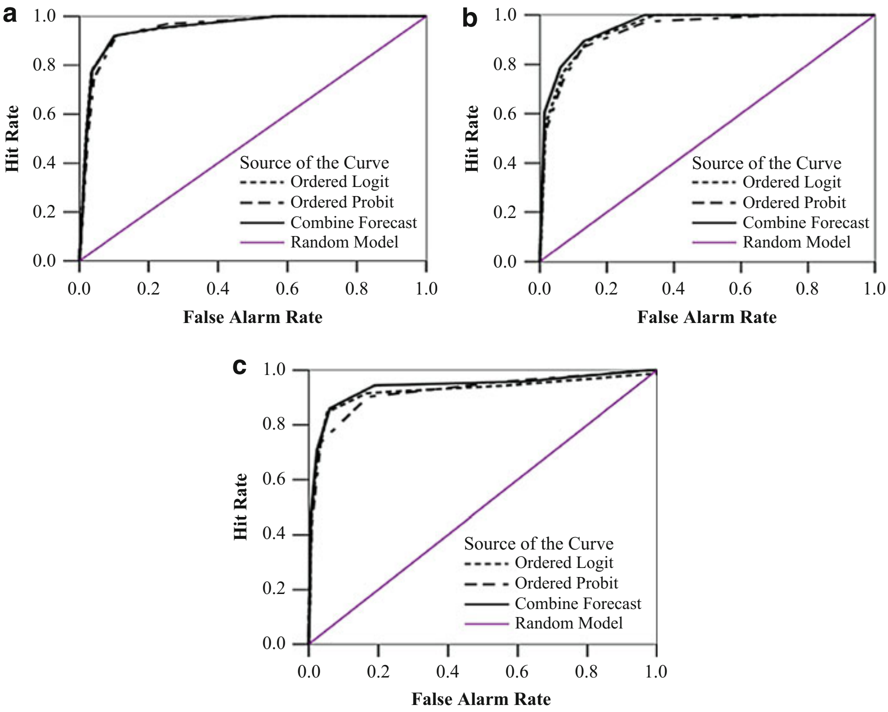


Fig. 27.1 ROC *curves*, (**a**) Traditional, (**b**) Manufacturing, (**c**) Electronics

Table 27.10 Performance evaluation for each model

	Traditional	Manufacturing	Electronics
Panel A: Ordered logit			
AUC	95.32 %	94.73 %	92.43 %
AR	90.63 %	89.46 %	84.86 %
McFadden's R-square	35.63 %	38.25 %	39.63 %
Panel B: Ordered probit			
AUC	95.15 %	93.66 %	92.30 %
AR	90.30 %	87.32 %	84.60 %
McFadden's R-square	34.45 %	40.05 %	41.25 %
Panel C: Combining forecasting			
AUC	95.32 %	95.51 %	94.07 %
AR	90.63 %	91.03 %	88.15 %
McFadden's R-square	42.34 %	43.16 %	46.28 %

CAP represents the cumulative accuracy profile, *AR* represents accuracy ratio. McFadden's R^2 is defined as 1−(unrestricted log-likelihood function/restricted log-likelihood function) are divided into three industries, i.e., traditional, manufacturing, and electronics, for analysis. Sixty-two explanatory variables consisting of financial, market, and macroeconomics factors are considered. We utilize the ordered logit, the ordered probit, and the combining forecasting model to estimate the parameters and conduct the out-of-sample tests. The main result is that the combining forecasting method leads to a more accurate rating prediction than that of any single use of the ordered logit or ordered probit analysis. By means of cumulative accuracy profile, the receiver operating characteristics, and McFadden R^2, we can measure the goodness-of-fit and the accuracy of each prediction model. These performance evaluations depict consistent results that the combining forecast performs better.

Appendix 1: Ordered Probit Procedure for Credit Rating Forecasting[2]

Credit rating is an ordinal scale from which the credit category of a firm can be ranked from high to low, but the scale of the difference between them is unknown. To model the ordinal outcomes, we follow Zavoina and McKelvey (1975) to begin with a latent regression

$$Y^* = X\beta + \varepsilon \tag{27.8}$$

where

$$Y^* = \begin{bmatrix} Y_1^* \\ \vdots \\ Y_N^* \end{bmatrix}, X = \begin{bmatrix} 1 & X_{11} & \cdots & X_{K1} \\ \vdots & \vdots & & \vdots \\ 1 & X_{1N} & \cdots & X_{KN} \end{bmatrix}$$

$$\beta = \begin{bmatrix} \beta_0 \\ \vdots \\ \beta_K \end{bmatrix}, \varepsilon = \begin{bmatrix} \varepsilon_0 \\ \vdots \\ \varepsilon_N \end{bmatrix}.$$

Here, β is a vector of unknown parameters, X is a set of explanatory variables, and ε is a random disturbance term assumed to follow the multivariate normal distribution with mean 0 and variance-covariance matrix $\sigma^2 I$, that is,

$$\varepsilon \sim N\left(0, \sigma^2 I\right). \tag{27.9}$$

Y^*, the dependent variable of theoretical interests, is unobserved, but from which we can classify the category j:[3]

[2]There are detailed discussion about ordered data in Ananth and Kleinbaum (1997), McCullagh (1980), Wooldridge (2010), and Greene (2011).

[3]In our case, $J = 10$ and $K = 62$.

$$Y_i = j \text{ if } \tau_{j-1} < Y_i^* \leq \tau_j \ (i = 1, 2, \ldots, N; j = 1, 2, \ldots, J) \tag{27.10}$$

where Y is the one we do observe, an ordinal version of Y^*, and the τ's are unknown parameters to be estimated with β. We assume that $-\infty = \tau_0 \leq \tau_1 \leq \cdots \leq \tau_J = +\infty$.

From Eqs. 27.8 and 27.10, we have

$$\begin{aligned} \tau_{j-1} < Y_i^* \leq \tau_j &\Leftrightarrow \tau_{j-1} < X_i\beta + \varepsilon_i \leq \tau_j \\ &\Leftrightarrow \frac{\tau_{j-1} - X_i\beta}{\sigma} < \frac{\varepsilon_i}{\sigma} \leq \frac{\tau_j - X_i\beta}{\sigma} \end{aligned} \tag{27.11}$$

where X_i is the i_{th} row of X.

From Eqs. 27.9 and 27.11, the probability of $Y_i = j$ can be written as

$$\Pr(Y_i = j) = \Phi\left(\frac{\tau_j - X_i\beta}{\sigma}\right) - \Phi\left(\frac{\tau_{j-1} - X_i\beta}{\sigma}\right) \tag{27.12}$$

where $\Phi(.)$ represents the cumulative density function of standard normal distribution. The model (27.12) is under-identified since any linear transformation of the underlying scale variable Y^*, if applied to the parameters and $\tau_0, \ldots, \tau_J$ as well, would lead to the same model. We will assume without loss of generality that, $\tau_1 = 0$ and $\sigma = 1$ in order to identify the model. The model we will estimate turns out to be

$$\Pr(Y_i = j) = \Phi\big(\tau_j - X_i\beta\big) - \Phi\big(\tau_{j-1} - X_i\beta\big). \tag{27.13}$$

Maximum likelihood estimation can be used to estimate the $J + K - 1$ parameters, $\tau_2, \ldots, \tau_{J-1}$ and $\beta_0, \beta_1, \ldots, \beta_K$, in Eq. 27.13. To form the likelihood function, first, define a dummy variable $Y_{i,j}$:

$$Y_{i,j} = \begin{cases} 1 \text{ if } Y_i = j \\ 0 \text{ otherwise} \end{cases}.$$

Then, for simple notation, we set $Z_{i,j} = \tau_j - X_i\beta$ and $\Phi_{i,j} = \Phi(Z_{i,j})$. The likelihood function, L, is

$$L = L(\beta_0, \ldots, \beta_K, \tau_2, \ldots, \tau_{J-1} | Y) = \prod_{i=1}^{n} \prod_{j=1}^{J} \big(\Phi_{i,j} - \Phi_{i,j-1}\big)^{Y_{i,j}}. \tag{27.14}$$

So, the log-likelihood function, ln L, is

$$\ln L = \sum_{i=1}^{n} \sum_{j=1}^{J} Y_{i,j} \ln\big(\Phi_{i,j} - \Phi_{i,j-1}\big). \tag{27.15}$$

Now, we want to maximize ln L subject to $\tau_1 \le \tau_2 \le \ldots \le \tau_{J-1}$.
Let $N_{i,j} = \frac{1}{\sqrt{2\pi}} e^{-\frac{z_{i,j}^2}{2}}$ for $1 \le j \le J$ and $1 \le i \le N$, and let $\delta_{l,j} = \begin{cases} 1 \text{ if } l = j \\ 0 \text{ if } l \ne j \end{cases}$, it follows that

$$\begin{aligned} \frac{\partial \Phi_{i,j}}{\partial \beta_u} &= N_{i,j} \frac{\partial Z_{i,j}}{\partial \beta_u} = -N_{i,j} X_{u,i} \text{ for } 0 \le u \le K \\ \frac{\partial \Phi_{i,j}}{\partial \tau_l} &= N_{i,j} \frac{\partial Z_{i,j}}{\partial \tau_l} = N_{i,j} \delta_{l,j} \text{ for } 2 \le l \le J-1 \end{aligned} \tag{27.16}$$

and

$$\begin{aligned} \frac{\partial N_{i,j}}{\partial \beta_u} &= Z_{i,j} N_{i,j} X_{u,i} \quad \text{for } 0 \le u \le K \\ \frac{\partial N_{i,j}}{\partial \tau_l} &= -Z_{i,j} N_{i,j} \delta_{l,j} \text{ for } 2 \le l \le J-1 \end{aligned} \tag{27.17}$$

By using Eqs. 27.16 and 27.17, we can calculate the $J + K - 1$ partial derivatives of Eq. 27.15 with respect to the unknown parameters, β and τ, respectively:

$$\begin{aligned} \frac{\partial \ln L}{\partial \beta_u} &= \sum_{i=1}^{n} \sum_{j=1}^{J} Y_{i,j} \frac{(N_{i,j-1} - N_{i,j}) X_{u,i}}{\Phi_{i,j} - \Phi_{i,j-1}} \qquad \text{for } 0 \le u \le K \\ \frac{\partial \ln L}{\partial \tau_l} &= \sum_{i=1}^{n} \sum_{j=1}^{J} Y_{i,j} \frac{(N_{i,j}\delta_{l,j} - N_{i,j-1}\delta_{l,j-1}) X_{u,i}}{\Phi_{i,j} - \Phi_{i,j-1}} \quad \text{for } 2 \le l \le J-1. \end{aligned} \tag{27.18}$$

And the elements in the $(J + K - 1) \times (J + K - 1)$ matrix of second partials are

$$\frac{\partial^2 \ln L}{\partial \beta_u \partial \beta_V} = \sum_{i=1}^{n} \sum_{j=1}^{J} Y_{i,j} \frac{(\Phi_{i,j} - \Phi_{i,j-1})(Z_{i,j-1} N_{i,j-1} - Z_{i,j} N_{i,j}) - (N_{i,j-1} - N_{i,j})^2 X_{u,i} X_{v,i}}{(\Phi_{i,j} - \Phi_{i,j-1})^2},$$

$$\frac{\partial^2 \ln L}{\partial \beta_u \partial \tau_l} = \frac{\partial^2 \ln L}{\partial \tau_l \partial \beta_u} = \sum_{i=1}^{n} \sum_{j=1}^{J} Y_{i,j} \left[\frac{(\Phi_{i,j} - \Phi_{i,j-1})(Z_{i,j} N_{i,j} \delta_{l,j} - Z_{i,j-1} N_{i,j-1} \delta_{l,j-1})}{(\Phi_{i,j} - \Phi_{i,j-1})^2} - \frac{(N_{i,j-1} - N_{i,j})(N_{i,j}\delta_{l,j} - N_{i,j-1}\delta_{l,j-1})}{(\Phi_{i,j} - \Phi_{i,j-1})^2} \right] X_{u,i},$$

$$\frac{\partial^2 \ln\ L}{\partial \tau_l \partial \tau_m} = \sum_{i=1}^{n}\sum_{j=1}^{J} Y_{i,j}\left[\frac{\left(\Phi_{i,j}-\Phi_{i,j-1}\right)\left(Z_{i,j-1}N_{i,j-1}\delta_{m,j-1}\delta_{l,j-1}-Z_{i,j}N_{i,j}\delta_{m,j}\delta_{l,j}\right)}{\left(\Phi_{i,j}-\Phi_{i,j-1}\right)^2} - \frac{\left(N_{i,j}\delta_{l,j}-N_{i,j-1}\delta_{l,j-1}\right)\left(N_{i,j}\delta_{m,j}-N_{i,j-1}\delta_{m,j-1}\right)}{\left(\Phi_{i,j}-\Phi_{i,j-1}\right)^2}\right]. \tag{27.19}$$

We then set the $J + K - 1$ equations in Eq. 27.18 to zero to get the MLE of the unknown parameters. The matrix of second partials should be negative definite to insure that the solution is a maximum. The computer program NPROBIT, which uses the Newton-Raphson method, can solve the nonlinear equations in Eq. 27.18.

Appendix 2: Ordered Logit Procedure for Credit Rating Forecasting

Consider a latent variable model where Y^* is the unobserved dependent variable, X a set of explanatory variables, β an unknown parameter vector, and ε a random disturbance term:

$$Y^* = X\beta + \varepsilon \tag{27.20}$$

ε is assumed to follow a standard logistic distribution, so the probability density function of ε is

$$\lambda(\varepsilon) = \frac{\exp(\varepsilon)}{[1+\exp(\varepsilon)]^2} \tag{27.21}$$

and the cumulative density function is

$$\Lambda(\varepsilon) = \frac{\exp(\varepsilon)}{1+\exp(\varepsilon)}. \tag{27.22}$$

Y^*, the dependent variable of theoretical interests, is unobserved, but from which we can classify the category j:

$$Y_i = j \ \text{ if } \ \tau_{j-1} < Y_i^* \le \tau_j \ \ (i = 1, 2, \ldots, N; j = 1, 2, \ldots, J) \tag{27.23}$$

where Y is the one we do observe, an ordinal version of Y^*, and the τ's are unknown parameters satisfying $\tau_1 \le \ldots \le \tau_J$ and to be estimated with β.

From Eqs. 27.20 and 27.22, we form the proportional odds model:

$$\Pr(Y_i \leq j|X) = \frac{\exp\left(\tau_j - X_i\beta\right)}{1 + \exp\left(\tau_j - X_i\beta\right)} \tag{27.24}$$

or equivalently

$$\begin{aligned} \log \mathrm{it}\left(\Pi_{ij}\right) &= \log\left[\frac{\Pi_{ij}}{1 - \Pi_{ij}}\right] \\ \log\left[\frac{\Pr(Y_i \leq j|X_i)}{\Pr(Y_i > j|X_i)}\right] &= \tau_j - X_i\beta \end{aligned} \tag{27.25}$$

where $\Pi_{ij} = \Pr(Y_i \leq j)$. Notice that in the proportional odds model, β is assumed to be constant and not depend on j. The validity of this assumption can be checked based on a χ^2 score test. The model that relaxes the proportional odds assumption can be represented as

$$\log \mathrm{it}\left(\Pi_{ij}\right) = \tau_j - X_i\beta_j, \tag{27.26}$$

where the regression parameter vector β is allowed to vary with j. Both models can be fit through the procedure of maximum likelihood estimation.

Appendix 3: Procedure for Combining Probability Forecasts

To combine the ordered logit and the ordered probit models for credit forecasting, the logit regression method as described in Kamstra and Kennedy (1998) is applied. We first assume that firm's credit classification is determined by an index θ. Suppose there are J rating classes, ordered from 1 to J. If θ exceeds the threshold value τ_j, $j = 1,\ldots J - 1$, credit classification changes from j rating to $j + 1$ rating. The probability of company i being in rating j is given by the integral of a standard logit from $\tau_{j-1} - \theta_i$ to $\tau_j - \theta_i$.

Each forecasting method is considered as producing $J - 1$ measures, $\omega_{ji} = \tau_j - \theta$, $j = 1, \ldots, J - 1$ for each firm. These measures can be estimated as

$$\omega_{ji} = \ln\left[\frac{P_{1i} + \cdots + P_{ji}}{1 - P_{1i} - \cdots - P_{ji}}\right] \tag{27.27}$$

where P_{ji} is a probability estimate for firm i in rating j. For firm i we have

$$P_{ji} = \begin{cases} \dfrac{e^{\omega_{ji}}}{1 + e^{\omega_{ji}}} & \text{for } j = 1 \\ \dfrac{e^{\omega_{ji}}}{1 + e^{\omega_{ji}}} - \dfrac{e^{\omega_{j-1i}}}{1 + e^{\omega_{j-1i}}} & \text{for } j = 2\cdots J - 1\,. \\ \dfrac{1}{1 + e^{\omega_{j-1i}}} & \text{for } j = J \end{cases} \tag{27.28}$$

In our case, there are two forecasting techniques A and B. For firm i, the combining probability estimate is

$$P_{ji} = \begin{cases} \dfrac{e^{\pi_j+\pi_A\omega_{jiA}+\pi_B\omega_{jiB}}}{1+e^{\pi_j+\pi_A\omega_{jiA}+\pi_B\omega_{jiB}}} & \text{for } j=1 \\ \dfrac{e^{\pi_j+\pi_A\omega_{jiA}+\pi_B\omega_{jiB}}}{1+e^{\pi_j+\pi_A\omega_{jiA}+\pi_B\omega_{jiB}}} - \dfrac{e^{\pi_{j-1}+\pi_A\omega_{j-1iA}+\pi_B\omega_{j-1iB}}}{1+e^{\pi_{j-1}+\pi_A\omega_{j-1iA}+\pi_B\omega_{j-1iB}}} & \text{for } j=2\cdots J-1\cdot \\ \dfrac{1}{1+e^{\pi_{j-1}+\pi_A\omega_{j-1iA}+\pi_B\omega_{j-1iB}}} & \text{for } j=J \end{cases} \tag{27.29}$$

This can be estimated using an ordered logit software package with the ω_A and ω_B values as explanatory variables and the τ_1 and π_2 parameters playing the role of the unknown threshold values.

References

Altman, E. I. (1968). Financial ratios discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance, 23*, 589–609.

Altman, E. I., & Katz, S. (1976). Statistical bond rating classification using financial and accounting data. In *Proceedings of the conference on topical research in accounting*, New York.

Altman, E. I., & Sounders, A. (1997). Credit risk measurement: Developments over the last 20 years. *Journal of Banking and Finance, 21*, 1721–1742.

Altman, E. I., Hadelman, R. G., & Narayanan, P. (1977). Zeta analysis, a new model to identify bankruptcy risk of corporations. *Journal of Banking and Finance, 10*, 29–51.

Ananth, C. V., & Kleinbaum, D. G. (1997). Regression models for ordinal responses: A review of methods and applications. *International Journal of Epidemiology, 26*, 1323–1333.

Beaver, W. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research, 4*, 71–111.

Blume, M. E., Lim, F., & Mackinlay, A. C. (1998). The declining credit quality of U.S. corporate debt: Myth or reality? *Journal of Finance, 53*, 1389–1413.

Bonfim, D. (2009). Credit risk drivers: Evaluating the contribution of firm level information and of macroeconomic dynamics. *Journal of Banking and Finance, 33*, 281–299.

Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting, 5*, 559–583.

Crouhy, M., Galai, D., & Mark, R. (2000). A comparative analysis of current credit risk models. *Journal of Banking and Finance, 24*, 59–117.

Ederington, L. H. (1985). Classification models and bond ratings. *The Financial Review, 20*, 237–261.

Granger, C. W. (1989). Invited review: Combining forecasts-twenty years later. *Journal of Forecasting, 8*, 167–173.

Greene, W. H. (2011). *Econometric analysis*. NJ: Prentice Hall.

Hand, D., & Henley, W. (1997). Statistical classification methods in consumer credit scoring: A review. *Journal of the Royal Statistical Society, Series A, 160*, 523–541.

Kamstra, M., & Kennedy, P. (1998). Combining qualitative forecasts using logit. *International Journal of Forecasting, 14*, 83–93.

Kaplan, R. S., & Urwitz, G. (1979). Statistical models of bond ratings: A methodological inquiry. *Journal of Business, 52*, 231–261.

Lawrence, E. C., & Arshadi, N. (1995). A multinomial logit analysis of problem loan resolution choices in banking. *Journal of Money, Credit and Banking, 27*, 202–216.

Mahmoud, E. (1984). Accuracy in forecasting: A survey. *Journal of Forecasting, 3*, 139–159.

McCullagh, P. (1980). Regression models for ordinal data. *Journal of the Royal Statistical Society, 42*, 109–142.

Ohlson, J. A. (1980). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research, 18*, 109–131.

Pinches, G. E., & Mingo, K. A. (1973). A multivariate analysis of industrial bond ratings. *Journal of Finance, 28*, 1–18.

Pompe, P. M., & Bilderbeek, J. (2005). The prediction of bankruptcy of small- and medium-sized industrial firms. *Journal of Business Venturing, 20*, 847–868.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). Cambridge, MA: MIT Press.

Zavoina, R., & McKelvey, W. (1975). A statistical model for the analysis of ordinal level dependent variables. *Journal of Mathematical Sociology, 4*, 103–120.

Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research, 22*, 59–68.

Can We Use the CAPM as an Investment Strategy?: An Intuitive CAPM and Efficiency Test

28

Fernando Gómez-Bezares, Luis Ferruz, and Maria Vargas

Contents

Abstract

The aim of this chapter is to check whether certain playing rules, based on the undervaluation concept arising from the CAPM, could be useful as investment strategies, and can therefore be used to beat the Market. If such strategies work, we will be provided with a useful tool for investors, and, otherwise, we will

The authors wish to thank Professor José Vicente Ugarte for his assistance in the statistical analysis of the data presented in this chapter.

F. Gómez-Bezares
Universidad de Deusto, Bilbao, Spain
e-mail: f.gomez-bezares@deusto.es

L. Ferruz (✉)
Facultad de Economía y Empresa, Departamento de Contabilidad y Finanzas, Universidad de Zaragoza, Zaragoza, Spain
e-mail: lferruz@unizar.es

M. Vargas
Universidad de Zaragoza, Zaragoza, Spain
e-mail: mvargas@unizar.es

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_28,

obtain a test whose results will be connected with the Efficient Market Hypothesis (EMH) and with the CAPM.

The basic strategies were set out in Gómez-Bezares, Madariaga, and Santibáñez (Análisis Financiero 68:72–96, 1996). Our purpose now is to reconsider them, to improve the statistical analysis, and to examine a more recent period for our study.

The methodology used is both intuitive and rigorous: analyzing how many times we beat the Market with different strategies, in order to check whether beating the Market happens by chance. Furthermore, we set out to study, statistically, when and by how much we beat it, and to analyze whether this is significant.

Keywords

ANOVA • Approximately normal distribution • Binomial distribution • CAPM • Contingency tables • Market efficiency • Nonparametric tests • Performance measures

28.1 Introduction

Finance, as it is currently taught in business schools and brought together in the most prestigious text books, places great importance on the Capital Asset Pricing Model (CAPM)[1] and on the Efficient Market Hypothesis (EMH)[2]. For example, Brealey, Myers, and Allen (2008) conclude their famous book by saying that within what we know about Finance, there are seven key ideas: the first of which is Net Present Value (NPV), the second CAPM, and the third EMH[3]; leaving aside for the moment NPV,[4] we have as the two clearly outstanding concepts the CAPM and the EMH. In our opinion, asset pricing models (above all the CAPM) and the efficiency of the Markets are among the foundations of the current paradigm which has been in vogue since the 1970s.[5] And this is perfectly logical; the financial objective of a company is to maximize its Market value[6]; to this end it must make investments with a positive NPV, and to calculate the NPV financiers require an asset pricing model such as the CAPM; ultimately the Markets need to be efficient so that they can notice increases in value provided by investments. We can thus see by this simple reasoning how the three concepts highlighted by Brealey, Myers, and Allen are interrelated.

[1]Sharpe (1964), Lintner (1965).

[2]One of its main defenders has been Fama (1970, 1991, 1998).

[3]There are also other four key ideas.

[4]A fundamental concept, predating the other two and clearly related to the CAPM and the EMH.

[5]See Gómez-Bezares and Gómez-Bezares (2006). Also of interest could be Dimson and Mussavian (1998, 1999).

[6]The interesting work by Danielson, Heck and Shaffer (2008) is also worth looking at.

An alternative way of viewing this is to say that Finance, as it is currently understood, is the science of pricing; we have to valuate so that we can know which decisions will result in the greatest added value, and to valuate we need asset pricing models (like the CAPM). Finally, the result of our labor will be recognized by the Market, if the Market is efficient and values more highly those stocks with a higher intrinsic value.

The Golden Age of these two principles was the decade of the 1970s, appearing in works such as Fama (1970), Black, Jensen, and Scholes (1972) and Fama and MacBeth (1973): the Market correctly values the assets (the EMH) and we have a good model of asset pricing (the CAPM). However, since then there have been many critiques of both of these principles: the efficient Market principle has been criticized by psychologists and behavioral economists, who question the rationality of human beings, as well as by econometricians, who claim that prices are susceptible to prediction (Malkiel 2003). There is ample criticism in the literature, but the EMH also has important apologists; among the classic defenders are the works of Fama (1991, 1998) as well as the very noteworthy and recent study by Fama and French (2010), in which they show that portfolio managers, on average, do not beat the Market, thereby proving that the Market is indeed efficient.

The detractors of the efficient Market hypothesis (currently, above all, psychologists and behavioral economists) underline again and again the inefficiencies they observe. These problems with the model, when they are seen to occur repeatedly, are termed "anomalies" (for a summary, see Malkiel 2003). Fama and French (2008) studied different anomalies, concluding that some are more significant than others. However, they concluded that we cannot be sure whether to attribute abnormal returns to inefficiencies of the Market or to rewards for risk-taking.

This brings us back to the debate regarding asset pricing models. According to the CAPM, the expected return on an asset should be a positive linear function of its systematic risk measured by beta, which is the sole measurement of risk. This statement has been questioned in many ways; one very important contribution in this respect is that of Fama and French (1992) in which they comment that the beta can contribute little to an explanation of expected returns and, in fact, that there are other variables that can explain a lot more. Their work gave rise to the famous Fama and French three-factor model. The existence of variables other than the beta which can help to explain mean returns is not in itself an argument against the efficiency of the Market if we consider them to be risk factors: as risk is our enemy we demand higher returns for higher risk, and we can measure this risk in terms of a range of factors. This approach may be compatible with Arbitrage Pricing Theory (APT), but it conflicts with the CAPM.

But we could also consider that the problem is that the Markets are inefficient, and hence expected returns respond to variables that are not risk factors, variables to which they should not react. Let us take as example momentum, which tells us that recent past high (or low) returns can help to predict high (or low) returns in the near future. Some may think that this could be used as a risk factor to explain returns whereas others would say that it is an inefficiency of the Market. The truth is that the asset pricing model and efficiency are tested jointly (Fama 1970, 1991, 1998), in such a way that, when they work, they both work,[7] but when they fail we cannot know which of them has failed.[8]

The CAPM tests fall far short of giving definitive results. Although some have given it up for dead and buried because it cannot fit the data or because of theoretical problems, one recent magnificent study by Levy (2010) rejects the theoretical problems posed by psychologists and behavioral economists and, based on ex ante data, says that there is experimental support for the CAPM.

Brav, Lehavy, and Michaely (2005), in an approach related to the above, set out to test the model, not by looking at past returns (as a *proxy* of expected returns) but instead at the expectations of Market analysts, on the basis that these may be assumed to be unbiased estimates of the Market expectations. From their analysis, they found a positive relationship between expected returns and betas.

Asset pricing models (such as the CAPM) look at the relationship between return and risk. Recently, there has been more focus on liquidity. Liu (2006) builds an enhanced CAPM based on two factors: the Market and liquidity – he obtains their corresponding factorial weighting which he uses to explain the expected returns. The model works correctly and allows him to account several anomalies; according to Liu, his enhanced CAPM lends new support to the risk-return paradigm.

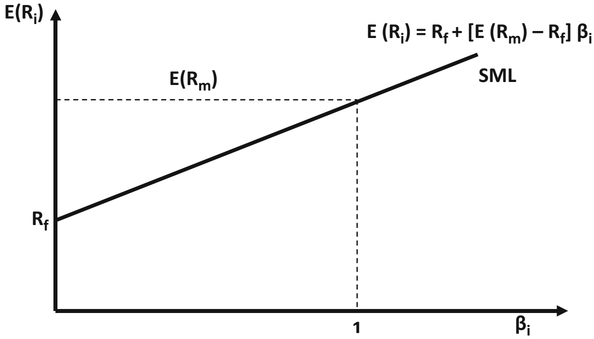
The methods most commonly used to contrast the CAPM are time-series and cross-sectional studies, which each present different statistical problems. What we set out to achieve in this chapter is to replicate a simple and intuitive but also rigorous methodology that we believe is much less problematic from a statistical point of view. The procedure was first proposed by Gómez-Bezares, Madariaga, and Santibáñez (1996), and in this chapter we attempt to replicate it with a more updated sample and greater statistical rigor. The basic idea is simple: an individual who knows the CAPM calculates at the beginning of each month the return that each stock has rendered in the past and compares it with the return it ought to have given according to the CAPM. The stocks which gave higher returns than they ought to have are cheap (undervalued) while those which gave lower return are expensive (overvalued). If we assume that return levels are going to persist, the investor should buy the cheap stocks and sell off the expensive ones in order to beat the Market. We illustrate this line of reasoning in Fig. 28.1.

In Fig. 28.1, the betas appear on the horizontal axis and the expected return on the vertical axis (R_f is the riskless rate; R_m is the Market return). We also show the Security Market Line (SML), with its formula above it, which is the formula for the CAPM. Our investor decides to buy the stocks which are cheap and refrain from selling short the expensive stocks (due to the limitations on short selling). Based on the data at his disposal, the stocks which have gained a value more than that the SML indicates are undervalued, that is to say, they are above

[7]We refer to the situation we have described previously which occurred at the beginning of the 1970s.

[8]See also Copeland, Weston, and Shastri (2005, p. 244).

Fig. 28.1 shows the Security Market Line (SML). The betas appear on the horizontal axis, and the expected return on the vertical axis



the SML. Believing that this situation will continue, he will buy all of these stocks, trusting that he will obtain an adjusted return higher than the Market return. If this happens (which he can find out by using Jensen's alpha), and this strategy consistently proves to result in abnormal returns (positive Jensen alphas, or to put it another way, the stocks stay above the SML in the next period), he will have found a way to beat the Market; therefore it is not efficient. Moreover, he shall be able to predict which stocks will, in the future, outperform the SML and therefore do not comply with the CAPM, which goes against the CAPM. There would then be one way left to save the EMH: risk should not be measured using the beta (i.e., the CAPM is mistaken) and therefore a positive value for alpha is not synonymous with beating the Market (which could not be done consistently in an efficient Market); we could also save the CAPM as follows: risk must be measured with the beta; however, the Market fails to value stocks accurately and hence it can be beaten; what we cannot do is to save both the EMH and the CAPM simultaneously.

On the other hand, if our investor cannot consistently beat the Market with his strategy but can only achieve about a 50 % success rate, more or less, we must conclude that the portfolio assembled with the previously undervalued stocks sometimes performs better than the SML and other times not as well, purely by chance, and then settles down to the value it ought to have according to the CAPM. We will not have a Market-beating strategy, and the results are compatible with the EMH; likewise, we will see that on a random basis the portfolios perform better than the SML at times and other times worse, which is compatible with the CAPM (every month there may be random oscillations around the SML); therefore two key elements of the aforementioned paradigm would be rescued.

In our research, we use Jensen's (1968, 1969) and Treynor's (1965) indices, since they consider systematic risk. These indices have been very widely used in the literature.[9]

[9]Recent studies that use these indices are, for example, Samarakoon and Hasan (2005), Abdel-Kader and Qing (2007), Pasaribu (2009), Mazumder and Varela (2010), and Ferruz, Gómez-Bezares, and Vargas (2010).

One of the virtues of this method is that it is perfectly replicable, as we have hypothesized an investor who makes decisions based on information in the public domain whenever he makes them. Meanwhile, we feel, reasonably enough, that the beta values and the risk premium may vary over time. We also use the most liquid stocks (to avoid problems with lack of liquidity) and portfolios (which reduces the measuring problems).

The results of our analysis clearly indicate that our strategy is not capable of beating the Market consistently, and therefore it is compatible with both the CAPM and the EMH. The result, to which we have come via different routes, which supports the strategy's robustness, supports the results reported by Fama and French (2010) and Levy (2010) although they used very different methods. Doubtless, both the CAPM and the EMH are simplifications of reality, but they help us to explain that reality.

The rest of our chapter is organized as follows: in Sect. 28.2, we describe the sample analyzed with information about the assets that it comprises, and we comment on the method used to calculate the returns and betas; in Sect. 28.3, we comment on the method employed, with reference to the proposed strategies and how these were formulated, we then go on to analyze the scenarios in which the strategies succeed in beating the Market and we describe the statistical tests used for the analyses; in Sect. 28.4, we show the results obtained by the different strategies and we analyze the number of occasions on which we were able to beat the Market and whether these results occur by chance. Furthermore, we analyze when and in what magnitude we beat the Market; in Sect. 28.5, we summarize the main conclusions drawn, highlighting the implications of our results in relation to the possibilities of using strategies to beat the Market according to the CAPM, and, therefore, we set out our conclusions as to the efficiency of the Market and the validity of the CAPM. In addition, there is a list of the references on which our research is based and an Appendix which contains, in greater detail, the statistical evidence gathered during our study.

28.2 Data

Our analysis is based on the 35 securities that comprise the IBEX 35 (the official index of the Spanish Stock Exchange Market) in each month from practically the beginning of trading in Continuous Market in Spain (1989) up to March 2007.[10] We chose this index because it includes the 35 most liquid companies in the Spanish Market and, taking into account that the CAPM in Sharpe-Lintner's version is valid for liquid companies, it would have been difficult to find such companies in Spain outside the IBEX 35.

[10]We have not extended our analysis beyond March 2007 in order to exclude the period of the global financial crisis which had a very damaging effect on major companies listed on the IBEX 35, such as banks, as this could have distorted our results.

The IBEX 35 is a capitalization-weighted index comprising the 35 most liquid companies which quote in the Continuous Market in Spain. It fulfills all of the criteria required of an indicator which aspires to be a benchmark for trading: it is representative (almost 90 % of the trading volume in the Continuous Market and approximately 80 % of the Market capitalization is held by the 35 firms listed on the IBEX 35), it can be replicated (ease of replication), its computation is guaranteed, it is widely publicized and impartial (supervised by independent experts).

The selection criteria for securities listed on the IBEX 35 are as follows:

- Securities must be included in the SIBE trading system.
- Stocks must be representative with a big Market capitalization and trading volume.
- There must be a number of "floating" stocks which is sufficient to ensure that the index's Market capitalization is sufficiently widespread and allows for hedge and arbitrage strategies in the Market for derivatives on the IBEX 35.
- The average Market capitalization of a stock measurable on the index[11] must be greater than 0.30 % of the average capitalization of the index during the monitoring period.[12]
- The stock must have been traded in at least 1/3 of the sessions during the monitoring period or be among the top 15 stocks measured in terms of Market capitalization.

The IBEX 35 is revised on a half-yearly basis in terms of its composition and the number of stocks considered of each firm; nevertheless, if financial operations are carried out which affect significantly any of the listed stocks, it can be adjusted accordingly. In general, the index is adjusted when there are increases in capital with preemptive rights, extraordinary dividend distributions, stocks integration as a result of increases in capital excluding preemptive rights, reductions in capital due to stocks redemption, capital reductions against own funds with distribution of the value to the shareholders (this is not the payment of an ordinary dividend), as well as mergers, takeovers, and demergers. Special adjustments of the index are often carried out.

In the 6-monthly selection of the 35 most liquid stocks, there is no minimum or maximum number of adjustments made with regard to the previous period. No changes at all may be required, or as many adjustments as necessary may be made, based on the results obtained when measuring the liquidity.

Having described the context in which we intend to carry out our study, we will propose a series of strategies that will enable us to test the validity of the CAPM and the degree to which the Market is efficient. For this we will take a hypothetical investor who, each month, examines the 35 stocks of the IBEX 35 listed at that moment and for which there are at least 36 monthly data available prior to that moment.

[11]The average Market capitalization of a stock in the Index is the arithmetic average of the result we obtain when multiplying the stocks allowed to be traded in each session during the monitoring period by the closing price of the stock in each of those sessions.

[12]The monitoring period is the 6-month period ending prior to each ordinary meeting of the Commission.

The data on returns for all the stocks were obtained from *Bloomberg*.

In Table 28.1, we show the annual descriptive statistics from our database.

There are two panels in Table 28.1: in panel A we show the descriptive statistics for the monthly returns of the previous period (December 1989 to November 1992). In this period, the index did not exist as such, so we have looked at the return of those stocks which, being included in the IBEX 35 at the beginning of the contrasting period, provide data in the corresponding month as they then were quoted in the Continuous Market. In panel B, we bring together the descriptive

Table 28.1 Annual descriptive statistics for the monthly returns on the stocks

PANEL A: PRECEDING PERIOD					
YEAR	**MEAN**	**MEDIAN**	**MAXIMUM**	**MINIMUM**	**St. Dev.**
DEC 1989/NOV 1990	0.187118	−0.002995	11.17	−0.359280	1,195
DEC 1990/NOV 1991	0.010761	0.0000	0.470758	−0.288640	0.095169
DEC 1991/NOV 1992	−0.013669	−0.011575	0.448602	−0.375127	0.113660
DEC 1989/NOV 1992	**0.054038**	**−0.004545**	**11.17**	**−0.375127**	**0.661538**
PANEL B: CONTRASTING PERIOD					
YEAR	**MEAN**	**MEDIAN**	**MAXIMUM**	**MINIMUM**	**St. Dev.**
DEC 1992/NOV 1993	0.035059	0.028576	0.343220	−0.237244	0.083789
DEC 1993/NOV 1994	0.009628	−0.000857	0.510076	−0.604553	0.101984
DEC 1994/NOV 1995	0.006480	0.002051	0.554368	−0.220966	0.077322
DEC 1995/NOV 1996	0.022723	0.022042	0.204344	−0.164185	0.062786
DEC 1996/NOV 1997	0.038109	0.034819	0.429381	−0.308638	0.097809
DEC 1997/NOV 1998	0.027631	0.028225	0.432660	−0.331141	0.120737
DEC 1998/NOV 1999	−0.001611	−0.003229	0.351075	−0.223282	0.080742
DEC 1999/NOV 2000	0.002103	−0.005529	0.556553	−0.357094	0.110788
DEC 2000/NOV 2001	0.003081	0.000000	0.468603	−0.282024	0.098028
DEC 2001/NOV 2002	−0.005393	−0.002908	0.369085	−0.403408	0.104471
DEC 2002/NOV 2003	0.013977	0.016353	0.551835	−0.347280	0.083299
DEC 2003/NOV 2004	0.018440	0.015852	0.297322	−0.147819	0.050610
DEC 2004/NOV 2005	0.023066	0.014685	0.257841	−0.110474	0.057863
DEC 2005/NOV 2006	0.029634	0.022335	0.372409	−0.169830	0.062640
DEC 2006/MAR 2007	0.021989	0.013010	0.264933	−0.285306	0.068472
DEC 1992/MAR 2007	**0.016678**	**0.012686**	**0.556553**	**−0.604553**	**0.087537**

Table 28.1 reports the annual descriptive statistics for our database. In Panel A, we show the figures for the monthly returns for the previous period and in Panel B the figures for monthly returns for the contrasting period

For Panel B, we consider only the returns on stocks which we have included in our study, that is to say, those stocks which were part of the IBEX 35 at a given time and which had a track record of monthly returns of at least 36 months. For panel A, as the index as such did not yet exist, we include those stocks listed on the IBEX 35 at the beginning of the contrasting period, for which there is data in the corresponding previous month; for example, the stocks we included for the month of December 1989 are those which were quoted on the IBEX 35 in December 1992 but which also were quoted in the Continuous Market back in December 1989. For the subsequent months the number of stocks considered in the study rises consistently since the number of stocks quoted continued to grow data for the monthly returns during the contrasting period, that is, it includes the returns on the stocks which are included in our study (namely, the stocks which comprised the IBEX 35 at each moment and for which there were at least 36 prior quotations).

We built the Market portfolio,[13] and as a risk-free asset we took the monthly return on the 1-month Treasury Bills.

28.3 Methods

The aim of our study is to check whether it is possible to obtain abnormal returns using the CAPM, that is to say, whether the returns derived from the use of the model are greater than those which could be expected according to the degree of systematic risk undertaken.

To this end, we analyzed the period December 1989 to March 2007. The study focuses on the 35 securities comprising the IBEX 35 at any given moment during the said period.

From this starting point, we propose two possible strategies for testing the efficient Market hypothesis and the validity of the CAPM: the first strategy assumes an individual who adjusts his portfolio at the end of each month, selling those stocks he bought at the end of the previous month and buying those he considers to be undervalued according to the CAPM; the second strategy is similar to the first, but with the distinction that in this case the investor does not buy all of the undervalued stocks but just 75 % of them, namely, those which are furthest removed from the SML. We use two methods to calculate which securities are the most undervalued: Jensen's ratio and Treynor's ratio. The reason for using a second method is in response to the critique by Modigliani (1997) of Jensen's alpha, in which she argues that differences in return cannot be compared when the risks are significantly different.

To conduct our analysis, we begin by dividing the overall period analyzed in our study (December 1989 to March 2007) into two subperiods: the first (December 1989 to November 1992) allows us to calculate the betas of the model at the beginning of the following (we carried out this calculation with the 36 previous month data). The second subperiod (December 1992[14] to March 2007) allows us to carry out the contrast.

[13]This has been constructed as a simple average of returns for the stocks comprising the IBEX 35 each month, corrected for dividend distributions, splits, etc. Of course, in this case, it is not necessary that there be at least 36 months of prior quotations for the stocks to be included in the portfolio.

[14]Initially, we intended to let the beginning of the second subperiod be the launch date of the IBEX 35 (January 1992), however, given the requirement that we set ourselves of a prior 36-month period to compute the betas of the model, we had to move this forward to December. Likewise, it was originally our intention that the beginning of the first subperiod be the launch date of the Continuous Market (April 1989) but most of the securities included in our study were listed for the first time in December 1989, so data could only be obtained from that date onward.

So our hypothetical investor will firstly, at the end of each month, observe the stocks that comprise the IBEX 35[15] and will buy those which are undervalued (in our second analysis method he will buy just 75 % of the stocks which are the most undervalued). A stock is undervalued if its return is higher than what is ought to be based on the CAPM.

To obtain this we compute the monthly betas for each stock (β_i) during the contrasting period by regressing the monthly returns on each stock[16] on the returns on the Market portfolio[17] in the 36 months immediately prior to this. Then, we calculate the mean monthly returns on stock i ($\overline{\Re_i}$), on the Market portfolio ($\overline{R_M}$) and on the risk-free asset ($\overline{R_F}$) during the same period as a simple average of the corresponding 36 monthly returns. From all of the foregoing we can compute the mean monthly return ($\overline{R_i}$), which, according to the CAPM, the stock ought to have gained during this 36-month period,

$$\overline{R_i} = \overline{R_F} + \left[\overline{R_M} - \overline{R_F}\right]\beta_i \tag{28.1}$$

We then compare it with the actual return gained ($\overline{\Re_i}$) to determine whether the stock is undervalued.

Once we have determined which stocks are undervalued, according to our first strategy the investor buys all of the undervalued stocks each month,[18] while with the second strategy, he disregards the first quartile and just buys 75 % of the most undervalued stocks. The quartiles are assembled based on either Jensen's alphas or Treynor's ratio.[19]

The next step consists in evaluating how well the CAPM functions; if the undervalued stocks continue to be undervalued indefinitely, we would expect any

[15]And for which there is a minimum of 36-monthly data prior to the month analyzed so that the beta can be calculated. In the case of companies which have recently merged, the beta is calculated by looking at the weighted average monthly returns on the companies prior to the merger. The weighting is proportional to the relative importance of each company involved in the merger.

[16]We look at the return obtained by our investor from capital gains, dividends, and the sale of preemptive rights. The returns are also adjusted to take splits into account.

[17]The return on the Market portfolio is calculated based on a simple average of the returns gained by the stocks which comprise the IBEX 35 in each month. The reason we decided to build our own portfolio rather than relying on a stock Market index such as the IBEX 35 itself is that we wanted to obtain an index corrected by capital gains, dividends, splits, and preemptive rights. Moreover, an equally weighted portfolio is theoretically superior.

[18]Overvalued stocks are not included in our analysis as we decided to exclude short selling.

[19]In this second variant, although the quartiles are built based on Treynor's ratio, we continue to use Jensen's index to work out whether or not a stock is undervalued. This is because of the problems we have encountered when determining which stocks are undervalued using Treynor's ratio. These are related to the return premiums and the negative betas which appear in some cases in our database. Naturally, we take on board the problems of using Treynor's ratio to construct the quartiles, as there could be stocks with positive alphas and low Treynor ratios which would exclude such stocks from the analysis. This is due to the form of Treynor's ratio as a quotient.

portfolio assembled with them to beat the Market in a given month, but this will not occur if in that month the CAPM operates perfectly.

Jensen's index allows us to determine whether we have been able to beat the Market, in which case it must yield a positive result:

$$\alpha_p = \left(R_p - R_F\right) - \beta_p(R_M - R_F) \tag{28.2}$$

where,

β_p is the beta for the portfolio, calculated as a simple average of the individual betas[20] of the stocks included in the portfolio (all of the undervalued stocks, or alternatively the top 75 % thereof as we have outlined above).

R_p is the return on the portfolio and is calculated as the simple average of the individual returns obtained each month for the stocks that comprise it.

R_M and R_F are the monthly returns on the Market portfolio and the risk-free asset respectively, for the month in question.

α_p is Jensen's alpha for the portfolio p.

The Z-statistic, which follows an approximately normal distribution, (0.1), allows us to determine whether the number of months in which our investor beats the Market is due to chance:

$$Z = (Y - np)/\sqrt{np(1-p)} \tag{28.3}$$

Y indicates the number of periods in which the portfolio comprising undervalued stocks beats the Market, n represents the number of months analyzed and to p we give a value of 0.5 as this is the probability of beating the Market if the CAPM and the efficient Market hypothesis are fulfilled,[21] and we want to find out if the difference (Y – np) is due to chance.

A value of $|Z| > 1.96$ would lead us to reject the null hypothesis of a difference due to chance, with a significance level of 5 %.

Moreover, in order to make our results more robust, we developed a series of statistical tests which are presented in Appendix. We carried out a mean test for each of the proposed strategies to ascertain whether we beat the Market or not, both parametric – which means normality – as well as nonparametric (Wilcoxon's test); in addition we tested the hypothesis that we could beat the Market 50 % of the time, and for this we used a binomial test. We also analyzed, on a monthly and on a yearly basis, whether or not there were differences between the performance means achieved in each month/year, using the ANOVA, a nonparametric test (Kruskal and Wallis), and also the technique of contingency tables and the chi-square

[20]Which have been calculated with the 36 previous months' data.

[21]If the CAPM and the efficient Market hypothesis are fulfilled exactly, the results of any portfolio (adjusted for risk) should match, as an average, with those of the Market. Here we suppose that, by pure chance, for any given month, they will end up 50 % above and 50 % below the Market average.

Pearson test together with the likelihood ratio. Finally, we conducted regression analysis of the performance measurement as a dependent variable and the time dimension (months) as an independent variable to see whether the performance improved or worsened over time. These tests allow us to confirm our results, which we shall see in the next section.

28.4 Results

28.4.1 Strategy 1: The Investor Buys All the Undervalued Stocks

In this first strategy, the investor observes, in a given month, the track record (for the previous 36 months) of the 35 stocks which comprise the IBEX 35 at the moment in question,[22] and buys all those that are undervalued according to Jensen's index. The next step is to check whether he has managed to beat the Market with the portfolio of undervalued stocks.

Table 28.2 shows the results of this strategy. We show for each of the 14 years in our study and for the other 4 months the number of months (and the percentage) in which the investor's portfolio has beaten the Market. In addition, we show this result for the whole period. Furthermore, and also with respect to the whole period, we show the result for the Z-statistic.

As we can see in Table 28.2, only in 92 of the 172 months analyzed in our study is the Market beaten, which is equal to 53.49 %[23]; moreover, the figure for Z-statistic (which is below 1.96) confirms that the CAPM does not offer a significantly better strategy than investing in the Market portfolio. In other words, we could say that the success of the investor (in the months in which he succeeds in beating the Market) could be due to chance. This conclusion is further confirmed in Appendix by the mean tests which allow us to accept that the mean of Jensen's alphas is zero and by the binomial test which allows accepting a success probability rate of 50 %.

If we now focus on the analysis of each year in our time frame, we can see that in only two of the 14 years covered in our study, to be precise from December 2004 to November 2006, the CAPM proves to be a useful tool for the investor allowing him to use a strategy to beat the Market; in fact, in 75 % and 83.33 % of the months of each of those 2 years our investor beats the Market, thus confirming the poor performance of the model, or the inefficiency of the Market.

[22]Which have a track record of at least 36 months in the Market. If, in a given month, there is a stock which is listed on the IBEX35 but does not have at least 36 months of prior data, this will be included in our study in the month in which it reaches this figure, unless by that stage it has already been deleted from the IBEX35.

[23]We have not included the transaction costs, which undoubtedly would be higher for the managed portfolio and, therefore, would make that strategy less attractive.

Table 28.2 Results of the Strategy 1 (buying all of the undervalued stocks)

YEAR	No. of successful months	% success	
DEC 1992/NOV 1993	4	33.33%	
DEC 1993/NOV 1994	4	33.33%	
DEC 1994/NOV 1995	5	41.67%	
DEC 1995/NOV 1996	7	58.33%	
DEC 1996/NOV 1997	7	58.33%	
DEC 1997/NOV 1998	6	50%	
DEC 1998/NOV 1999	7	58.33%	
DEC 1999/NOV 2000	6	50%	
DEC 2000/NOV 2001	6	50%	
DEC 2001/NOV 2002	5	41.67%	
DEC 2002/NOV 2003	7	58.33%	
DEC 2003/NOV 2004	7	58.33%	
DEC 2004/NOV 2005	9	75%	
DEC 2005/NOV 2006	10	83.33%	
DEC 2006/MAR 2007	2	50%	Z-statistic
DEC 1992/MAR 2007	**92**	**53.49%**	0.91

In Table 28.2, we bring together the results of a strategy in which our hypothetical investor buys all the undervalued stocks on the Index. To be specific, we provide for each of the 14 years in our study as well as for the other 4 months, the number of months (and the percentage) in which the portfolio beats the Market. We also show this result for the period as a whole. Moreover, for the period as a whole, we show the result for the Z-statistic

In the first 2 years (from December 1992 to November 1994), the opposite happens: in just 33.33 % of the months the investor's strategy beats the Market. In those months it seems, therefore, that he could use the opposite strategy to beat it.

The statistical tests carried out and included in the Appendix do not provide us with clear conclusions as to the basis for the differences in performance of the investor's strategy between the different months and years, because the different tests used gave us different results. However, the regression analysis does confirm that the performance of the portfolio improves over time, as we obtain a positive and significant beta.

28.4.2 Strategy 2: The Investor Buys the Three Stock-Quartiles with the Highest Degree of Undervaluation

With this strategy, which has two variants, what lets us find out whether a stock is or is not undervalued is, just as with the previous strategy, the difference between the mean monthly return actually achieved by the stock in a given month (using information on the previous 36 months) and that which it ought to have achieved according to the CAPM. However, with this second strategy, the investor disregards the first quartile of undervalued stocks and buys the other 75 %, those that are the most undervalued.

In the 1st variant of this second strategy, the ranking of the stocks (in each month) by quartiles is done using Jensen's index and in the 2nd variant it is done using Treynor's ratio.

28.4.2.1 Strategy 2, 1st Variant: The Quartiles Are Built Based on Jensen's Index

In Table 28.3, we show the results of this strategy:

The results, although slightly better, are not very different from those achieved by the previous strategy, as once again the CAPM proves not to be a useful tool for beating the Market. In only 94 of the 172 months in our study (54.65 %) does the strategy manage to beat the Market. The value for Z-statistic again confirms this result. And again, the mean tests show that the mean for the alphas for the investor's portfolio can be zero; while the binomial test shows that the probability of success in beating the Market can be 50 % (see Appendix).

There is also a slight improvement in the results by years in comparison with the results achieved by the previous strategy; the CAPM proves to be a useful Market-beating tool in 5 of the 14 years analyzed (December 1995 to November 1996, December 1998 to November 1999, December 2002 to November 2003, December 2004 to November 2005, and December 2005 to November 2006), which implies that the model either works badly or that the Market is inefficient since it is possible to beat it in most months.

Table 28.3 Results of Strategy 2, 1st Variant (buying the top three quartiles of the most undervalued stocks, constructed based on Jensen's Index)

YEAR	No. of successful months	% success	
DEC 1992/NOV 1993	3	25%	
DEC 1993/NOV 1994	4	33.33%	
DEC 1994/NOV 1995	5	41.67%	
DEC 1995/NOV 1996	8	66.67%	
DEC 1996/NOV 1997	6	50%	
DEC 1997/NOV 1998	7	58.33%	
DEC 1998/NOV 1999	8	66.67%	
DEC 1999/NOV 2000	6	50%	
DEC 2000/NOV 2001	5	41.67%	
DEC 2001/NOV 2002	5	41.67%	
DEC 2002/NOV 2003	9	75%	
DEC 2003/NOV 2004	6	50%	
DEC 2004/NOV 2005	10	83.33%	
DEC 2005/NOV 2006	10	83.33%	
DEC 2006/MAR 2007	2	50%	Z-statistic
DEC 1992/MAR 2007	**94**	**54.65%**	1.22

Table 28.3 shows the results of the strategy in which the investor disregards the first quartile of undervalued stocks and buys the remaining 75 %, which are the stocks that are most undervalued. The quartiles are built using Jensen's Index

The opposite results are obtained for the first 2 years (from December 1992 to November 1994), in which the percentages of success (25 % and 33.33 %, respectively) are the lowest.

The ANOVA, the Kruskal–Wallis test, and the contingency table with Pearson's chi-square-statistic and the likelihood ratio all lead us to the conclusion[24] that there are no differences between the performance means achieved by the portfolio in the months in our analysis, so the small differences could be due to mere chance. However, these tests do not allow us to offer conclusions when we look at the years that comprise our sample, as different conclusions can be drawn from the different tests used.

Moreover, the regression analysis once again allows us to confirm that, as with the previous strategy, the performance of the portfolio improves over time.

28.4.2.2 Strategy 2, 2nd Variant: The Quartiles Are Constructed Based on Treynor's Ratio

In Table 28.4, we show the results of this strategy.

The conclusions are very similar to those which the two previous strategies lead us to, thus we can see, for the whole period, a scant success rate for the strategy.

Table 28.4 Results of Strategy 2, 2nd Variant (buying the top three quartiles of the most undervalued stocks, constructed based on Treynor's Ratio)

YEAR	No. of successful months	% success	
DEC 1992/NOV 1993	3	25%	
DEC 1993/NOV 1994	4	33.33%	
DEC 1994/NOV 1995	7	58.33%	
DEC 1995/NOV 1996	7	58.33%	
DEC 1996/NOV 1997	6	50%	
DEC 1997/NOV 1998	7	58.33%	
DEC 1998/NOV 1999	7	58.33%	
DEC 1999/NOV 2000	7	58.33%	
DEC 2000/NOV 2001	4	33.33%	
DEC 2001/NOV 2002	5	41.67%	
DEC 2002/NOV 2003	8	66.67%	
DEC 2003/NOV 2004	6	50%	
DEC 2004/NOV 2005	10	83.33%	
DEC 2005/NOV 2006	10	83.33%	
DEC 2006/MAR 2007	2	50%	Z-statistic
DEC 1992/MAR 2007	**93**	**54.07%**	1.07

Table 28.4 shows the results of the strategy in which the investor disregards the first quartile of undervalued stocks and buys the remaining 75 %, which are the stocks that are most undervalued. The quartiles are built using Treynor's Ratio

[24] See Appendix.

Moreover, the value for Z-statistic once again leads us to accept the hypothesis that any possible success or failure is due to chance. Meanwhile, the tests we include in Appendix confirm the previous results: the mean tests support the interpretation that the mean performance of the portfolios is zero, therefore we do not beat the Market, and the binomial test suggests that the strategy's success rate is 50 %.

Focusing now on the analysis by years, there are 3 years (December 2002 to November 2003 and December 2004 to November 2006) in which we can confirm that the Market is not efficient or the model does not work well, since in those years the CAPM seems to be a useful Market-beating tool. However, in the first 2 years of our study as well as in the ninth year, the lowest success rates are delivered, which does not allow us to confirm the efficient Market hypothesis, as by using the opposite strategy we could have beaten it.

With regard to the statistical tests to analyze the robustness of the above-mentioned results for the various years and months in our database and those shown in Appendix, we find that there are no significant differences between the performance means achieved by the portfolio in the different months, so any difference can be due to chance. Nonetheless, we cannot draw conclusions on a yearly basis as the tests produce different results.

Finally, once again the result of the previous strategies is confirmed: the performance of the investor's portfolio improves over time. Overall, simply looking at the graphics for the three strategies where we can see the Jensen alphas achieved by our investor over time, if the reader disregards the initial data, he will see that the adjusted line is flat.

28.5 Conclusion

The aim of our study was to analyze whether the design of strategies based on the CAPM can enable an investor to obtain abnormal returns. We also set out to explore this using a methodology that was both intuitive and scientifically rigorous.

We also set out to determine whether the efficient Market hypothesis was fulfilled and whether we could accept the validity of the CAPM, since if strategies that can beat the Market with a degree of ease exist, we cannot confirm either the efficiency of the Market or the workability of the CAPM, as this defends that the only way to obtain extra returns is to accept a higher degree of systematic risk. Conversely, we would then be in possession of a tool enabling the investor to achieve higher returns than the Market for a given level of systematic risk.

Analyzing the behavior of an investor who can buy the stocks comprising the IBEX 35 at any given moment, that is, the benchmark for the Spanish stock

Market, and who reconfigures his portfolio on a monthly basis, we find that, regardless of the strategy used (buying all of the undervalued stocks or buying the 75 % most undervalued stocks, either measured with Jensen's alpha or with Treynor's ratio), the investor manages to beat the Market about 50 % of the time. We opted to exclude from our calculations the transaction costs, so we in fact overvalued the investor's results. Therefore, we can conclude that the CAPM is not an attractive tool for an investor who wishes to achieve abnormal returns. It seems that undervalued stocks rapidly fit the SML, and so from another perspective, we can confirm the efficient Market hypothesis and the workability of the CAPM. These conclusions are backed up by a series of statistical tests included in Appendix.

A positive aspect of our study from the point of view of its applicability is that the behavior we have envisaged for our virtual investor is perfectly replicable as he acts only with the information available in any given month. Meanwhile, by focusing on the stocks that make up the IBEX 35, our study's conclusions are applicable to the most representative and consolidated stocks on the Spanish Market.

Furthermore, our system possesses interesting statistical advantages: it allows for variation of the beta and of the risk premium over time, it avoids the risk of illiquidity, and it reduces errors in measurement. We have also considered its robustness.

Finally, it is important to point out that at all times we have accepted implicitly the logic of the CAPM, when measuring performance with Jensen's index (which implies basing ourselves on the CAPM). For this reason our results enable us to confirm that it is the efficiency of the Market which prevents abnormally high returns compared to those the Market itself is able to achieve, and that the CAPM is a model that works reasonably well; hence it cannot be used as a Market-beating tool.

Appendix

In this section, we report a series of statistical tests done using the "Stata" program, which support the results obtained in our study, thereby, we believe, making it more robust.[25] These statistics are drawn up for each of the strategies we tried out: buying all of the undervalued stocks, buying the top 75 % most undervalued stocks using Jensen's Index to select them, and buying the 75 % most undervalued stocks according to Treynor's Ratio.

[25] See Agresti (2007), Anderson et al. (2011), Conover (1999) and Newbold et al. (2009) for a widening of the statistical processing used in this Appendix.

Strategy 1: Investor Buys All Undervalued Stocks

1. *Summary statistics*

We work here, as for all of the strategies, with the Jensen values recorded for each of the portfolios (172).

	Percentiles	Smallest		
1%	−0.1064671	−0.1087249		
5%	−0.0279054	−0.1064671		
10%	−0.0180511	−0.0708467	Obs.	172
25%	−0.0077398	−0.0540319	Sum of Wgt.	172
50%	0.0012046		Mean	0.001246
		Largest	Std. Dev.	0.0213603
75%	0.0123422	0.0389192		
90%	0.0237416	0.0444089	Variance	0.0004563
95%	0.0309298	0.0509601	Skewness	−1.397642
99%	0.0509601	0.0773151	Kurtosis	10.85719

We would point out that in this table the Jensen values are far from normality, as can be seen in the readings for asymmetry and kurtosis.

2. *Mean test*

Variable	Obs.	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]
jensen	172	0.001246	0.0016287	0.0213603	−0.001969 0.004461

mean = mean (jensen) t = 0.7650
Ho: mean = 0 degrees of freedom = 171

Ha: mean < 0 Ha: mean != 0 Ha: mean > 0
Pr(T < t) = 0.7773 Pr(|T| > |t|) = 0.4453 Pr(T > t) = 0.2227

From this mean test, we can accept that the mean for the Jensen alphas is zero; in fact, if we focus on the two-sided test we obtain a probability of 0.4453, higher than 5 %. This allows us to conclude that we are not able to beat the Market with this strategy.

3. *Nonparametric mean test*

sign	Obs.	sum ranks	expected
positive	92	8451	7439
negative	80	6427	7439
zero	0	0	0
all	172	14878	14878

unadjusted variance 427742.50
adjustment for ties 0.00
adjustment for zeros 0.00
----
adjusted variance 427742.50

Ho: jensen = 0
z = 1.547
Prob > |z| = 0.1218

We can see again that Wilcoxon's nonparametric test leads to the same conclusion: we cannot beat the Market with this strategy.

4. *Frequencies*

BEAT	BEAT		
	NO	YES	Total
NO	80 46.51	0 0	80 46.51
YES	0 0	92 53.49	92 53.49
Total	80 46.51	92 53.49	172 100

From this contingency table, we see that the proportion of months in which we beat the Market is very similar to the proportion in which we fail to do so.

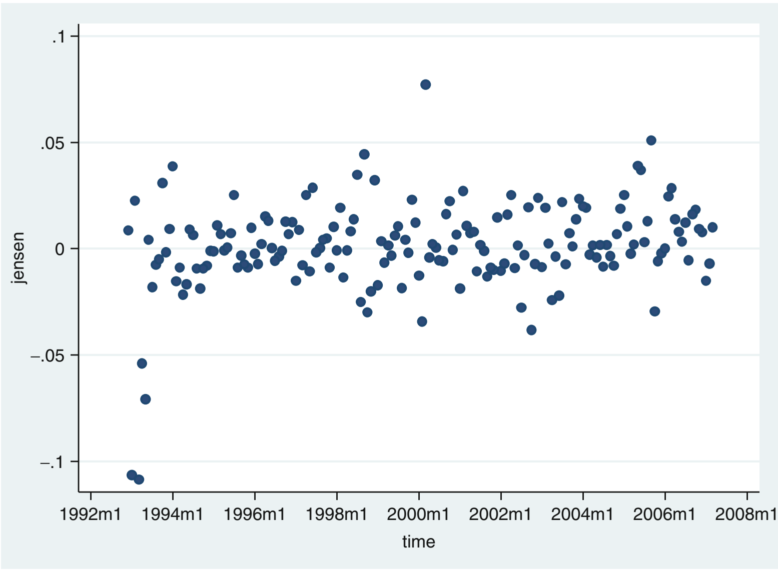
5. *Probabilities test*

Variable	N	Observed K	Expected K	Assumed p	Observed p
Jensen	172	92	86	0.50000	0.53488

Pr(k >= 92) = 0.200842 (one-sided test)
Pr(k <= 92) = 0.839213 (one-sided test)
Pr(k <= 80 or k >= 92) = 0.401684 (two-sided test)

We use this binomial test to determine whether the probability of beating the Market is similar to the probability of failing to beat it (50–50). If we look at the two-sided test, we obtain a probability of 0.401684, above 5 %, and so we accept the success rate is around 50 %, which supports the results obtained with the contingency table.

6. *Fit of Jensen by month*



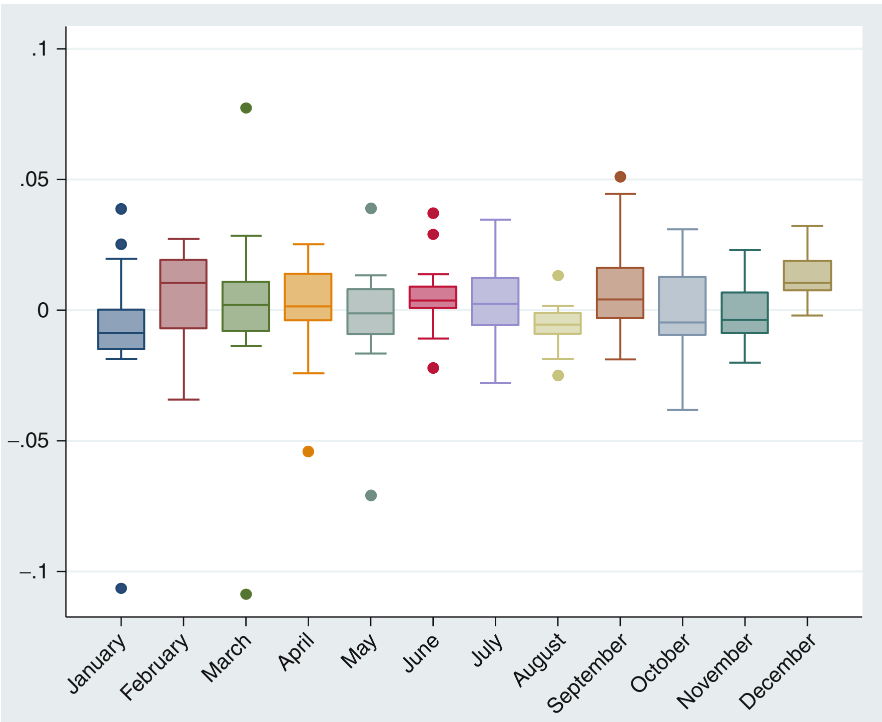
7. *Linear fit*

Source	SS	df	MS	Number of obs.	172
Model	0.00459285	1	0.00459285	F(1, 170)	10.63
Residual	0.07342831	170	0.00043193	Prob > F	0.0013
Total	0.07802116	171	0.00045626	R-squared	0.0589
				Adj. R-squared	0.0533
				Root MSE	0.02078

jensen	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
time	0.0001041	0.0000319	3.26	0.001	0.0000411	0.0001671
_cons	–0.048762	0.0154174	–3.16	0.002	–0.0791963	–0.0183277

From the above table we can observe that the regression slope which connects Jensen's index (the dependent variable) and time measured in months (the independent variable) gives a positive and significant coefficient for beta which allows us to conclude that this strategy gives results that improve over time.

8. *One-way analysis of Jensen by month*



9. *One-way ANOVA (by month)*

Source	SS	df	MS		
Model	0.00607108	11	0.00055192	Number of obs	172
Residual	0.07195009	160	0.00044969	F(11, 160)	1.23
Total	0.07802116	171	0.00045626	Prob > F	0.2729
				R-squared	0.0778
				Adj R-squared	0.0144
				Root MSE	0.02121

jensen	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
_cons	0.0084177	0.0056675	1.49	0.139	−0.0027751	0.0196104
1	−0.0094303	0.0080151	−1.18	0.241	−0.0252592	0.0063987
2	−0.0142236	0.0080151	−1.77	0.078	−0.0300526	0.0016053
3	0.0040468	0.0078803	0.51	0.608	−0.0115161	0.0196097
4	−0.0021065	0.0078803	−0.27	0.79	−0.0176694	0.0134564
5	−0.0167755	0.0078803	−2.13	0.035	−0.0323384	−0.0012126
6	−0.0049838	0.0080151	−0.62	0.535	−0.0208128	0.0108451
7	−0.0026922	0.0080151	−0.34	0.737	−0.0185211	0.0131368
8	−0.0082308	0.0078803	−1.04	0.298	−0.0237937	0.0073321
9	−0.0112643	0.0080151	−1.41	0.162	−0.0270932	0.0045647
10	−0.0092672	0.0080151	−1.16	0.249	−0.0250961	0.0065618
11	−0.0115344	0.0080151	−1.44	0.152	−0.0273633	0.0042946
12	(dropped)					

We carry out an ANOVA to determine whether the means for Jensen's indices obtained in the different months are uniform. We can see that F has a value of 1.23 and a related probability of 0.2729, above 5 %, which leads us to accept that the means for Jensen's indices are uniform throughout the months included in our study.

10. *Kruskal-Wallis tests (Rank sums), by month*

Month	Obs.	Rank Sum
January	15	949
February	15	1551
March	15	1310
April	14	1217
May	14	1093
June	14	1386
July	14	1270
August	14	798
September	14	1382
October	14	1025
November	14	1032
December	15	1865

chi-squared = 22.716 with 11 d.f.
probability = 0.0194

We use the Kruskal–Wallis nonparametric test to measure the same phenomenon as with the ANOVA, and we see that for a significance level of 1 % we would come to the same conclusion as we did with the ANOVA, while for a significance level of 5 %, we can rule out uniformity from month to month.

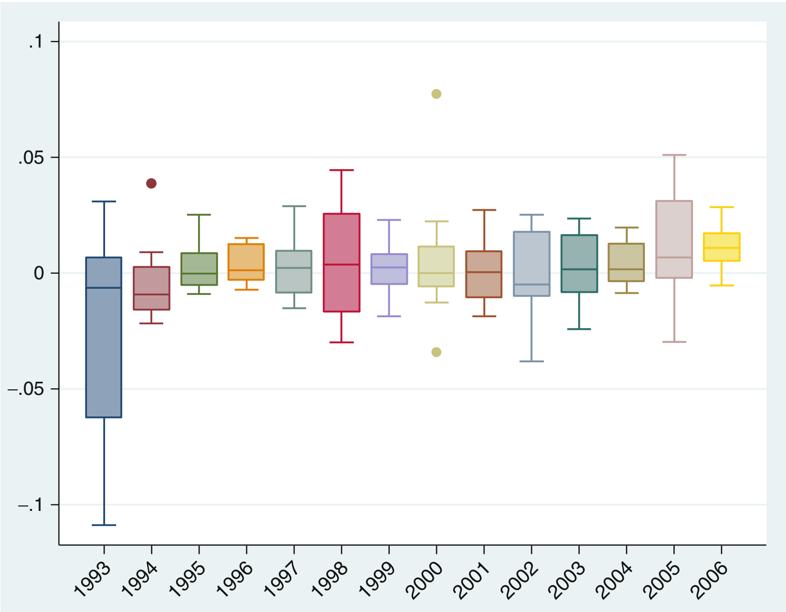
11. *Contingency table (by month)*

Month	BEAT		
	NO	YES	Total
January	11	4	15
February	5	10	15
March	7	8	15
April	6	8	14
May	7	7	14
June	2	12	14
July	6	8	14
August	11	3	14
September	6	8	14
October	8	6	14
November	9	5	14
December	2	13	15
Total	80	92	172

Pearson chi2(11) = 26.3578 Pr = 0.006
likelihood-ratio chi2(11) = 28.4294 Pr = 0.003

We use this contingency table to test the same point as with the ANOVA and the Kruskal–Wallis test above, to determine whether there are any differences between the months, and we find that both with Pearson's chi-square test and with the likelihood ratio the probability is less than 1 %, which leads us to conclude that the chances of beating the Market are not the same month on month, so that there will be some months in which it is easier to do so than others.

12. *One-way analysis of Jensen by year*



13. *One-way ANOVA (by year)*

Source	SS	df	MS		
Model	0.01216506	13	0.00093577	Number of obs	168
Residual	0.06538048	154	0.00042455	F(13, 154)	2.2
Total	0.07754554	167	0.00046435	Prob > F	0.0117
				R-squared	0.1569
				Adj R-squared	0.0857
				Root MSE	0.0206

jensen	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
_cons	0.0113553	0.005948	1.91	0.058	−0.000395	0.0231055
1	−0.0368126	0.0084118	−4.38	0	−0.0534299	−0.0201952
2	−0.0159892	0.0084118	−1.9	0.059	−0.0326065	0.0006282
3	−0.0088588	0.0084118	−1.05	0.294	−0.0254762	0.0077586
4	−0.0077928	0.0084118	−0.93	0.356	−0.0244102	0.0088246
5	−0.0082156	0.0084118	−0.98	0.33	−0.024833	0.0084018
6	−0.0062178	0.0084118	−0.74	0.461	−0.0228352	0.0103996
7	−0.0102373	0.0084118	−1.22	0.225	−0.0268546	0.0063801
8	−0.0061693	0.0084118	−0.73	0.464	−0.0227866	0.0104481
9	−0.0108019	0.0084118	−1.28	0.201	−0.0274192	0.0058155
10	−0.0127709	0.0084118	−1.52	0.131	−0.0293883	0.0038465
11	−0.009494	0.0084118	−1.13	0.261	−0.0261114	0.0071234
12	−0.0078177	0.0084118	−0.93	0.354	−0.0244351	0.0087997
13	0.0003579	0.0084118	0.04	0.966	−0.0162595	0.0169752
14	(dropped)					

We performed an ANOVA based on years, to see whether the strategy gives similar Jensen values from year to year for those years included in our study (we excluded 1992 and 2007 because they could provide only 1 and 3 months of data, respectively. These years are also excluded from the other tests done based on years). We obtained a figure F of 2.20 with a related probability of 0.0117, thus for a significance level of 1 % we would accept that the means are uniform from one year to another, however, for a significance level of 5 % we must rule out this hypothesis and conclude that the Jensen values differ from one year to another.

14. *Kruskal-Wallis tests (Rank sums), by year*

Year	Obs.	Rank Sum
1993	12	722
1994	12	661
1995	12	1017
1996	12	1092
1997	12	1024
1998	12	1083
1999	12	993
2000	12	1002
2001	12	931
2002	12	906
2003	12	1019
2004	12	1058
2005	12	1275
2006	12	1413

chi-squared = 16.527 with 13 d.f.
probability = 0.2218

We now use the Kruskal–Wallis test to find out the same point as we did with the ANOVA. We see that we obtain a probability greater than 5 % which means we can accept that the mean Jensen values are uniform from one year to another.

15. *Contingency table (by year)*

Year	BEAT		
	NO	YES	Total
1993	8	4	12
1994	9	3	12
1995	6	6	12
1996	5	7	12
1997	5	7	12
1998	6	6	12
1999	5	7	12
2000	6	6	12
2001	6	6	12
2002	7	5	12
2003	5	7	12
2004	5	7	12
2005	4	8	12
2006	1	11	12
Total	78	90	168

Pearson chi2(13) = 15.2205 Pr = 0.294
likelihood-ratio chi2(13) = 16.7608 Pr = 0.210

The contingency table confirms the result of the previous test, namely, that there are no differences between the Jensen values for the different years so that any small difference can be due to chance.

Strategy 2: The Investor Buys the Most Undervalued Stocks

Strategy 2, 1st Variant: Quartiles Constructed Using Jensen's Index

1. *Summary statistics*

	Percentiles	Smallest		
1%	−0.1414719	−0.1486773		
5%	−0.0391241	−0.1414719		
10%	−0.0192669	−0.1117398	Obs.	172
25%	−0.0103728	−0.0683232	Sum of Wgt.	172
50%	0.0017188		Mean	0.0011871
		Largest	Std. Dev.	0.0280576
75%	0.0154674	0.0545513		
90%	0.0294085	0.0690987	Variance	0.0007872
95%	0.0390722	0.0696097	Skewness	−1.842765
99%	0.0696097	0.0745625	Kurtosis	11.929

We would point out that in this table the Jensen values are far from normality, as can be seen in the readings for asymmetry and kurtosis.

2. *Mean test*

Variable	Obs.	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]	
jensen (Q)	172	0.001187	0.0021394	0.0280576	–0.0030359	0.0054101

mean = mean (jensen) t = 0.5549
Ho: mean = 0 degrees of freedom = 171

Ha: mean < 0	Ha: mean != 0	Ha: mean > 0
Pr(T < t) = 0.7101	Pr(|T| > |t|) = 0.5797	Pr(T > t) = 0.2899

From this mean test we assume that the mean for Jensen's alphas is zero, in fact, if we focus on the two-sided test we can detect a probability of 0.5797, higher than 5 %. This allows us to conclude that we are not able to beat the Market with this strategy.

3. *Nonparametric mean test*

sign	Obs.	sum ranks	expected
positive	94	8567	7439
negative	78	6311	7439
zero	0	0	0
all	172	14878	14878

unadjusted variance 427742.50
adjustment for ties 0.00
adjustment for zeros 0.00
----
adjusted variance 427742.50
Ho: jensen = 0
z = 1.725
Prob > |z| = 0.0846

We can see again, that Wilcoxon's nonparametric test leads to the same conclusion: we cannot beat the Market with this strategy.

4. *Frequencies*

BEAT	BEAT		
	NO	YES	Total
NO	78 45.35	0 0	78 45.35
YES	0 0	94 54.65	94 54.65
Total	78 45.35	94 54.65	172 100

From this contingency table, we see that the proportion of months in which we beat the Market is very similar to the proportion of months in which we do not, with a slightly higher probability of beating the Market.

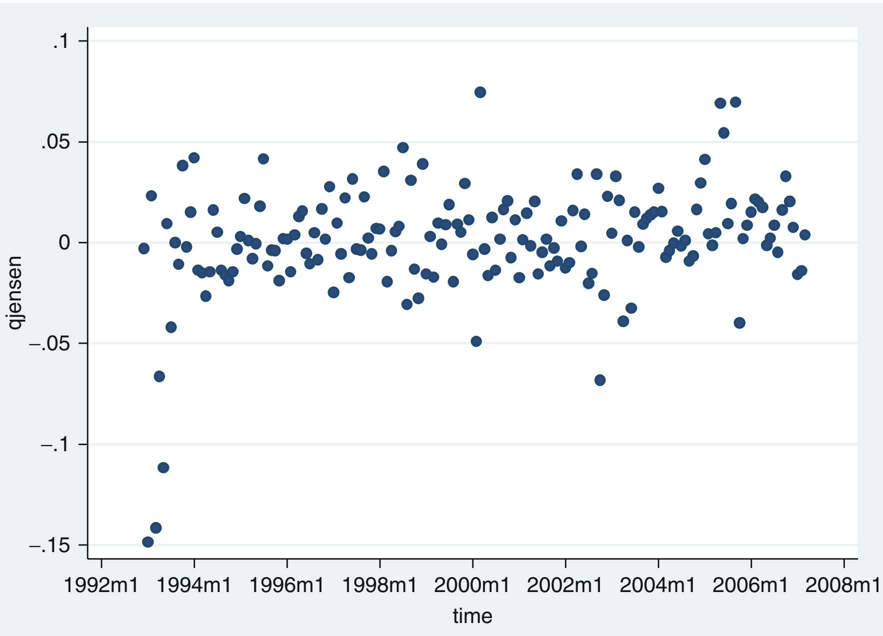
5. *Probabilities test*

Variable	N	Observed K	Expected K	Assumed p	Observed p
Jensen (Q)	172	94	86	0.50000	0.54651

Pr(k >= 94) = 0.126331 (one-sided test)
Pr(k <= 94) = 0.902626 (one-sided test)
Pr(k <= 78 or k >= 94) = 0.252662 (two-sided test)

With this binomial test we set out to find whether the probability of beating the Market is the same as the probability of not beating it; (50–50). If we look at the two-sided test, we get a probability of 0.252662, above 5 %, hence we can accept that the success rate is 50 %, which backs up the results obtained with the contingency table.

6. *Fit of Jensen (quartiles) by month*



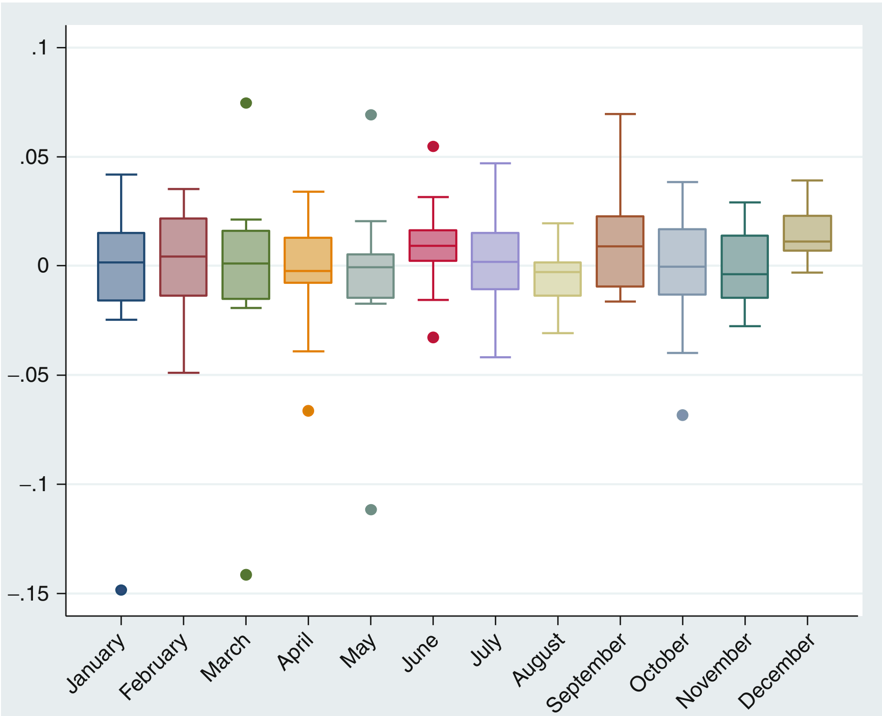
7. *Linear fit*

Source	SS	df	MS	Number of obs	172
Model	0.00979599	1	0.00979599	F(1, 170)	13.34
Residual	0.1248204	170	0.00073424	Prob > F	0.0003
Total	0.13461639	171	0.00078723	R-squared	0.0728
				Adj R-squared	0.0673
				Root MSE	0.0271

jensen (Q)	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
time	0.000152	0.0000416	3.65	0	0.0000699	0.0002341
_cons	−0.0718465	0.0201013	−3.57	0	−0.1115268	−0.0321663

From the above table we can observe that the regression slope, which connects Jensen's index (the dependent variable) and time measured in months (the independent variable), gives a positive and significant coefficient for beta which allows us to conclude that this strategy gives results that improve over time.

8. *One-way analysis of Jensen (quartiles) by month*



9. *One-way ANOVA (by month)*

Source	SS	df	MS	Number of Obs.	172
Model	0.00735542	11	0.00066868	F(11, 160)	0.84
Residual	0.12726097	160	0.00079538	Prob > F	0.5997
Total	0.13461639	171	0.00078723	R-squared	0.0546
				Adj. R-squared	–0.0104
				Root MSE	0.0282

jensen (Q)	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
_cons	0.0105061	0.0075374	1.39	0.165	–0.0043796	0.0253918
1	–0.0142298	0.0106595	–1.33	0.184	–0.0352813	0.0068218
2	–0.015749	0.0106595	–1.48	0.142	–0.0368006	0.0053025
3	0.0028747	0.0104804	0.27	0.784	–0.017823	0.0235724
4	–0.0060897	0.0104804	–0.58	0.562	–0.0267874	0.0146081
5	–0.0172233	0.0104804	–1.64	0.102	–0.037921	0.0034744
6	–0.0069927	0.0106595	–0.66	0.513	–0.0280443	0.0140588
7	–0.0014246	0.0106595	–0.13	0.894	–0.0224761	0.019627
8	–0.014007	0.0104804	–1.34	0.183	–0.0347047	0.0066907
9	–0.0143055	0.0106595	–1.34	0.181	–0.035357	0.0067461
10	–0.0125205	0.0106595	–1.17	0.242	–0.0335721	0.008531
11	–0.0123631	0.0106595	–1.16	0.248	–0.0334147	0.0086884
12	(dropped)					

We perform an ANOVA to see whether the means for Jensen values obtained for the different months are uniform. We find that F-statistic has a value of 0.84 with a related probability of 0.5997, above 5 %, which leads us to accept that the means for Jensen values are uniform among the months included in our sample.

10. *Kruskal-Wallis tests (Rank sums), by month*

Month	Obs.	Rank Sum
January	15	1167
February	15	1428
March	15	1228
April	14	1105
May	14	1077
June	14	1451
July	14	1216
August	14	885
September	14	1385
October	14	1156
November	14	1020
December	15	1760

chi-squared = 14.369 with 11 d.f.
probability = 0.2133

We now use the Kruskal–Wallis test to find out the same point as we did with the ANOVA. We come to the same conclusion as we did with the ANOVA.

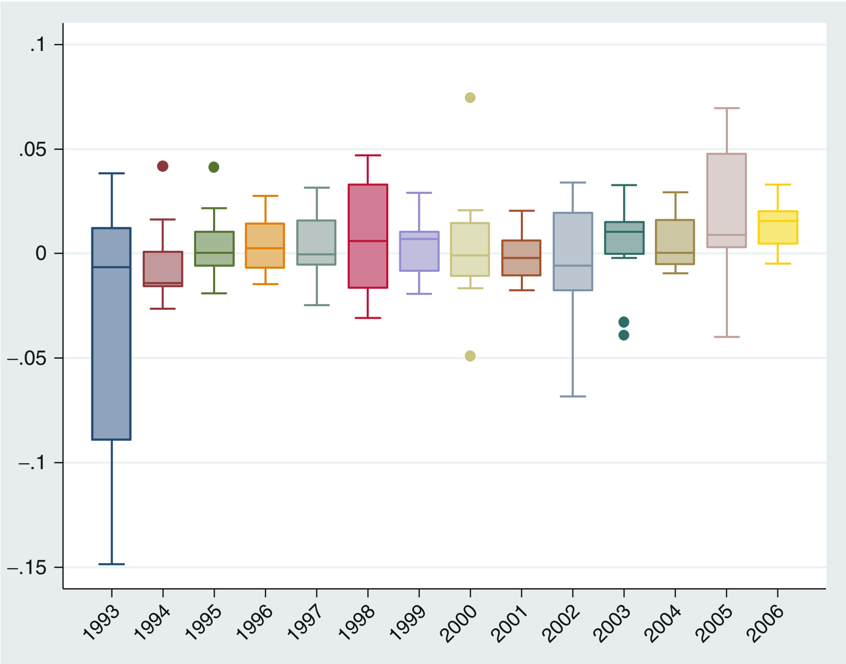
11. *Contingency table (by month)*

Month	BEAT		
	NO	YES	Total
January	7	8	15
February	5	10	15
March	7	8	15
April	8	6	14
May	9	5	14
June	3	11	14
July	7	7	14
August	9	5	14
September	6	8	14
October	7	7	14
November	8	6	14
December	2	13	15
Total	78	94	172

Pearson chi2(11) = 16.2331 Pr = 0.133
likelihood-ratio chi2(11) = 17.3939 Pr = 0.097

The contingency table allows us to test the same point as above with the ANOVA and the Kruskal–Wallis test, and leads to the same conclusions; hence, we can accept that all months show a similar tendency in terms of beating the Market.

12. *One-way analysis of Jensen (quartiles) by year*



13. *One-way ANOVA (by year)*

Source	SS	df	MS		
Model	0.02482403	13	0.00190954	Number of obs.	168
Residual	0.10923702	154	0.00070933	F(13, 154)	2.69
Total	0.13406105	167	0.00080276	Prob > F	0.0019
				R-squared	0.1852
				Adj. R-squared	0.1164
				Root MSE	0.02663

jensen (Q)	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
_cons	0.0129934	0.0076884	1.69	0.093	−0.0021949	0.0281817
1	−0.0494527	0.010873	−4.55	0	−0.0709322	−0.0279733
2	−0.0191305	0.010873	−1.76	0.08	−0.04061	0.002349
3	−0.00969	0.010873	−0.89	0.374	−0.0311695	0.0117894
4	−0.0092178	0.010873	−0.85	0.398	−0.0306973	0.0122617
5	−0.0100839	0.010873	−0.93	0.355	−0.0315634	0.0113956
6	−0.0066183	0.010873	−0.61	0.544	−0.0280978	0.0148612
7	−0.0095188	0.010873	−0.88	0.383	−0.0309983	0.0119607
8	−0.0095735	0.010873	−0.88	0.38	−0.031053	0.0119059
9	−0.0142031	0.010873	−1.31	0.193	−0.0356825	0.0072764
10	−0.0158029	0.010873	−1.45	0.148	−0.0372824	0.0056765
11	−0.0087823	0.010873	−0.81	0.421	−0.0302617	0.0126972
12	−0.0075224	0.010873	−0.69	0.49	−0.0290019	0.013957
13	0.0071267	0.010873	0.66	0.513	−0.0143528	0.0286061
14	(dropped)					

We now perform an ANOVA by years to see whether the strategy gives similar Jensen values among the years in our sample. We obtain a value for F-statistic of 2.69 with a related probability of 0.0019; hence we conclude that the means for Jensen values are not uniform from one year to another.

14. *Kruskal-Wallis tests (Rank sums), by year*

Year	Obs.	Rank Sum
1993	12	735
1994	12	665
1995	12	980
1996	12	1042
1997	12	1002
1998	12	1064
1999	12	1045
2000	12	972
2001	12	846
2002	12	913
2003	12	1136
2004	12	1070
2005	12	1353
2006	12	1373

chi-squared = 17.864 with 13 d.f.
probability = 0.1627

We now use the Kruskal–Wallis nonparametric test to find out the same point as we did with the ANOVA. We see that we obtain a probability greater than 5 % which means we can accept that the mean Jensen values are uniform from one year to another.
15. *Contingency table (by year)*

Year	BEAT		
	NO	YES	Total
1993	8	4	12
1994	9	3	12
1995	6	6	12
1996	4	8	12
1997	6	6	12
1998	5	7	12
1999	4	8	12
2000	6	6	12
2001	7	5	12
2002	7	5	12
2003	3	9	12
2004	6	6	12
2005	2	10	12
2006	2	10	12
Total	75	93	168

Pearson chi2(13) = 19.9673 Pr = 0.096
likelihood-ratio chi2(13) = 21.0731 Pr = 0.071

The contingency table confirms the result of the previous test, namely, that there are no differences between the Jensen values for the different years so that any small difference can be due to chance.

Strategy 2, 2nd Variant: Quartiles Built Using Treynor's Ratio

1. *Summary statistics*

	Percentiles	Smallest		
1%	–0.0968895	–0.104714		
5%	–0.0382262	–0.0968895		
10%	–0.0231222	–0.0940315	Obs.	172
25%	–0.0106754	–0.0702515	Sum of Wgt.	172
50%	0.0023977		Mean	0.0014829
		Largest	Std. Dev.	0.0256145
75%	0.0157561	0.0484799		
90%	0.0285186	0.0629996	Variance	0.0006561
95%	0.04065	0.0648316	Skewness	–0.9274022
99%	0.0648316	0.0745625	Kurtosis	6.503481

We note that in this table the Jensen values are far from normality, as can be seen in the readings for asymmetry and kurtosis, although this is less obvious than for the previous strategies.

2. *Mean test*

Variable	Obs.	Mean	Std. Err.	Std. Dev.	[95% Conf. Interval]
Jensen (Q)	172	0.001483	0.0019531	0.0256145	−0.0023724 0.0053381

mean = mean (jensen) t = 0.7592

Ho: mean = 0 degrees of freedom = 171

Ha: mean < 0	Ha: mean != 0	Ha: mean > 0
Pr(T < t) = 0.7756	Pr(|T| > |t|) = 0.4487	Pr(T > t) = 0.2244

From this mean test, we can accept that the mean for the Jensen alphas is zero; in fact, if we focus on the two-sided test, we obtain a probability of 0.4487, which is higher than 5 %. This allows us to conclude that we cannot beat the Market with this strategy.

3. *Nonparametric mean test*

sign	Obs.	sum ranks	expected
positive	93	8460	7439
negative	79	6418	7439
zero	0	0	0
all	172	14878	14878

unadjusted variance 427742.50
adjustment for ties 0.00
adjustment for zeros 0.00
----
adjusted variance 427742.50
Ho: jensen = 0
z = 1.561
Prob > |z| = 0.1185

We can see again that Wilcoxon's nonparametric test leads to the same conclusion: we cannot beat the Market with this strategy.

4. *Frequencies*

BEAT	BEAT		
	NO	YES	Total
NO	79 45.93	0 0	79 45.93
YES	0 0	93 54.07	93 54.07
Total	79 45.93	93 54.07	172 100

From this contingency table, we see that the proportion of months in which we beat the Market is very similar to the proportion in which we do not, with a slightly higher figure for those months in which we do beat the Market.

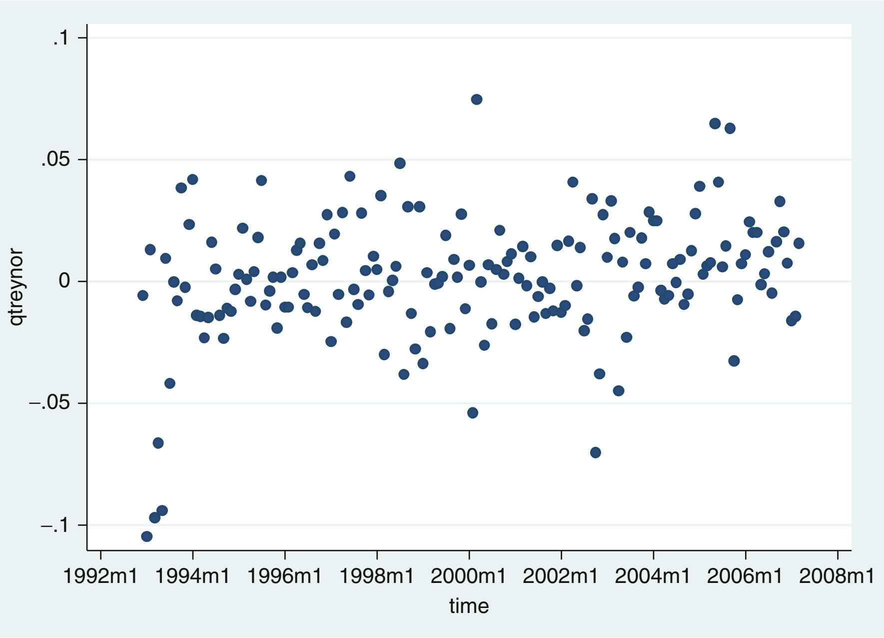
5. *Probabilities test*

Variable	N	Observed K	Expected K	Assumed p	Observed p
Jensen (Q)	172	93	86	0.50000	0.5407

Pr(k >= 93) = 0.160787 (one-sided test)
Pr(k <= 93) = 0.873669 (one-sided test)
Pr(k <= 79 or k >= 93) = 0.321574 (two-sided test)

We use this binomial test to determine whether the probability of beating the Market is similar to the probability of failing to beat it (50–50). If we look at the two-sided test, we obtain a probability of 0.321574, above 5 %, and so we accept the success rate is around 50 %, which supports the results obtained with the contingency table.

6. *Fit of Jensen (quartiles) by month*



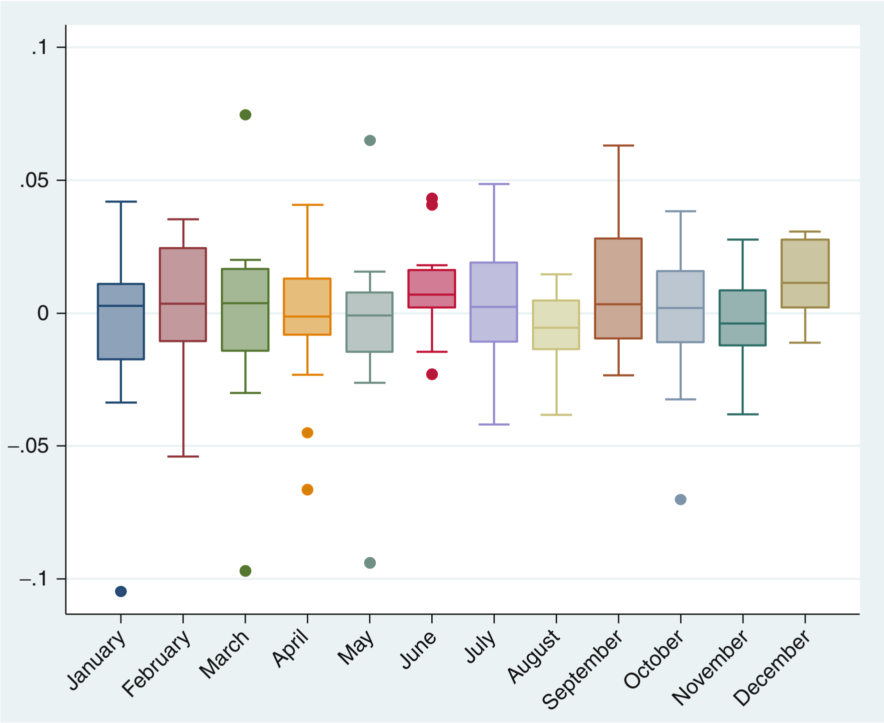
7. *Linear fit*

Source	SS	df	MS	Number of obs.	172
Model	0.00721202	1	0.00721202	F(1, 170)	11.68
Residual	0.1049811	170	0.00061754	Prob > F	0.0008
Total	0.11219312	171	0.0006561	R-squared	0.0643
				Adj R-squared	0.0588
				Root MSE	0.02485

jensen (Q)	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
time	0.0001304	0.0000382	3.42	0.001	0.0000551	0.0002058
_cons	−0.0611824	0.0184347	−3.32	0.001	−0.0975728	−0.024792

From the above table we can observe that the regression slope which connects Jensen's index (the dependent variable) and time measured in months (the independent variable) gives a positive and significant coefficient for beta which allows us to conclude that this strategy gives results that improve over time.

8. *One-way analysis of Jensen (quartiles) by month*



9. *One-way ANOVA (by month)*

Source	SS	df	MS		
Model	0.00656026	11	0.00059639	Number of Obs.	172
Residual	0.10563286	160	0.00066021	F(11, 160)	0.9
Total	0.11219312	171	0.0006561	Prob > F	0.5387
				R-squared	0.0585
				Adj. R-squared	–0.0063
				Root MSE	0.02569

jensen (Q)	Coef.	Std. Err.	t	P>|t|	[95% Conf.	Interval]
_cons	0.0092267	0.0068671	1.34	0.181	–0.0043352	0.0227886
1	–0.0125889	0.0097116	–1.3	0.197	–0.0317683	0.0065906
2	–0.0150679	0.0097116	–1.55	0.123	–0.0342474	0.0041115
3	0.003986	0.0095484	0.42	0.677	–0.0148711	0.0228431
4	–0.0040562	0.0095484	–0.42	0.672	–0.0229133	0.0148009
5	–0.0145235	0.0095484	–1.52	0.13	–0.0333805	0.0043336
6	–0.0055123	0.0097116	–0.57	0.571	–0.0246917	0.0136671
7	–0.0003726	0.0097116	–0.04	0.969	–0.0195521	0.0188068
8	–0.0092886	0.0095484	–0.97	0.332	–0.0281457	0.0095684
9	–0.013379	0.0097116	–1.38	0.17	–0.0325584	0.0058005
10	–0.0120712	0.0097116	–1.24	0.216	–0.0312506	0.0071083
11	–0.0105584	0.0097116	–1.09	0.279	–0.0297378	0.008621
12	(dropped)					

We perform an ANOVA to see whether the means for the Jensen values obtained for the different months are uniform. We find that F-statistic has a value of 0.90 with a related probability of 0.5387, above 5 %, which leads us to accept that the means for the Jensen values are uniform among the months included in our sample.

10. *Kruskal-Wallis tests (Rank sums), by month*

Month	Obs.	Rank Sum
January	15	1153
February	15	1442
March	15	1297
April	14	1107
May	14	1046
June	14	1430
July	14	1220
August	14	899
September	14	1352
October	14	1160
November	14	1044
December	15	1728

chi-squared = 12.840 with 11 d.f.
probability = 0.3039

We now use the Kruskal–Wallis nonparametric test to find out the same point as we did with the ANOVA. We come to the same conclusion as we did with the ANOVA.

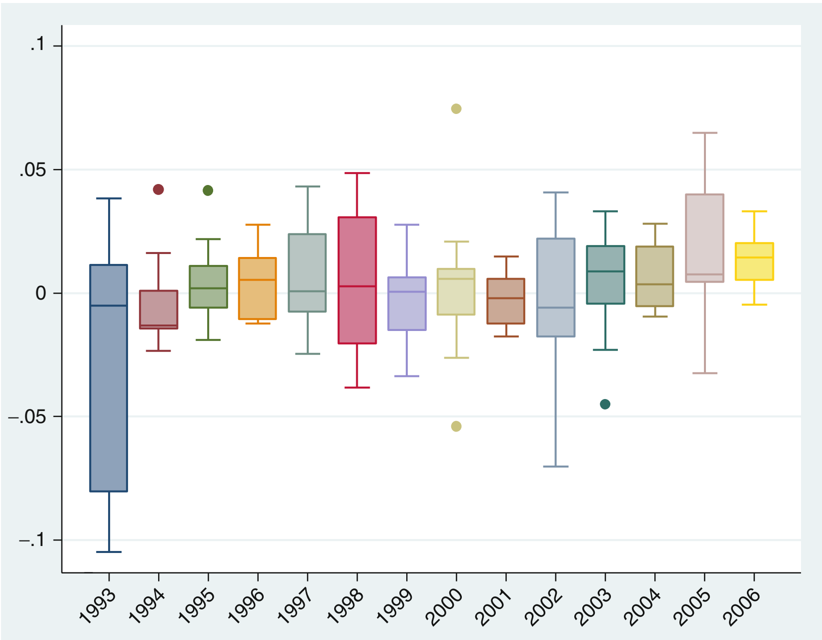
11. *Contingency table (by month)*

Month	BEAT		
	NO	YES	Total
January	7	8	15
February	5	10	15
March	6	9	15
April	9	5	14
May	8	6	14
June	3	11	14
July	7	7	14
August	10	4	14
September	7	7	14
October	6	8	14
November	8	6	14
December	3	12	15
Total	79	93	172

Pearson chi2 (11) = 15.8416 Pr=0.147
likelihood-ratio chi2(11) = 16.5468 Pr = 0.122

The contingency table allows us to test the same point as above with the ANOVA and the Kruskal–Wallis test, and leads to the same conclusions; hence we can accept that all months show a similar tendency in terms of beating the Market.

12. *One-way analysis of Jensen (quartiles) by year*



13. *One-way ANOVA (by year)*

Source	SS	df	MS		
Model	0.0170417	13	0.0013109	Number of obs.	168
Residual	0.09434205	154	0.00061261	F(13, 154)	2.14
Total	0.11138374	167	0.00066697	Prob > F	0.0147
				R-squared	0.153
				Adj. R-squared	0.0815
				Root MSE	0.02475

jensen	Coef.	Std. Err.	t	P>|t|	[95% Conf. Interval]	
_cons	0.0135106	0.007145	1.89	0.061	−0.0006043	0.0276254
1	−0.041029	0.0101045	−4.06	0	−0.0609904	−0.0210676
2	−0.0190657	0.0101045	−1.89	0.061	−0.0390271	0.0008957
3	−0.0091611	0.0101045	−0.91	0.366	−0.0291225	0.0108003
4	−0.0100757	0.0101045	−1	0.32	−0.0300372	0.0098857
5	−0.0077821	0.0101045	−0.77	0.442	−0.0277435	0.0121793
6	−0.0098938	0.0101045	−0.98	0.329	−0.0298552	0.0100676
7	−0.0154658	0.0101045	−1.53	0.128	−0.0354272	0.0044956
8	−0.0103002	0.0101045	−1.02	0.31	−0.0302616	0.0096612
9	−0.0158053	0.0101045	−1.56	0.12	−0.0357667	0.0041561
10	−0.0165104	0.0101045	−1.63	0.104	−0.0364718	0.003451
11	−0.0080536	0.0101045	−0.8	0.427	−0.028015	0.0119078
12	−0.0072511	0.0101045	−0.72	0.474	−0.0272125	0.0127103
13	0.004184	0.0101045	0.41	0.679	−0.0157774	0.0241454
14	(dropped)					

We now perform an ANOVA by years to see whether the strategy gives similar Jensen values among the years in our sample. We obtain a value for F-statistic of 2.14 with a related probability of 0.0147; hence we conclude that for a significance level of 1 % we can accept that the means for the Jensen values are the same from one year to another, however for a significance level of 5 % we must reject this hypothesis and we would conclude that there are differences between the Jensen values of different years.

14. *Kruskal-Wallis tests (Rank sums), by year*

Year	Obs.	Rank Sum
1993	12	740
1994	12	679
1995	12	1025
1996	12	1033
1997	12	1065
1998	12	1022
1999	12	877
2000	12	1026
2001	12	824
2002	12	920
2003	12	1166
2004	12	1110
2005	12	1322
2006	12	1387

chi-squared = 18.336 with 13 d.f.
probability = 0.1452

We now use the Kruskal–Wallis test to find out the same point as we did with the ANOVA. We see that we obtain a probability greater than 5 % which means we can accept that the means for the Jensen values are uniform from one year to another.

15. *Contingency table (by year)*

Year	BEAT		
	NO	YES	Total
1993	8	4	12
1994	9	3	12
1995	4	8	12
1996	5	7	12
1997	6	6	12
1998	5	7	12
1999	6	6	12
2000	4	8	12
2001	8	4	12
2002	7	5	12
2003	4	8	12
2004	6	6	12
2005	2	10	12
2006	2	10	12
Total	76	92	168

Pearson chi2 (13) = 19.9908 Pr = 0.095
likelihood-ratio chi2(13) = 21.0581 Pr = 0.072

The contingency table confirms the result of the previous test, namely, that there are no differences between the Jensen values for the different years and that any small difference can be due to chance.

References

Abdel-Kader, M. G., & Qing, K. Y. (2007). Risk-adjusted performance, selectivity, timing ability and performance persistence of Hong Kong mutual funds. *Journal of Asia-Pacific Business, 8*(2), 25–58.

Agresti, A. (2007). *An introduction to categorical data analysis* (2nd ed.). Hoboken: Wiley Interscience.

Anderson, D. R., Sweeney, D. J., & Williams, T. A. (2011). Statistics for business and economics (11th ed.). Mason: South-Western, Cengage Learning.

Black, F., Jensen, M. C., & Scholes, M. (1972). The capital asset pricing model: Some empirical tests. In Jensen, M. C. (Ed.), *Studies in the theory of capital Markets* (pp. 79–121). New York: Praeger.

Brav, A., Lehavy, R., & Michaely, R. (2005). Using expectations to test asset pricing models. *Financial Management, 34*(3), 31–64.

Brealey, R. A., Myers, S. C., & Allen, F. (2008). *Principles of corporate finance* (9th ed.). New York: McGraw-Hill.

Conover, W. J. (1999). *Practical nonparametric statistics* (3rd ed.). New York: Wiley.

Copeland, T. E., Weston, J. F., & Shastri, K. (2005). *Financial theory and corporate policy* (4th ed.). Boston: Pearson Addison Wesley.

Danielson, M. G., Heck, J. L., & Shaffer, D. R. (2008). Shareholder theory – How opponents and proponents both get it wrong. *Journal of Applied Finance, 18*(2), 62–66.

Dimson, E., & Mussavian, M. (1998). A brief history of Market efficiency. *European Financial Management, 4*(1), 91–103.

Dimson, E., & Mussavian, M. (1999). Three centuries of asset pricing. *Journal of Banking and Finance, 23*(12), 1745–1769.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance, 25*(2), 383–417.

Fama, E. F. (1991). Efficient capital markets: II. *Journal of Finance, 46*, 1575–1617.

Fama, E. F. (1998). Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics*, *49*, 283–306.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance, 47*, 427–465.

Fama, E. F., & French, K. R. (2008). Dissecting anomalies. *Journal of Finance, 63*(4), 1653–1678.

Fama, E. F., & French, K. R. (2010). Luck versus skill in the cross-section of mutual fund returns. *Journal of Finance, 65*(5), 1915–1947.

Fama, E. F., & MacBeth, J. D. (1973). *Risk, return and equilibrium: empirical tests* (pp. 607–636). May-June: Journal of Political Economy.

Ferruz, L., Gómez-Bezares, F., & Vargas, M. (2010). Portfolio theory, CAPM, and performance measures. In Lee, C. F. et al. (Eds.), *Handbook of quantitative finance and risk management* (pp. 267–283). New York: Springer.

Gómez-Bezares, F., & Gómez-Bezares, F. R. (2006). On the use of hypothesis testing and other problems in financial research. *The Journal of Investing, 15*(4), 64–67.

Gómez-Bezares, F., Madariaga, J. A., & Santibáñez, J. (1996). Valuation and efficiency models: Does the CAPM beat the Market? *Análisis Financiero, 68*, 72–96.

Jensen, M. C. (1968). *The performance of mutual funds in the period 1945–1964* (pp. 389–416). May: Journal of Finance.

Jensen, M. C. (1969). Risk, the pricing of capital assets, and the evaluation of investment portfolios. *Journal of Business, 42*, 167–247.

Levy, H. (2010). The CAPM is alive and well: A review and synthesis. *European Financial Management, 16*(1), 43–71.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics, 47*, 13–37.

Liu, W. (2006). A liquidity-augmented capital asset pricing model. *Journal of Financial Economics, 82*, 631–671.

Malkiel, B. G. (2003). The efficient Market hypothesis and its critics. *Journal of Economic Perspectives, 17*(1), 59–82.

Mazumder, M. I., & Varela, O. (2010). Market timing, the trading of international mutual funds: weekend, weekday and serial correlation strategies. *Journal of Business, Finance & Accounting, 37*(7–8), 979–1007.

Modigliani, L. (1997). Don't pick your managers by their alphas. *U.S. Investment Research, U.S. Strategy*, Morgan Stanley Dean Witter, June.

Newbold, P., Carlson, W. L., & Thorne, B. M. (2009). *Statistics for business and economics* (7th ed.). New Jersey: Pearson Education Canada.

Pasaribu, R. B. (2009). Relevance of accrual anomaly information in stock portfolio formation – case study: Indonesia stock exchange. *Journal of Accounting and Business, 10*(2), 1–45.

Samarakoon, L. P., & Hasan, T. (2005). Portfolio performance evaluation. In Lee, C. F. (Ed.), *The encyclopedia of finance* (pp. 617–622). New York: Springer.

Sharpe, W. F. (1964). Capital asset prices: A theory of Market equilibrium under conditions of risk. *Journal of Finance, 19*, 425–442.

Treynor, J. L. (1965). How to rate management of investment funds. *Harvard Business Review, 43*, 63–75.

Group Decision-Making Tools for Managerial Accounting and Finance Applications

29

Wikil Kwak, Yong Shi, Cheng-Few Lee, and Heeseok Lee

Contents

W. Kwak (✉)
University of Nebraska at Omaha, Omaha, NE, USA
e-mail: wkwak@unomaha.edu

Y. Shi
University of Nebraska at Omaha, Omaha, NE, USA

Chinese Academy of Sciences, Beijing, China
e-mail: yshi@unomaha.edu

C.-F. Lee
Department of Finance and Economics, Rutgers Business School, Rutgers, The State University of New Jersey, Piscataway, NJ, USA

Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
e-mail: cflee@business.rutgers.edu

H. Lee
Korea Advanced Institute of Science and Technology, Yuseong-gu, Daejeon, South Korea
e-mail: hsl@business.kaist.ac.kr

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_29,

Abstract

To deal with today's uncertain and dynamic business environments with different background of decision makers in computing trade-offs among multiple organizational goals, our series of papers adopts an analytic hierarchy process (AHP) approach to solve various accounting or finance problems such as developing a business performance evaluation system and developing a banking performance evaluation system. AHP uses hierarchical schema to incorporate nonfinancial and external performance measures. Our model has a broader set of measures that can examine external and nonfinancial performance as well as internal and financial performance. While AHP is one of the most popular multiple goals decision-making tools, multiple-criteria and multiple-constraint (MC^2) linear programming approach also can be used to solve group decision-making problems such as transfer pricing and capital budgeting problems. This model is rooted by two facts. First, from the linear system structure's point of view, the criteria and constraints may be "interchangeable." Thus, like multiple criteria, multiple-constraint (resource availability) levels can be considered. Second, from the application's point of view, it is more realistic to consider multiple resource availability levels (discrete right-hand sides) than a single resource availability level in isolation. The philosophy behind this perspective is that the availability of resources can fluctuate depending on the decision situation forces, such as the desirability levels believed by the different managers. A solution procedure is provided to show step-by-step procedure to get possible solutions that can reach the best compromise value for the multiple goals and multiple-constraint levels.

Keywords

Analytic hierarchy process • Multiple-criteria and multiple-constraint linear programming • Business performance evaluation • Activity-based costing system • Group decision making • Optimal trade-offs • Balanced scorecard • Transfer pricing • Capital budgeting

29.1 Introduction

In this chapter, we provide an up-to-date review on our past works in AHP and MC^2 linear programming models to solve real-world problems faced by managers in accounting and finance. These applications include developing a business performance evaluation system using an analytic hierarchical model, a banking performance evaluation system, capital budgeting, and transfer pricing problems. A good performance measurement system should incorporate strategic success factors, especially to be successful in today's competitive environment. Balanced scorecard is a hot topic, but it lacks linkages among different basic units of financial and nonfinancial measures or across different levels of managers. The model proposed in this study uses a three-level hierarchical schema to combine financial and nonfinancial performance measures systematically. Its emphasis is on an external as well as an internal business performance measures such as the balanced scorecard method. This method is more likely to cover a broader set of measures that include operational control as well as strategic control.

The purpose of this chapter is to provide additional insight for managers who face group decision-making problems in accounting and finance and want to find practical solutions.

29.2 Designing a Comprehensive Performance Evaluation System: Using the Analytic Hierarchy Process

AHP is one of the most popular multiple goals decision-making tools (Ishizaka et al. 2011). Designing a comprehensive performance measurement system has frustrated many managers (Eccles 1991). The traditional performance measures enterprises have used may not well fit in for the new business environment and competitive realities. The figures that enterprises have traditionally used are not very useful for the information-based society we are becoming. We suspect that firms are much more productive than these out-of-date measures.

A broad range of firms is deeply engaged in redefining how to evaluate the performance of their businesses. New measurements for quantification are needed to perform business evaluation. Drucker (1993) put the ever-increasing measurement dilemma this way:

> Quantification has been the rage in business and economics these past 50 years. Accountants have proliferated as fast as lawyers. Yet we do not have the measurements we need. Neither our concepts nor our tools are adequate for the control of operations, or for managerial control. And, so far, there are neither the concepts nor the tools for business control, i.e., for economic decision making. In the past few years, however, we have become increasingly aware of the need for such measurements.

Drucker's message is clear: a traditional measure is not adequate for business evaluation. A primary reason why traditional measures fail to meet new business needs is that most measures are lagging indicators (Eccles and Pyburn 1992).

The emphasis of accounting measures has been on historical statements of financial performance. They are the result of management performance, not the cause of it; i.e., they are better at measuring the consequence of yesterday's decisions but unlikely to provide useful indicators for future success. As a result, they easily conflict with new strategies and current competitive business realities.

To ameliorate this accounting lag situation, researchers have frequently attempted to provide new measuring procedures (Kaplan 1986; Wu et al. 2011). Yet most measuring guidelines may not be well represented analytically. Managers keep asking: What are the most important measures of performance? What are the associations among those measures? Unfortunately, we know little about how measures are integrated into a performance measurement system regarding a particular business. For example, is the customer service perceived as a more important measure than cost of quality? Another question often raised is: What is the association between the customer service and on-time delivery or manufacturing cycle time?

The current wave of dissatisfaction with traditional accounting systems has been intensified partly because most measures have internal and financial focus. The new measure should broaden the basis of nonfinancial performance measurement. Measures must truly predict long-term strategic success. External performance relative to competitors such as market share is as of importance as internal measures. In addition, the recent rise of global competitiveness reemphasizes the primacy of operational, i.e., nonfinancial, performance over financially oriented performance. Nonfinancial measures reflect the actionable steps needed for surviving in today's competitive environment (Fisher 1992).

When a company uses an activity-based costing system (Campi 1992; Ayvaz and Pehlivanli 2011) or just-in-time manufacturing system, using nonfinancial measures is inevitable. Nonfinancial measures reduce communication gap between workers and managers; i.e., workers can better understand what they are measured by, and managers can get timely feedback and link them to strategic decision making.

The answer proposed in this study is to use hierarchical schema to incorporate nonfinancial and external performance measures. The model has a broader set of measures that can examine external and nonfinancial performance as well as internal and financial performance. On the basis of the schema, this chapter demonstrates how Saaty's analytic hierarchy process (e.g., see Saaty (1980) and Harker and Vargas (1987)) can be merged with the performance measurement. The analytic hierarchy process is a theory of measurement that has been widely applied in modeling human judgment process. In this sense, the performance measuring method proposed in this study is referred to as the Analytic Hierarchical Performance Model (AHPM).

While the AHP has been applied in a number of cases of capital budgeting, auditing, preference analysis, and balanced scorecard to product planning, enterprise risk management, and internal control structure study (see Arrington et al. 1984; Boucher and MacStravic 1991; Liberatore et al. 1992; Hardy and Reeve 2000; Huang et al. 2011; Li et al. 2011), little attention is devoted to the problem of an analytical and comprehensive business performance model to cover a broader base of measures in the currently changing environment in accounting information systems. Although Chan and Lynn (1991) originally investigated the application of the AHP in business performance evaluation, the problem of structuring the decision hierarchy in an appropriate manner is yet to be explored. The methodology proposed in this chapter will resolve this issue.

29.2.1 Hierarchical Schema for Performance Measurement

The AHP is a theory of measurement that has been extensively applied in modeling the human judgment process (e.g., Lee 1993; Muralidhar et al. 1990). It decomposes a complex decision operation into a multilevel hierarchical structure. The primary advantage of the AHP is its simplicity and the availability of the software.

Several desirable features of the AHP can help resolve issues in performance evaluation. For example, nonfinancial and external effects can be identified and integrated with financial and internal aspects of business performance through the AHP. Furthermore, the AHP is a participation-oriented methodology that can aid coordination and synthesis of multiple evaluators in the organizational hierarchy. Participation makes a positive contribution to the quality of the performance evaluation process. This point is further explored within the context of hierarchical schema of performance evaluation as follows.

Performance measures have the relationship with management levels. They need to be filtered at each superior/subordinate level in an organization; i.e., measures do not need to be the same across management levels. Performance measures at each level, however, should be linked to performance measures at the next level up. Performance measurement information is tailored to match the responsibility of each management level. For example, at the highest level, the CEO has responsibility for performance of the total business. In contrast, the production manager's main interest may be in cost control because he or she is responsible for this. Taking this idea into account leads to the notion that the performance measuring process consists of different levels according to management levels.

Depending on organizational levels, therefore, a three management level model will be suggested: top, middle, and operational. To recognize the ever-increasing importance of nonfinancial measures, at the top management level, the hierarchy consists of two criteria: nonfinancial and financial performance. One level lower (middle-level management) includes performance criteria such as market share, customer satisfaction, productivity, ROI, and profitability.

The lowest level (operational management level) includes measures that lead to simplifications in the manufacturing process as related to high-level performance measures. Examples are quality, delivery, cycle time, inventory turnover, asset turnover, and cost. Typically, the criteria at the lowest level may have several sub-criteria. For example, quality measures may have four sub-criteria: voluntary, (appraisal and prevention,) and failure, (internal and external,) costs. With relation to these criteria, accounting information is controlled typically with respect to each product (or service) or division (or department). As a result, product and division levels are added. One should keep in mind that the above three-level hierarchical schema is dynamic over time. As a company evolves, the hierarchy must be accordingly adjusted.

Another interesting aspect is that the hierarchy is not company invariant. The hierarchy must be adjusted depending on unique situation faced by each individual company or division. Basically, the hierarchal schema is designed to accommodate any number of levels and alternatives. New level or alternative can be easily added or deleted to the hierarchy once introduced. For example, another level regarding products may be added at the bottom in order to evaluate performance of different products.

29.2.2 Analytic Hierarchical Performance Model

Based on the hierarchical structure, the performance indices at each level of the AHPM are derived by the analytic hierarchy process. The AHPM collects input judgment in the form of matrix by pairwise comparisons of criteria. An eigenvalue method is then used to scale weights of criteria at each level; i.e., the relative importance of each criterion at each level is obtained. The relative importance is defined as a performance index with respect to each alternative (i.e., criterion, product, or division).

From now on, each step of the AHPM in obtaining the weights is explored. First, a set of useful strategic criteria must be identified. Let n_t be the total number of criteria under consideration at the top management level. Typically, $n_t = 2$ (nonfinancial and financial measures). The relative weight of each criteria may be evaluated by pairwise comparison; i.e., two criteria are compared at one time until all combinations of comparison are considered (only one pairwise comparison is needed if $n_t = 2$). The experience of many users of this method and the experiments reported (Saaty 1980) are likely to support that the 1–9 scale for pairwise comparisons captures human judgment fairly well while the scale can be altered to suit each application. The result from all of pairwise comparisons is stored in an input matrix:

$$A_t = \left(\alpha_{tij}\right) \text{ (an } n_t \text{ by } n_t \text{ matrix).}$$

The element α_{tij} states the importance of alternative i compared to alternative j. For instance, if at $\alpha_{t12} = 2$, then criterion 1 is twice as important as criterion 2. Applying an eigenvalue method to A_t results in a vector $W_t = (w_{ti})$ that has n_t elements.

In addition to the vector, the inconsistency ratio (γ) is obtained to estimate the degree of inconsistency in pairwise comparisons. The common guideline is that if the ratio surpasses 0.1, a new input matrix must be generated. Generally speaking, each element of the vector resulting from an eigenvalue method is the estimated relative weight of the corresponding criterion of one level with respect to one level higher; i.e., the element w_{ti} is the relative weight of the ith criteria at this level.

At the second level of hierarchy, consider the ith criterion of the top management level. Then, we have one input matrix of pairwise comparisons of criteria (at middle management level) that corresponds to the ith criterion of the top management level. The result is stored in an n_m by n_m matrix A^i_m. Here, n_m is the total number of criteria at this level. Applying an eigenvalue method to A^i_m results in the relative weight for each criterion with relation to the i^{th} criteria at one level higher. This "local relative weight" is stored in a local weighing vector W^i_m. We need n_t local weighing vectors at this level. The "global relative weight" at this level is computed and then stored in a vector $W_m = (w_{mi})$ that has the n_m elements as

$$W_m = W_t \times \left[W_m^1 \ldots W_m^{nt}\right].$$

The similar computing process continues at the operational management level until we have all global relative weights as W_o.

A prototype for the AHPM with three-level hierarchies was built via commercial software for the analytic hierarchy process, called *Expert Choice* (Forman et al. 1985).

Above relative weights can be utilized in a number of ways for performance measurement. Clearly, it implies the relative importance among criteria at each level. For example, consider an automobile company where market share and ROI are two important criteria in performance evaluation. If the AHPM generates their weights as 0.75 and 0.25, respectively, it is reasonable to conclude that market share affects the company's performance three times higher than ROI. The weights can be used as a measure for allocating future resources in products or divisions. Assume the automobile company produces two types of autos, sedan and minivan. If the AHPM generates global relative weights as 0.8 and 0.2, respectively, i.e., the performance of sedans is four times higher than minivans, then this may provide a good reason for the CEO to invest in sedans four times higher than minivans.

One important performance control measure is the rate of performance change that can be computed at any level. The change of performance can be measured. This measure is useful in estimating the elasticity of the performance of any alternative. This elasticity can aid in resource allocation decisions; i.e., further resources may be assigned to more elastic products of divisions.

For the example of computing this elasticity, at the middle management level, the rate is then

$$\varepsilon_m = \sum\nolimits_{i=1}^{nm} W_{mi} \Delta C_{mi} / C_{mi}.$$

Here, ΔC_{mi} is the amount of change of the ith criteria. For example, if the ROI is increased from 5 % to 7 % and the market share is changed from 20 % to 25 % in the automobile company discussed above,

$$\varepsilon_m = 0.25 \times \frac{2}{5} + 0.75 \times \frac{5}{20} = 0.2875.$$

We may conclude that the overall business performance has been increased by \$28.75 % with respect to middle management level. Similarly, the performance change rates at any level can be obtained. Typically, they vary depending on levels of the AHPM; i.e., they have different implications.

29.2.3 An Example

The suitability of the AHPM is illustrated with a study on a hypothetical automobile company. Nonfinancial criteria include market share, customer satisfaction, and productivity. Financial criteria include ROI and profitability. At the operational level, six criteria such as quality, delivery, cycle time, inventory turnover, asset turnover, and cost are considered.

First, nonfinancial and financial criteria are compared and stored in a vector:

$$\mathrm{W_t} = (0.4, 0.6).$$

Next, the middle management level is considered. The relative weights of middle-level performance criteria with relation to each top-level criterion are to be computed. First, the local relative weights were computed.

For the nonfinancial criteria,

$$\mathrm{A_m^1} = \begin{bmatrix} 1 & 2 & 4 \\ 1/2 & 1 & 3 \\ 1/4 & 1/3 & 1 \end{bmatrix}$$

For the sake of convenience, ROI and profitability are not listed in this comparison matrix.

The lower triangle is not listed because, in the eigenvalue method, the lower triangle is simply the reciprocity of the upper triangle. As a result,

$$W_m^1 = (0.558,\ \ 0.320,\ \ 0.122).$$

For each weighing computing, an inconsistency ratio was computed and checked for acceptance; i.e., in this case, the ratio ($\gamma = 0.017$) was accepted because $\gamma \leq 0.1$.

For the financial criteria, ROI is estimated to be twice more important than profitability, i.e.,

$$W_m^2 = (0.667,\ \ 0.333).$$

Accordingly, the global relative weights of the managerial criteria (in the order of market share, customer satisfaction, productivity, ROI, and profitability) are then

$$\begin{aligned} \mathrm{W_m} &= (0.4,\ \ 0.6) \times \begin{bmatrix} 0.558 & 0.320 & 0.122 & 0 & 0 \\ 0 & 0 & 0 & 0.667 & 0.333 \end{bmatrix} \\ &= (0.223,\ \ 0.128,\ \ 0.049,\ \ 0.400,\ \ 0.200). \end{aligned}$$

Here, the global relative weights of market share, customer satisfaction, productivity, ROI, and profitability are 22.3 %, 12.8 %, 4.9 %, 40 %, and 20 %, respectively. Note that these percentages are elements of the above $\mathrm{W_m}$.

Let us move down to the operational level. The following are local relative weights of operational level criteria (in the order of quality, delivery, cycle time, cost, inventory turnover, and asset turnover).

For the market share,

$$\mathrm{A}_0^1 = \begin{bmatrix} 1 & 2 & 3 & 3 & 5 & 7 \\ 1/2 & 1 & 2 & 3 & 4 & 5 \\ 1/3 & 1/2 & 1 & 2 & 4 & 5 \\ 1/3 & 1/3 & 1/2 & 1 & 2 & 3 \\ 1/5 & 1/4 & 1/4 & 1/2 & 1 & 2 \\ 1/7 & 1/5 & 1/5 & 1/3 & 1/2 & 1 \end{bmatrix}$$

For convenience, the local weights are arranged in the order of quality, cycle time, delivery, cost, inventory turnover, and asset turnover:

$$W_o^1 = (0.372,\ 0.104,\ 0.061,\ 0.250,\ 0.174,\ 0.039).$$

Similarly,

$$W_o^2 = (0.423,\ 0.038,\ 0.270,\ 0.185,\ 0.055,\ 0.029);$$

$$W_o^3 = (0.370,\ 0.164,\ 0.106,\ 0.260,\ 0.064,\ 0.037);$$

$$W_o^4 = (0.220,\ 0.060,\ 0.030,\ 0.4150,\ 0.175,\ 0.101);$$

$$W_o^5 = (0.246,\ 0.072,\ 0.065,\ 0.334,\ 0.160,\ 0.122).$$

Consequently, we get the global relative weights of operational criteria as

$$\begin{aligned} \mathrm{W}_o &= (0.223, 0.128, 0.049, 0.400, 0.200) \\ &\quad \times \begin{bmatrix} 0.372 & 0.104 & 0.061 & 0.250 & 0.174 & 0.039 \\ 0.423 & 0.038 & 0.270 & 0.185 & 0.055 & 0.029 \\ 0.370 & 0.164 & 0.106 & 0.260 & 0.064 & 0.037 \\ 0.220 & 0.060 & 0.030 & 0.414 & 0.175 & 0.101 \\ 0.246 & 0.072 & 0.065 & 0.334 & 0.160 & 0.122 \end{bmatrix} \\ &= (0.292, 0.074, 0.078, 0.324, 0.151, 0.079). \end{aligned}$$

Finally, the relative importance of operational level performance measures are 29.2 %, 7.4 %, 7.8 %, 32.4 %, 15.1 %, and 7.9 %, respectively. One should note that nonfinancial measures are integrated with financial measures in the scaling process.

Our automobile company can adopt these performance measures for further analysis. The measures serve as the basis for the rate of performance change. In addition, they can be further used for evaluating the performance of each product if the product level is connected to the operational management level in the AHPM.

29.2.4 Discussions

It is too optimistic to argue that there can be one right way of measuring performance. There are two factors to be considered. First, each business organization requires its own unique set of measures depending on its environment. What is most effective for a company depends upon its history, culture, and management style (Eccles 1991). To be used for any business performance measurement, a well-designed model must be flexible enough to incorporate a variety of measures while retaining major aspects. Second, managers should change measures over time. In the current competitive business environment, change is not incidental. It is essential. Change must be managed. The AHPM is highly flexible so that managers may adjust its structure to evolving business environments. This flexibility will allow a company to improve its performance measurement system continuously.

Generally, a performance measurement system exists to monitor the implementation of planning of an organization and aid to motivate desirable individual performance through a realistic communication of performance information in related goals of business. This premise of performance measurement requires a significant number of feedbacks and corrective actions in the practice of accounting information systems (Nanni et al. 1990), i.e., feedbacks between levels and also within level. In the implementation of AHPM, these activities are common procedures. A clean separation of levels in the hierarchy of the AHPM is an idealization for simplifying the presentation. The flexibility of the AHPM, however, accommodates as many feedbacks and corrective actions as possible.

Performance is monitored by group rather than by an individual because of its ever-increasing importance of teamwork in the success of businesses. The integrated structure of the AHPM facilitates group decision and thus increases the chance that managers put trust in the resulting measures. This structure will enhance more involvement of lower-level managers as well as workers when a firm implements the AHPM. Furthermore, the hierarchy of the AHPM corresponds to that of business organization. As a result, group decision at each level is facilitated. For example, middle-level managers will be responsible for determining weights for middle-level criteria while lower-level managers will be responsible for operational level criteria.

The iterative process of weighing goals among managers, as the implementation of AHPM progresses, will help them to understand which strategic factors are important and how these factors are linked to other goals to be a successful company as a group.

Information technology plays a critical role in designing a performance measurement system to provide timely information to management. A computer-based decision support system can be used to utilize this conceptual foundation in a real-world situation. The AHPM can be easily stored in an enterprise database because of the commercial software, *Expert Choice*. As a result, the AHPM can be readily available to each management level via the network system of an enterprise. The AHPM fits any corporate information architecture to pursue the company's long-term strategy.

29.2.5 Conclusions

A well-designed performance model is a must for an enterprise to gain competitive edges. The performance measurement model proposed in this study is the first kind of analytical model to cover a wide variety of measures while providing operational control as well as strategic control. With comparison to previous evaluation methods, the model shows advantages such as flexibility, feedbacks, group evaluation, and computing simplicity. A prototype was built via a personal computer so that the model can be applied to any business situations.

The possible real-time control is of importance for the competitive business environment we are facing today. In sum, the contribution of this study is of both conceptual and practical importance.

29.3 Developing a Comprehensive Performance Measurement System in the Banking Industry: An Analytic Hierarchy Approach

The objective of this study is to design a practical model for a comprehensive performance measurement system that incorporates strategic success factors in the banking industry. A performance measurement system proposed in this study can be used to evaluate top managers of a main bank or managers of a branch office.

Performance measurement should be closely tied to goal setting of an organization because it feeds back information to the system on how well strategies are being implemented (Chan and Lynn 1991; Eddie et al. 2001). A balanced scorecard approach (Kaplan and Norton 1992; Tapinos et al. 2011) is a hot topic currently in this area, but it does not provide systematic aggregation of each level as well as different levels of managers' performance for the overall company. In other words, there is no systematic linkage between financial and nonfinancial measures across different levels of management hierarchy. A traditional performance measurement system, which focuses on financial measures such as return on assets (ROA), however, may not serve this purpose well for middle- or lower-level managers in the new competitive business environment either.

The model proposed in this study will have a broader set of measures that incorporate traditional financial performance measures such as return on assets and debt to equity ratio as well as nonfinancial performance measures such as the quality of customer service and productivity. Using Saaty's analytic hierarchy process (AHP) (Saaty 1980; Harker and Vargas 1987), the model will demonstrate how multiple performance criteria can be systematically incorporated into a comprehensive performance measurement system. The AHP enables decision makers to structure a problem in the form of a hierarchy of its elements according to an organization's structure or ranks of management levels and to capture managerial decision preferences through a series of comparisons of relevant factors or criteria. The AHP has been applied recently to several business problems (e.g., divisional performance evaluation (Chan and Lynn 1991), capital budgeting

(Liberatore et al. 1992), and real estate investment (Kamath and Khaksari 1991), marketing applications (Dyer and Forman 1991), information system project selection (Schniederjans and Wilson 1991), activity-based costing cost driver selection (Schniederjans and Garvin 1997), developing lean performance measures (DeWayne 2009), and customer's choice analysis in retail banking (Natarajan et al. 2010)). The AHP is relatively easy to use and its commercial software is available. This study will be an analytical and comprehensive performance evaluation model to cover a broader base of measures in the rapidly changing environment of today's banking industry. The following section presents background. The third section discusses methodology. The fourth section presents a numerical example. The last section summarizes and concludes this chapter.

29.3.1 Background

Recent research topic guide from Institute of Management Accountants lists "performance measurement" as one of top priority research issues. Therefore, this project will be interesting to bank administrators as well as managerial accountants.

With unprecedented competitive pressure from nonbanking institutions, deregulation, and the rapid acquisition of smaller banks by large national or regional banks, the most successful banks in the new millennium will be the ones that adapt strategically to a changing environment (Calvert 1990). Management accounting and finance literature have emphasized using both financial and nonfinancial measures as performance guidelines in the new environment (e.g., Chan and Lynn 1991; Rotch 1990). However, most studies do not propose specifically how we should incorporate these financial and nonfinancial factors into a formal model.

The performance measurement system proposed in this study is the formal model applied using the AHP in the banking industry to cover a wide variety of measures while providing operational control as well as strategic control. The AHP can incorporate multiple-subjective goals into a formal model (Dyer and Forman 1991). Unless we design a systematic performance measurement system that includes financial as well as nonfinancial control factors, there may be incorrect behavior by employees because they misunderstand the organization's goals and how they relate to their individual performance.

Compared with previous evaluation methods, the model proposed in this study will have advantages such as flexibility, continuous feedback, teamwork in goal setting, and computational simplicity. To be used for any business performance measurement, a well-designed model must be flexible enough to incorporate a variety of measures while retaining major success factors. The AHP model is flexible enough for managers to adjust its structure to a changing business environment through an iterative process of weighing goals. This flexibility will allow a company to improve its performance measurement system continuously. Through the iterative process of goal comparisons, management could get continuous feedback for the priority of goals and work as a team. The possible real-time control is of importance in the competitive business environment we are facing today.

29.3.2 Methodology

The analytic hierarchy process (AHP) collects input judgments in the form of a matrix by pairwise comparisons; i.e., two criteria are compared at one time. The experience of many users of this method supports use of a 1–9 scale for pairwise comparisons to capture human judgment while the scale can be altered to fit each application (Saaty 1980).

A simple example will be provided to explain how AHP operates. Consider the situation where a senior executive has to decide on which of three managers to promote to a senior position in the firm. The candidate's profiles have been studied and rated on three criteria: leadership, human relations skills, and financial management ability. First, the decision maker compares each of the criteria in pairs to develop a ranking of the criteria. In this case, the comparisons would be:

1. Is leadership more important than human relations skills for this job?
2. Is leadership more important than financial management ability for this job?
3. Are human relations skills more important than financial management ability for this job?

The response to these questions would provide an ordinal ranking of the three criteria. By adding a ratio scale of 1–9 for rating the relative importance of one criterion over another, a decision maker could make statements such as "leadership is four times as important as human relations skills for this job," "financial management ability is three times as important as leadership," and "financial management ability is seven times as important as human relations skills." These statements of pairwise comparisons can be summarized in a square matrix. The preference vectors are then computed to determine the relative rankings of the three criteria in selecting the best candidate. For example, the preference vectors of the three criteria are 0.658 for financial management ability, 0.263 for leadership, and 0.079 for human relations skills.

Once the preference vector of the criteria is determined, each of the candidates can be compared on the basis of the criteria in the following manner:

1. Is candidate A superior to candidate B in leadership skills?
2. Is candidate A superior to candidate C in leadership skills?
3. Is candidate B superior to candidate C in leadership skills?

Again, rather than using an ordinal ranking, the degree of superiority of one candidate over another can be assessed. The same procedures can be applied to human relations skills and financial management ability. The responses to these questions can be summarized in matrices where the preference vectors are again computed to determine the relative ranking of the three candidates for each criterion. Accordingly, the best candidate should be the one who ranks "high" on the "more important" criteria. The matrix multiplication of preference vectors of candidates on evaluation criteria and the preference vector of evaluation criteria will provide the final ranking of the candidates. In this example, the candidates are ranked A, B, and C. This example provides the usefulness of the AHP for setting priorities for both qualitative and quantitative measures (Chan and Lynn 1991).

We could apply the same procedures to a bank. Based on the hierarchical structure of a banking institution, the relative weights of criteria at each level of managers are derived by AHP. Here the relative importance of performance measures can be defined as a performance index with respect to each alternative (i.e., criterion, service, or branch office). For example, derive the relative weights of financial and nonfinancial criteria at the highest management level through the pairwise comparisons of criteria. Next, the relative weights of middle-level performance criteria with relation to each top-level criterion are to be computed then enter these relative weights into an *n* x *n* matrix format. Finally, this matrix is multiplied by the relative weights of criteria of the top management level. The same procedures can be applied to the next lower-level management. This AHP approach could link systematically the hierarchical structure of business performance measurement between different levels of organizational structures.

Consider a local bank where market share and return on assets are two important criteria in performance evaluation. If the AHP generates their weights as 0.4 and 0.6, respectively, it is reasonable to conclude that return on assets affects the bank's performance one and a half times higher than market share. The weights can be used as a measure for allocating future resources in products or branch offices. Assume the bank has two types of services, commercial loans and residential mortgage loans. If the AHP generates relative weights as 0.8 and 0.2 (i.e., the performance of commercial loans is four times higher than mortgage loans), then this may provide a good reason for top management to invest resources in the commercial loan market four times higher than the residential loan market.

The use of the AHP for multiple-criteria situation is superior to ad hoc weighing because it has the advantage of forcing the decision maker to focus exclusively on the criteria at one time and the way in which they are related to each other (Saaty 1980).

A model could be built using a microcomputer program called *Expert Choice* so that the model can be applied to any bank easily.

29.3.3 A Numerical Example

This section presents a numerical example for a commercial bank. The Commercial Omaha Bank (COB) is a local bank that specializes in commercial loans. Their headquarters are located in Omaha, Nebraska, and they have several branch offices throughout rural areas of Nebraska. The top management of COB realized that the current measurement system is not adequate for their strategic performance management and identified the following measures based on the hierarchy of the organization for their new performance measurement using AHP. These measures are shown in Table 29.1.

The OCB uses financial criteria such as return on assets and debt to equity and nonfinancial criteria such as market share, productivity, and quality of service. At the lowest management level, income to interest expense, service charges, interest revenue, growth of deposits, default ratio, and customer satisfaction can be used. Each computing step of the AHP is discussed as follows.

First, nonfinancial and financial criteria are computed and the result is entered in a vector:

$$W_t = (0.5, 0.5).$$

Table 29.1 New performance measures

Level of organization	High	Middle	Low
	Financial measures	Return on assets	Income to interest expenses
		Debt to equity	Service charges
			Interest revenue
	Nonfinancial measures	Market share	Growth of deposits
		Productivity	Default ratio
		Quality	Customer satisfaction

Next, the mid-level management is considered. The relative weight of mid-level performance criteria with relation to each top-level criterion is to be computed. Here, the local relative weights are computed.

For the nonfinancial criteria,

$$A_m^1 \begin{bmatrix} 1 & 3 & 4 \\ 1/3 & 1 & 3 \\ 1/4 & 1/3 & 1 \end{bmatrix}.$$

Here, the market share is estimated to be three times more important than productivity and four times more important than the quality of service. Productivity is estimated to be three times more important than the quality of service. From this result,

$$W_m^1 = (0.608, 0.272, 0.120).$$

For each weight computation, an inconsistency ratio (γ) was computed and checked for the acceptance level. If $\gamma \leq 0.1$, it is acceptable. For this example, it is acceptable since $\gamma = 0.065$. If it is not acceptable, the input matrix should be adjusted or recomputed.

For the financial criteria, ROA is estimated to be three times more important than debt to equity ratio. Therefore,

$$W_m^2 = (0.75, 0.25).$$

The global relative weights of the criteria are

$$\begin{aligned} \mathrm{W_m} &= (0.5, 0.5) \times \begin{pmatrix} 0.608, 0.272, 0.120, 0, 0 \\ 0, 0, 0, 0.75, 0.25 \end{pmatrix} \\ &= (0.304, 0.136, 0.060, 0.375, 0.125). \end{aligned}$$

Here the global relative weights of market share, productivity, quality of service, ROA, and debt to equity ratio are 30.4 %, 13.6 %, 6 %, 37.5 %, and 12.5 %, respectively.

Let us move to the lower-level managers. Income to interest expense, service charges, interest revenue, growth of deposits, default ratio, and customer satisfaction are criteria at this level.

For market share,

$$A_0^1 = \begin{bmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 1/2 & 1 & 2 & 3 & 4 & 5 \\ 1/3 & 1/2 & 1 & 3 & 4 & 5 \\ 1/4 & 1/3 & 1/3 & 1 & 3 & 4 \\ 1/5 & 1/4 & 1/4 & 1/3 & 1 & 3 \\ 1/6 & 1/5 & 1/5 & 1/4 & 1/3 & 1 \end{bmatrix}$$

For simplicity of presentation, the local weights are arranged in the order of income to interest expense, service charges, interest revenue, growth of deposits, default ratio, and customer satisfaction:

$$W_o^1 = (0.367, 0.238, 0.183, 0.109, 0.065, 0.038).$$

Here the local weights with relation to market share are 36.7 %, 23.8 %, 18.3 %, 10.9 %, 6.5 %, and 3.8 %, respectively.

For other criteria at one level higher, the local weights can be calculated in the same way. These are

$$W_o^2 = (0.206, 0.163, 0.179, 0.162, 0.143, 0.146);$$

$$W_o^3 = (0.155, 0.231, 0.220, 0.103, 0.169, 0.122);$$

$$W_o^4 = (0.307, 0.197, 0.167, 0.117, 0.117, 0.094);$$

$$W_o^5 = (0.266, 0.133, 0.164, 0.159, 0.154, 0.124).$$

For the next step, the global relative weights of lower-level management criteria are

0.367 0.238 0.183 0.109 0.065 0.038

$$\mathrm{W}_0 = (0.304, 0.136, 0.060, 0.375, 0.125) \times \begin{bmatrix} 0.206 & 0.163 & 0.179 & 0.162 & 0.143 & 0.146 \\ 0.155 & 0.231 & 0.220 & 0.103 & 0.169 & 0.122 \\ 0.307 & 0.197 & 0.167 & 0.117 & 0.117 & 0.094 \\ 0.266 & 0.133 & 0.164 & 0.159 & 0.159 & 0.124 \end{bmatrix}$$
$$= (0.297, 0.199, 0.176, 0.125, 0.112, 0.089).$$

Finally, the relative importance of lower-level performance measures are 29.7 %, 19.9 %, 17.6 %, 12.54 %, 11.2 %, and 8.9 %, respectively. Note that financial measures are integrated with nonfinancial measures in the scaling process. The OCB can extend these performance measures into the next lower level of each product using the same method.

29.3.4 Summary and Conclusions

A performance measurement system should incorporate nonfinancial as well as financial measures to foster strategic success factors for a bank in the new environment. Generally, a good performance measurement system should monitor employees' behavior in a positive way and be flexible enough to adapt to the changing environment. To motivate employees, a bank should communicate performance information of an individual employee in relation to overall business goals. This characteristic of performance measurement requires a significant amount of feedback both between and within levels and corrective actions in the practice of accounting information (Nanni et al. 1990).

The AHP model proposed in this study is flexible enough to incorporate the "continuous improvement" philosophy of today's business environment by changing weighting values of measures. In addition, the integrated structure of AHP allows group performance evaluation, which is a buzzword for "teamwork" in today's business world. The iterative process of getting input data in the AHP procedure also helps each manager as well as employee to be aware of the importance of strategic factors of each performance measure of the bank.

29.4 Optimal Trade-offs of Multiple Factors in International Transfer Pricing Problems

Ever since DuPont and General Motors Corporation of the USA initiated transfer pricing systems for the interdivisional transfer of resources among their divisions, many large organizations, with the creation of profit centers, have used a transfer pricing system in one way or the other. Recently, transfer pricing problems have become more important because most corporations increase transfer of goods or services dramatically among their divisions as a result of restructuring or downsizing their organizations (Tang 1992). Therefore, designing a good transfer pricing strategy should be a major concern for both top management and divisional managers (Curtis 2010).

Transfer pricing problems have been extensively studied by a number of scholars. Many of them have recognized that a successful transfer pricing strategy should consider multiple criteria (objectives), such as overall profit, total market share, divisional autonomy, performance evaluation, and utilized production capacity (Abdel-khalik and Lusk 1974; Bailey and Boe 1976; Merville and Petty 1978; Yunker 1983; Lecraw 1985; Lin et al. 1993). However, few developed methods have a capability of dealing with all possible optimal trade-offs of multiple criteria in optimal solutions of the models with involvement of multiple decision makers.

In this chapter, we propose a multiple factor model to provide managers from different background, who are involved in transfer price decision making of a multidivisional corporation, with a systematic and comprehensive scenario about all possible optimal transfer prices depending on both multiple-criteria and multiple-constraint levels (in short, multiple factors). The trade-offs of the optimal transfer prices, which have rarely been considered in the literature, can be used as a basis for managers of corporations to make a high-quality decision in selecting their transfer pricing systems for business competition.

This chapter proceeds as follows. First, existing transfer pricing models will be reviewed. Then, the methodology of formulating and solving a transfer pricing model with multiple factors will be described. A prototype of a transfer pricing problem in a corporation will be illustrated to explain the implications of the multiple factor transfer pricing model. Finally, conclusions and remaining research problems will be presented.

29.4.1 Existing Transfer Pricing Models

In the literature of transfer pricing problems, the various approaches can be categorized into four groups: (i) market-based pricing, (ii) accounting-based pricing, (iii) marginal cost pricing (or opportunity cost pricing), and (iv) negotiation-based pricing. Market-based prices are ideal for transfer prices when external market prices are available. Even though empirical research has found that some corporations prefer cost-based prices to market-based prices, market-based pricing method is recommended when the emphasis is on the motivation of divisional managers (Borkowski 1990). Pricing intermediate goods based on the market price will motivate the supplying division to reduce its costs to achieve efficiency and to allow divisional autonomy for both the supplying division and the purchasing division. Statistics show that almost a third of corporations actually use market-based transfer pricing (Tang 1992).

However, if there is no outside market for intermediate goods or services, then accounting-based pricing, marginal cost pricing (economic models and mathematical programming techniques), or negotiation-based pricing (behavioral approach) is commonly recommended for finding a transfer pricing system.

In the accounting-based pricing approach, the divisional managers simply use accounting measurements of the divisions, such as full costs or variable costs, as their transfer prices. Thus, the transfer price of one division may differ from that of another division. These transfer prices may not be globally optimal for the corporation as a whole (Abdel-khalik and Lusk 1974; Eccles 1983).

In marginal cost pricing approaches, Hirshleifer (1956) recommended use of an economic model to set transfer pricing at a manufacturing division's marginal cost to achieve the global optimal output. A problem of this economic model is that it can destroy the divisional manager's autonomy, and the supplying division may not get the benefit of efficiencies. Moreover, the manufacturing division manager in this case should be evaluated based on cost, not on profit. Similarly, Gould (1964) and Naert (1973) recommended economic models based on current entry and current exit prices. Their models also focus on global profit maximization and have the same problems as Hirshleifer's. Note that when the transfer price is set based on marginal costs, the division should be either controlled as a standard cost center or merged into a larger profit center with a division that processes the bulk of its output.

Ronen and McKinney (1970) suggested dual prices in which a subsidy is given to the manufacturing division by the central office in addition to marginal costs. This subsidy would not be added to the price charged to the purchasing division. They believed that autonomy is enhanced because the corporate office is only a transmitter of information, not a price setter, and that the supplying division has the same autonomy as an independent supplier. However, there might be a gaming chance where all divisions are winners but the central office is a loser. There is also a "marginal cost plus" approach that charges variable manufacturing costs (usually standard variable costs) and is supplemented by periodic lump-sum charges to the transferees to cover the transferor's fixed costs, or their fixed costs plus profit, or some kind of subsidy. It is difficult, however, to set a fixed fee that will satisfy both the supplying division and the purchasing division. This method enables the purchasing division to absorb all the uncertainties caused by fluctuations in the marketplace. Moreover, the system begins to break down if the supplying division is operating at or above the normal capacity since variable costs no longer represent the opportunity costs of additional transfers of goods and services (Onsi 1970).

As an alternative method to identify marginal costs as the transfer prices, a mathematical programming approach becomes more attractive for the transfer pricing problem because it handles complex situations in a trade setting (Dopuch and Drake 1964; Bapna et al. 2005). The application of linear programming to the transfer pricing problem is based on the relationship between the primal and dual solutions in the linear programming problem. Shadow prices, which reflect the input values of scarce resources (or opportunity cost) implied in the primal problem, can be used as the basis for a transfer price system. However, these transfer prices have the following limitations for decision making: (i) those transfer prices based on dual values of a solution tend to reward divisions with scarce resources, (ii) the linear formulation requires a great deal of local information, and (iii) transfer prices based on shadow prices do not provide a guide for performance evaluation of divisional managers.

Based on the mathematical decomposition algorithm developed by Dantzig and Wolfe (1960), Baumol and Fabian (1964) demonstrated how to reduce complex optimization problems into sets of smaller problems solvable by divisions and the central office (corporation). Although the final analysis of output decisions is made by the central manager, the calculation process is sufficiently localized that central management does not have to know anything about the internal technological arrangements of the divisions. However, this approach does not permit divisional autonomy.

Ruefli (1971) proposed a generalized goal decomposition model for incorporating multiple criteria (objectives) and some behavioral aspects within a three-level hierarchical organization into the mathematical formulation of transfer pricing problems. Bailey and Boe (1976) suggested another goal programming model as a supplement of Ruefli's model. Both models overcome the shortcoming of linear programming in dealing with multiple criteria and the organizational hierarchy. Merville and Petty (1978) used a dual formulation of goal programming to directly find the shadow price as the optimal transfer price for multinational corporations. However, the optimal solution of these models that results in the transfer prices is determined by a particular distance (norm) function because of the mathematical structure of goal programming. Thus, the optimal solution represents only a single optimal trade-off of the multiple criteria, not all possible optimal trade-offs.

Watson and Bulmer (1975) and Thomas (1980) criticized the lack of behavioral considerations in the mathematical programming approaches. They suggested a negotiated transfer pricing to further the integration of differentiation within the organization. The differentiation of the organization exists because, for example, different managers can interpret the same organizational problem differently. Generally, the negotiated transfer pricing will be successful under conditions such as the existence of some form of outside markets for the intermediate product, freedom to buy or sell outside, and sharing all market information among the negotiators. Negotiated prices may require iterative exchanges of information with the central office as a part of a mathematical programming algorithm.

29.4.2 Methodology

29.4.2.1 MC^2 Linear Programming

Practically speaking, since linear programming has only a single criterion (objective) and a single resource availability level (right-hand side), it has limitations in handling real-world transfer pricing problems. For instance, linear programming cannot be used to solve the problem in which a corporation tries to maximize overall profit and the total market share simultaneously. This dilemma is overcome by a technique called multiple-criteria (MC) linear programming (Zeleny 1974; Goicoechea et al. 1982; Steuer 1986; Yu 1985). To extend the framework of MC linear programming, Seiford and Yu (1979) and Yu (1985) formulated a model of multiple-criteria and multiple-constraint level (MC^2) linear programming. This model is rooted by two facts. First, from the linear system structure's point of view, the criteria and constraints may be "interchangeable." Thus, like multiple criteria, multiple-constraint (resource availability) levels can be considered. Second, from the application's point of view, it is more realistic to consider multiple resource availability levels (discrete right-hand sides) than a single resource availability level in isolation. The philosophy behind this perspective is that the availability of resources can fluctuate depending on the decision situation forces, such as the desirability levels believed by the different managers. For example, if the differentiation of budget among managers in transfer pricing problems (Watson and Baumler 1975) is represented by different levels of budget, then this differentiation can be resolved by identifying some best compromise of budget levels as the consensus budget.

The theoretical connections between MC linear programming and MC^2 linear programming can be found in Gyetvan and Shi (1992). Decision problems related to MC^2 linear programming have been extensively studied in Lee et al. (1990), Shi (1991), and Shi and Yu (1992). Key ideas of MC^2 linear programming, a primary theoretical foundation of this chapter, are outlined as follows.

An MC^2 linear programming problem can be formulated as

$$\begin{aligned} &\text{Max}\, \lambda^t \text{Cx} \\ &\text{s.t.}\;\; \text{Ax} \leq \text{D}\gamma \\ &\text{x} \geq 0, \end{aligned}$$

where $C \in \mathbf{R}^{\mathbf{qxn}}$, $A \in \mathbf{R}^{\mathbf{mxn}}$, and $D \in \mathbf{R}^{\mathbf{mxp}}$ are matrices of qxn, mxn, and mxp dimensions, respectively; $x \in \mathbf{R}^{\mathbf{n}}$ are decision variables; $\lambda \in \mathbf{R}^{\mathbf{q}}$ is called the *criteria parameter*; and $\gamma \in \mathbf{R}^{\mathbf{p}}$ is called the *constraint level parameter*. Both (γ,λ) are assumed unknown.

The above MC^2 problem has q criteria (objectives) and p constraint levels. If the constraint level parameter γ is known, then the MC^2 problem reduces to an MC linear programming problem (e.g., Yu and Zeleny 1975). In addition, if the criteria parameter λ is known, it reduces to a linear programming problem (e.g., Charnes and Cooper 1961; Dantzig 1963).

Denote the index set of the basic variables $\{x_{ji}, \ldots, x_{jm}\}$ for the MC^2 problem by $J = \{j_1, \ldots, j_m\}$. Note that the basic variables may contain some slack variables. Without confusion, J is also called a basis for the MC^2 problem. Since a basic solution J depends on parameters (γ, λ), define that (i) a basic solution J is *feasible* for the MC^2 problem if and only if there exists a $\gamma^0 > 0$ such that J is a feasible solution for the MC^2 problem with respect to γ^0 and (ii) J is *potentially optimal* for the MC^2 problem if and only if there exist a $\gamma^0 > 0$ and a $\lambda^0 > 0$ such that J is an optimal solution for the MC^2 problem with respect to (γ^0, λ^0). Let $\Gamma(J)$ be the *constraint level parameter set* of all γ such that the basis J is feasible and $\Lambda(J)$ be the *criteria parameter set* of all λ that the basis J is dual feasible. Then, for a given basis J of the MC^2 problem, (i) J is a feasible solution if and only if the set $\Gamma(J)$ is not empty and (ii) J is potentially optimal if and only if both sets $\Gamma(J)$ and $\Lambda(J)$ are not empty. For an MC^2 problem, there may exist a number of potentially optimal solutions {J} as parameters (γ, λ) vary depending on decision situations.

Seiford and Yu (1979) derived a simplex method to systematically locate the set of all potentially optimal solutions {J}. The computer software of the simplex method (called MC^2 software) was developed by Chien et al. (1989). This software consists of five subroutines in each iteration: (i) pivoting, (ii) determining primal potential bases, (iii) determining dual potential bases, (iv) determining the effective constraints for the primal weight set, and (v) determining the effective constraints for the dual weight set (Chap. 8 of Yu 1985). It is written in PASCAL and operates in a man-machine interactive fashion. The user cannot only view the tableau of each iteration but also trace the past iterations. In the next section, the framework of the MC^2 problem, as well as its software, will be used to formulate and solve the multiple factor transfer pricing problems.

29.4.2.2 Multiple Factor Transfer Pricing Model

The review of existing transfer pricing models shows that two major shortcomings, from a management point of view, need to be overcome in the previous mathematical models of transfer pricing problems. First, neither the linear programming approach nor the goal programming approach can provide a comprehensive scenario of all possible optimal trade-offs between multiple objectives under consideration for a given transfer pricing problem, such as maximizing the overall profits for a corporation and minimizing the underutilization of production capacity (Tang 1992). Transfer pricing scheme by linear programming only reflects a single objective of a corporation. As a result, the linear programming approach cannot help the corporation seek to simultaneously achieve several objectives, some in conflict, in business competition. The transfer price determined by the goal programming approach is an optimal compromise (i.e., trade-off) among several objectives of the corporation. However, it misses other possible optimal compromises of the objectives that result from some linear combinations of objective weights. These compromises lead to different optimal transfer prices for different decision situations that the corporation may face. Second, none of the past mathematical models can deal with the organizational differentiation problems, as Watson and Baumler (1975) pointed out. In real-life cases, when a corporation designs its transfer prices for the divisions, the involved decision makers (executives or the members of the task force) can give different opinions on the same issue, such as production capacity and customer's demand. In mathematical models, these different interpretations can be represented by different "constraint levels." Because both linear programming and goal programming presume a fixed single constraint level, they fail to mathematically describe such an organizational differentiation problem.

The MC^2 linear programming framework can resolve the above shortcomings inherent in the previous transfer pricing models. Based on Yunker (1983) and Tang (1992), the four important objectives of transfer pricing problems in most corporations are considered: (i) maximizing the overall profit, (ii) maximizing the total market share, (iii) maximizing the subsidiary profit (note that the subsidiary profit maximization is used to reflect the degree of the subsidiary autonomy in decision making. It may differ from the overall profit), and (iv) maximizing the utilized production capacity. Even though the following model contains only these four specific objectives, the generality of the modeling process fits in all transfer pricing problems with multiple-criteria and multiple-constraint levels.

Let k be the number of divisions in a corporation under consideration and t be the index number of the products that each division of the corporation produces.

Define x_{ij} as the units of the jth product made by the ith division, $i = 1, \ldots, k$; $j = 1, \ldots, t$. For the coefficients of the objectives, let p_{ij} be the unit overall profit generated from the jth product made by the ith division, m_{ij} be the market share value for the jth product made by the ith division in the market, s_{ij} be the unit subsidiary profit generated from the jth product made by the ith division, and c_{ij} be the unit utilized production capacity of the ith division to produce the jth product. For the coefficients of the constraints, let b_{ij} be the budget allocation rate for producing the jth product by the ith division. For the coefficients of the constraint levels, let b_{ij}^{s}

be the budget availability level believed by the sth manager (or executive) for producing the jth product by the ith division, s = 1, . . . , h; d_i^s be the production capacity level believed by the sth manager for the ith division; d_{ij}^s be the production capacity level believed by the sth manager for the ith division to produce the jth product; and e_{ij}^s be the initial inventory level believed by the sth manager for the ith division to hold the jth product. Then, the multiple factor transfer pricing model is

$$\begin{aligned}
&\text{Max } \Sigma_{i=1}^{k} \quad \Sigma_{j=1}^{l} \quad \mathrm{p_{ij}x_{ij}}\\
&\text{Max } \Sigma_{i=1}^{k} \quad \Sigma_{j=1}^{l} \quad \mathrm{m_{ij}x_{ij}}\\
&\text{Max } \Sigma_{i=1}^{k} \quad \Sigma_{j=1}^{l} \quad s_{\mathrm{ij}}\mathrm{x_{ij}}\\
&\text{Max } \Sigma_{i=1}^{k} \quad \Sigma_{j=1}^{l} \quad c_{\mathrm{ij}}\mathrm{x_{ij}}\\
&\text{Subject to } \Sigma_{i=1}^{k} \quad \Sigma_{j=1}^{l} \quad \mathrm{bijxij} \leq \left(\mathrm{b^1{}_{ij}}, \ldots, \mathrm{b^h{}_{ij}}\right) \qquad (29.1)\\
&\Sigma_{j=1}^{k} \quad \mathrm{xij} \leq \left(\mathrm{d^1{}_{ij}}, \ldots \mathrm{d^h{}_i}\right)\\
&\mathrm{x_{ij}} \leq \left(\mathrm{d^1{}_{ij}}, \ldots, \mathrm{d^h{}_{ij}}\right)\\
&-\mathrm{x_{ij}} + \mathrm{x}_{\mathrm{i}+1,j} \leq \left(\mathrm{e^1{}_{ij}}, \ldots, \mathrm{e^h{}_{ij}}\right)\\
&\mathrm{X_{ij}} \geq 0, i = 1, \ldots, k, j = 1, \ldots, t.
\end{aligned}$$

In the next section, a prototype of the transfer pricing model in a corporation will be illustrated to demonstrate the implications for decision makers.

29.4.3 Model Implications

29.4.3.1 Numerical Example

As an illustration of the multiple factor transfer pricing model, the United Chemical Corporation has two divisions that process raw materials into intermediate or final products. Division 1, which is located in Kansas City, manufactures two kinds of chemicals, called Products 1 and 2. Product 1 in Division 1 is intermediate product and cannot be sold externally. It can, however, be processed further by Division 2 into a final product. Division 2, which is located in Atlanta, manufactures Product 2 and finalizes Product 1 in Division 1. The executives (president, vice president for production, and vice president for finance) and all divisional managers agree on the following multiple objectives:

(i) Maximize the overall company's profit.
(ii) Maximize the market share goal of Product 2 in Division 1 and Products 1 and 2 in Division 2.
(iii) Maximize the utilized production capacity of the company so that each division manager can avoid any underutilization of normal production capacity.

The data related to these objectives is given in Table 29.2.

In Table 29.2, Product 1 of Division 1 is a by-product that has no sale value in Division 1 at all. The unit profit −$4 of this product means its production cost.

Table 29.2 Data of the objectives in the United Chemical Corporation

Division 1	Product 1	Product 2	Division 2	Product 1	Product 2
Unit profit ($)	−4	8		13	5
Market price	0	40		46.2	38
Unit utilizing production capacity (h)	4	4		3	2

All other products generate profits. We use a market price of each product to maximize the market share goal. Note that the market prices of three products in Table 29.2 are different.

In the company, besides the president, the vice presidents for production and for finance are the main decision makers and may have different interpretations of the same resource availability across the divisions. The vice president for production views the constraint level based on the material, manpower, and equipment under control, while the vice president for finance views the constraint level based on the available cash flow. The president will make the final decision for the company's transfer price setting on the basis of the compromises of both vice presidents. All divisional managers will carry out the president's decision, although they have their autonomy to provide the information about the divisional profits and costs for the executives. The interpretations of the vice presidents for the resource constraint levels are summarized in Table 29.3.

All executives and divisional managers agree on the units of resources consumed to produce the products. The data is given in Table 29.4.

Let x_{ij} be the units of the jth product produced by the ith division, i = 1, 2; j = 1, 2. Using the information in Tables 29.2, 29.3, and 29.4, the multiple factor transfer pricing model is formulated as

$$\begin{array}{l} \text{Max} \quad -4x_{11} + 8x_{12} + 13x_{21} + 5x_{22} \\ \text{Max} \quad 40x_{12} + 46.2x_{21} + 38x_{22} \\ \text{Max} \quad 4x_{11} + 4x_{12} + 3x_{21} + 2x_{22} \\ \text{Subject to} \quad -x_{11} + x_{21} \le (0, 100) \\ \quad 0.4x_{12} + 0.4x_{21} + 0.4x_{22} \le (45{,}000, 40{,}000) \\ X_{12} \le (38{,}000, 12{,}000) \\ X_{21} \le (45{,}000, 50{,}000) \\ X_{22} \le (36{,}000, 10{,}000) \\ x_{ij} \le 0, i = 1, 2; j = 1, 2. \end{array} \tag{29.2}$$

Since this multiple factor transfer pricing problem is a typical MC^2 problem, the MC^2 software of Chien et al. (1989) can be used to solve the problem. Let $\lambda = (\lambda_1, \lambda_2, \lambda_3)$ be the weight parameter for the objectives, where $\lambda_1 + \lambda_2 + \lambda_3 = 1$ and $\lambda_1, \lambda_2, \lambda_3 \ge 0$. Let $\gamma = (\gamma_1, \gamma_2)$ be the weight parameter for the constraint levels, where $\gamma_1 + \gamma_2 = 1$ and $\gamma_1, \gamma_2 \ge 0$. Because both weight parameters (γ, λ) are unknown before design time, the solution procedure of MC^2 linear programming must be used to locate all possible potentially optimal solutions as (γ, λ) vary. The implications of the potentially optimal solutions for accounting decision makers will be explained in the next subsection. After putting (γ, λ) into the above model, it becomes

Table 29.3 The constraint levels of the vice presidents

	Vice president for production	Vice president for finance
Transfer product constraint	0	100
Budget constraint ($)	45,000	40,000
Production capacity of Product 2 in Division 1	38,000	12,000
Production capacity of Product 1 in Division 2	45,000	50,000
Production capacity of Product 2 in Division 2	36,000	10,000

Table 29.4 The unit consumptions of resources

Division 1	Product 1	Product 2	Division 2	Product 1	Product 2
Transfer product constraint	−1.0	0		1.0	0
Budget constraint ($)	0	0.4		0.4	0.4
Production capacity of Product 2 in Division 1		1.0			
Production capacity of Product 1 in Division 2				1.0	
Production capacity of Product 2 in Division 2					1.0

$$\begin{aligned}
\text{Max}\quad & \lambda_1(-4x_{11} + 8x_{12} + 13x_{21} + 5x_{22}) \\
& + \lambda_2(40x_{12} + 46.2x_{21} + 38x_{22}) \\
& + \lambda_3(4x_{11} + 4x_{12} + 3x_{21} + 2x_{22}) \\
\text{Subject to} & \\
& -x_{11} + x_{21} \le 100\gamma_2 \\
& 0.4x_{12} + 0.4x_{21} + 0.4x_{22} \le 45,000\gamma_1 + 40,000\gamma_2 \\
& x_{12} \le 38,000\gamma_1 + 12,000\gamma_2 \\
& x_{21} \le 45,000\gamma_1 + 50,000\gamma_2 \\
& x_{22} \le 36,000\gamma_1 + 10,000\gamma_2 \\
& x_{ij} \ge 0, i = 1,2; j = 1,2.
\end{aligned} \tag{29.3}$$

Let s_q, q = 1, . . . , 5, be the slack variables corresponding to the constraints. The MC^2 software yields two potentially optimal solutions $\{J_1, J_2\}$ and their associated values of (γ, λ) as shown in Table 29.5. Here, $V(J_i)$ denotes the objective value of potentially optimal solution J_i, which is a function of (γ, λ). When (γ, λ) are specified, $V(J_i)$ is the payoff of using J_i.

Table 29.5 means that if the γ takes the value from $\Gamma(J_1)$ and the λ takes the value from $\Lambda(J_1)$, J_1 is the optimal solution for the transfer pricing problem. In this case, products $\{x_{11}, x_{12}, x_{21}, x_{22}\}$ will be produced to achieve the objective payoff $V(J_1)$ and the resource of x_{22} has the amount of s5 unused. Similarly, the potentially optimal solution J_2 can be interpreted. However, J_2 is different from J_1

Table 29.5 All potentially optimal solutions

J_i	$\Gamma(J_i)$	$\Lambda(J_i)$	$V(J_i)$		
$J_1 = (x_{11}, x_{12}, x_{21}, x_{22}, s_5)$	$\gamma_1 + \gamma_2 = 1,$	$\lambda_1 + \lambda_2 + \lambda_3 = 1,$	856,500	736,400	γ
	$65\gamma_1 - 280\gamma_2 - 0,$	$\lambda_1 - \lambda_3,$	4,720,000	4,234,000	
	$\gamma_1, \gamma_2 - 0$	$\lambda_1, \lambda_2, \lambda_3 - 0$	526,000	47,444,000	
$J_2 = (x_{11}, x_{12}, x_{21}, x_{22}, s_2)$	$\gamma_1 + \gamma_2 = 1,$	$\lambda_1 + \lambda_2 + \lambda_3 = 1,$	889,000	596,400	γ
	$65\gamma_1 - 280\gamma_2 - 0,$	$\lambda_1 - \lambda_3,$	4,967,000	3,170,000	
	$\gamma_1, \gamma_2 - 0$	$\lambda_1, \lambda_2, \lambda_3 - 0$	539,000	41,184,000	

Table 29.6 All optimal transfer prices

J_i	$p_q(J_i)$
$J_1 = (x_{11}, x_{12}, x_{21}, x_{22}, s_5)$	$p_1(J_1) = 4\lambda_1 - 4\lambda_3,$
	$p_2(J_1) = 12.5\lambda_1 + 95\lambda_2 + 5\lambda_3,$
	$p_3(J_1) = 3\lambda_1 + 2\lambda_2 + 2\lambda_3$
	$p_4(J_1) = 4\lambda_1 + 8.2\lambda_2 + 5\lambda_3$
	$p_5(J_1) = 0$
$J_2 = (x_{11}, x_{12}, x_{21}, x_{22}, s_2)$	$p_1(J_2) = 4\lambda_1 - 4\lambda_3,$
	$p_2(J_2) = 0,$
	$p_3(J_2) = 8\lambda_1 + 40\lambda_2 + 4\lambda_3$
	$p_4(J_2) = 9\lambda_1 + 46.2\lambda_2 + 7\lambda_3$
	$p_5(J_2) = 5\lambda_1 + 38\lambda_2 + 2\lambda_3$

because there is the unused budget s_2 for J_2 and the parameter set $\Gamma(J_1)$ _ $\Gamma(J_2)$. Because Product 1 in Division 1 is a by-product and its unit profit is $-\$4$, whenever $\lambda_1 < \lambda_3$, both J_1 and J_2 are not optimal.

Let $p_q(J_i)$, $q = 1, \ldots, 5$; $i = 1, 2$, be the shadow price of J_i for the qth constraint. According to the marginal cost pricing approach, the optimal transfer prices of $\{J_1, J_2\}$ are designated as the shadow prices of $\{J_1, J_2\}$. These optimal transfer prices are found in Table 29.6. Table 29.6 shows that (i) the relative transfer price between x_{11} and x_{21} is $p_1(J_1) = 4\,\lambda_1 - 4\,\lambda_3$; (ii) the transfer price for budget across the divisions is $p_2(J_1) = 12.5\,\lambda_1 + 95\,\lambda_2 + 5\,\lambda_3$; (iii) the transfer price for x_{12} is $p_3(J_1) = 3\,\lambda_1 + 2\,\lambda_2 + 2\,\lambda_3$; (iv) the transfer price for x_{21} is $p_4(J_1) = 4\,\lambda_1 + 8.2\,\lambda_2 + 5\,\lambda_3$; and (v) the transfer price for x_{22} is $p_5(J_1) = 0$.

29.4.4 Optimal Trade-offs and Their Accounting Implications

Optimal trade-offs related to transfer prices of the multiple factor model consist of three components: (i) trade-offs among multiple objectives, (ii) trade-offs among multiple-constraint levels, and (iii) trade-offs between multiple objectives and multiple-constraint levels. The trade-offs among multiple objectives imply that all possible optimal compromises of the multiple objectives are determined by locating all possible weights of importance of these objectives. Similarly,

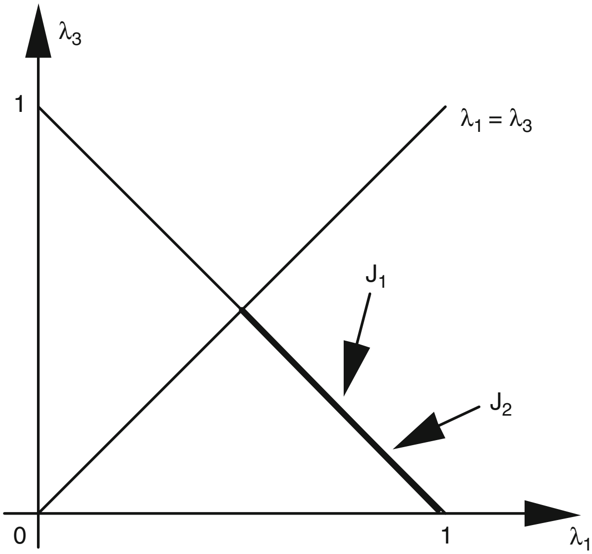


Fig. 29.1 Optimal trade-offs between profit (λ_1) and production capacity (λ_3)

the trade-offs among multiple-constraint levels imply that all possible optimal compromises of the multiple-constraint levels that represent the executives' different opinions are determined by locating all possible weights of importance of these opinions. The trade-offs between multiple objectives and multiple-constraint levels measure the effects on the transfer pricing problem by the interaction of the objectives and constraint levels. These three cases of the optimal trade-offs can be analyzed through the potentially optimal solutions of the multiple factor model. The optimal transfer prices also result from those solutions. In the following, three trade-off cases and accounting implications of the corresponding optimal transfer prices are explored in detail by using the above numerical example.

Because there are two potentially optimal solutions $\{J_1, J_2\}$ in the example, the optimal trade-offs should be studied in terms of both J_1 and J_2. For the trade-offs among three objectives that are the overall company's profit, the market share of the company, and the utilized production capacity of the company:

(i) If the overall company's profit is not considered (this implies $\lambda_1 = 0$ and $\lambda_3 \neq 0$), then either J_1 or J_2 is not optimal since $\lambda_3 < 0$ (see Table 29.5).

(ii) If the market share of the company is not considered (i.e., $\lambda_2 = 0$), then both J_1 and J_2 are optimal for which $\lambda_1 + \lambda_3 = 1$, $\lambda_1 \geq \lambda_3$, and λ_1, $\lambda 3 \geq 0$. The graphical representation is shown in Fig. 29.1. From Fig. 29.1, any weighting values of λ_1 and λ_3 taken from the feasible (dark) segment guarantee that the utility values of both the overall company's profit and the utilized production capacity of the company are maximized. The resulting optimal transfer prices associated with J_1 are $p_1(J_1) = 4\lambda_1 - 4\lambda_3$, $p_2(J_1) = 12.5\lambda_1 + 5\lambda 3$, $p_3(J_1) = 3\lambda_1 + 2\lambda_3$, $p_4(J_1) = 4\lambda_1 + 5\lambda_3$, and $p_5(J_1) = 0$, respectively. The resulting optimal transfer

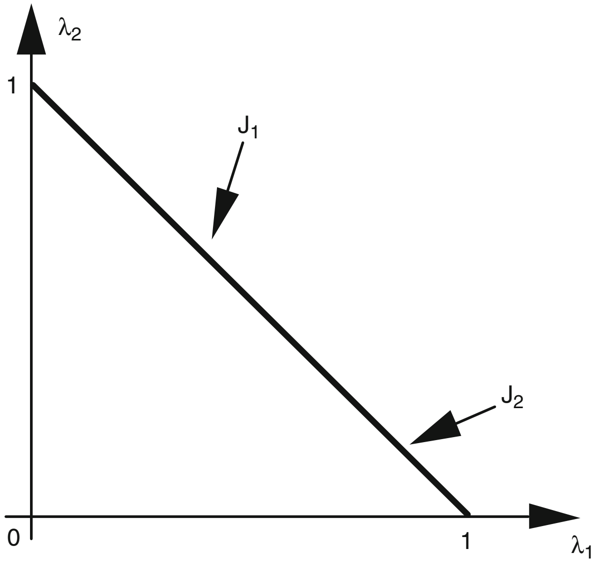


Fig. 29.2 Optimal trade-offs between profit (λ_1) and market share (λ_2)

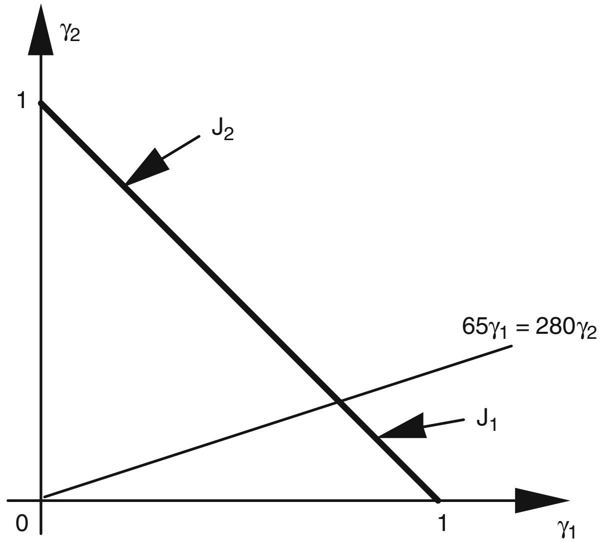
prices associated with J_2 are $p_1(J_2) = 4\lambda_1 - 4\lambda_3$, $p_2(J_2) = 0$, $p_3(J_2) = 8\lambda_1 + 4\lambda_3$, $p_4(J_2) = 9\lambda_1 + 7\lambda 3$, and $p_5(J_2) = 5\lambda 1 + 2\lambda 3$, respectively. The weighting values of λ_1 and λ_3 in the transfer prices for J_1 and J_2 are given in Fig. 29.1, because $\Lambda(J_1) = \Lambda(J_2)$.

(iii) If the utilized production capacity of the company is not considered (i.e., $\lambda_3 = 0$), then both J_1 and J_2 are optimal for which $\lambda_1 + \lambda_2 = 1$ and λ_1, λ_2 _ 0, where the graphical representation is shown in Fig. 29.2. Figure 29.2 implies that any weighted combination of λ_1 and λ_2 taken from the indicated feasible segment maximizes the utility values of both the overall company's profit and the market share of the company. The resulting optimal transfer prices of J_1 are $p_1(J_1) = 4\lambda_1$, $p_2(J_1) = 12.5\lambda_1 + 95\lambda_2$, $p_3(J_1) = 3\lambda_1 + 2\lambda_2$, $p_4(J_1) = 4\lambda_1 + 8.2\lambda_2$, and $p_5(J_1) = 0$, while the resulting optimal transfer prices of J_2 are $p_1(J_2) = 4\lambda_1$, $p_2(J_2) = 0$, $p_3(J_2) = 8\lambda_1 + 40\lambda_2$, $p_4(J_2) = 9\lambda_1 + 46.2\lambda_2$, and $p_5(J_2) = 5\lambda_1 + 38\lambda_2$, respectively.

For the trade-offs among two constraint levels (the different opinions of two vice presidents), the weight of the vice president for production is γ_1 while that of the vice president for finance is γ_2 such that $\gamma_1 + \gamma_2 = 1$, $\gamma_1, \gamma_2 \geq 0$. The range of γ_1 and γ_2 are decomposed into two subsets: $\Gamma(J_1) = \{\gamma_1, \gamma_2 \geq 0 \mid 65\gamma_1 - 280\gamma_2 \geq 0 \text{ and } \gamma_1 + \gamma_2 = 1\}$ and $\Gamma(J_2) = \{\gamma_1, \gamma_2 \geq 0 \mid -65\gamma_1 + 280\gamma_2 \geq 0 \text{ and } \gamma_1 + \gamma_2 = 1\}$ (see Table 29.5). The graphical representation of $\Gamma(J_1)$ and $\Gamma(J_2)$ is shown in Fig. 29.3. Whenever the weighting values of γ_1 and γ_2 are taken from $\Gamma(J_1)$, the corresponding compromise of two vice presidents will result in the optimal transfer prices of J_1 in Table 29.6. Similarly, the decision situation of constraint levels for J_2 can be explained.

Finally, there are many optimal trade-off situations between three objectives and two constraint levels involved with the transfer pricing problem. For example, two cases are illustrated as follows:

Fig. 29.3 Optimal trade-offs between V.P. for production (γ_1) and V.P. for finance (γ_2)



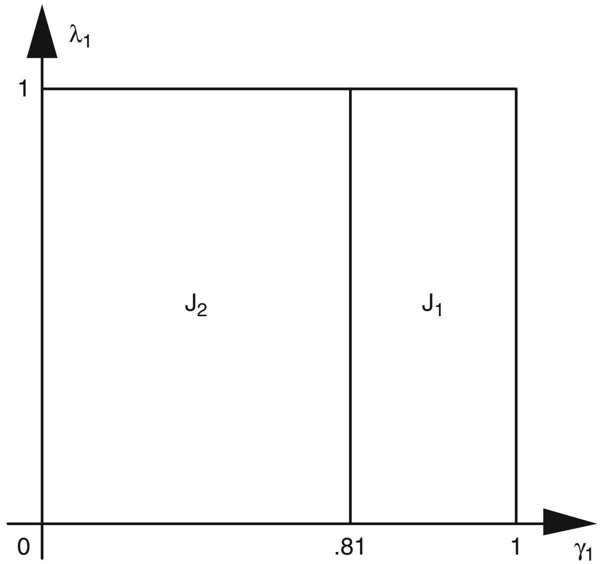


Fig. 29.4 Optimal trade-offs V.P. for production (γ_1) and total profit (λ_1)

(i) In the case that the utilized production capacity of the company is not considered (i.e., $\lambda_3 = 0$) in contrast to the weighting value of the market share and that of the vice president for finance (i.e., $0 \leq \lambda_2 \leq 1$ and $0 \leq \gamma_2 \leq 1$), the optimal trade-offs between the overall company's profit and the constraint level believed by the vice president for production for J_1 and J_2 are shown in Fig. 29.4. Here, if the values of (γ_1, λ_1) are taken from $0 \leq \lambda_1 \leq 1$ to $.81 \leq \gamma_1 \leq 1$, then the optimal transfer prices

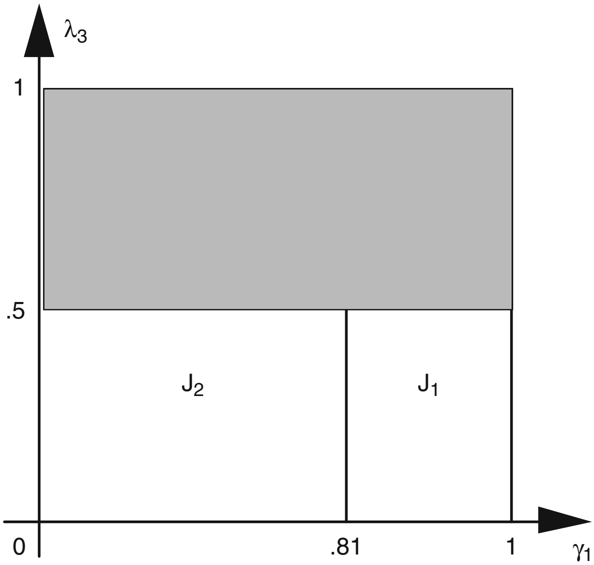


Fig. 29.5 Optimal trade-offs between V.P. for production (γ_1) and production capacity (λ_3)

associated with J_1 will be chosen. They are $p_1(J_1) = 4\lambda_1$, $p_2(J_1) = 12.5\lambda_1 + 95\lambda_2$, $p_3(J_1) = 3\lambda_1 + 2\lambda_2$, $p_4(J_1) = 4\lambda_1 + 8.2\lambda_2$, and $p_5(J_1) = 0$. Otherwise, the values of (γ_1, λ_1) are taken from $0 \le \lambda_1 \le 1$ to $0 \le \gamma_1 \le .81$, and the optimal transfer prices associated with J_2, $p_1(J_2) = 4\lambda_1$, $p_2(J_2) = 0$, $p_3(J_2) = 8\lambda_1 + 40\lambda_2$, $p_4(J_2) = 9\lambda_1 + 46.2\lambda_2$, and $p_5(J_2) = 5\lambda_1 + 38\lambda_2$ will be chosen.

(ii) In the case that the weighting value of the overall company's profit is fixed at .5 (i.e., $\lambda_1 = .5$) and the weighting value of the market share and that of the vice president for finance are any of $0 \le \lambda_2 \le 1$ and $0 \le \gamma_2 \le 1$, respectively, Fig. 29.5 shows the optimal trade-offs between the utilized production capacity of the company and the constraint level believed by the vice president for production for J_1 and J_2. In Fig. 29.5, if the values of (γ_1, λ_3) are taken from $0 \le \lambda_3 \le .5$ to $.81 \le \gamma_1 \le 1$, then the optimal transfer prices associated with J_1 are $p_1(J_1) = 2 - 4\lambda_3$, $p_2(J_1) = 6.25 + 95\lambda_2 + 5\lambda_3$, $p_3(J_1) = 1.5 + 2\lambda_2 + 2\lambda_3$, $p_4(J_1) = 2 + 8.2\lambda_2 + 5\lambda_3$, and $p_5(J_1) = 0$, respectively. If the values of (γ_1, λ_3) are taken from $0 \le \lambda_3 \le .5$ to $0 \le \gamma_1 \le .81$, then the optimal transfer prices associated with J_2 are $p_1(J_2) = 2 - 4\lambda_3$, $p_2(J_2) = 0$, $p_3(J_2) = 4 + 40\lambda_2 + 4\lambda_3$, $p_4(J_2) = 4.5 + 46.2\lambda_2 + 7\lambda_3$, and $p_5(J_2) = 2.5 + 38\lambda_2 + 2\lambda_3$, respectively. Note that when the weighting values fall in the range of $.5 \le \lambda_3 \le 1$ and $0 \le \gamma_1 \le 1$, there is not any optimal trade-off because both J_1 and J_2 are not optimal solutions (recall that this is caused by the by-product 1 in Division 1 that has -\$4 as the unit profit).

It is worth noting some important implications for accounting decision makers from the above trade-off analysis. First, the multiple factor transfer pricing model has a capability of systematically locating all possible optimal transfer prices through the optimal trade-offs of multiple objectives and multiple-constraint levels. Since the set of all possible optimal transfer prices found by this model describes every possible decision situation within the model framework, the optimal transfer price obtained from either linear programming or goal programming models is included in the subset as a special case.

Second, the multiple factor transfer pricing model can be applied to solve transfer pricing problems not only with complex structure but also with some organizational behavior contents, such as the organizational differentiation (see Table 29.3 for different constraint levels). Consequently, this model can foster more autonomous flexibility than any other mathematical programming model by allowing central management or local managers to express their own preference structures and weights.

Third, the proposed method facilitates decision makers' participation that may make a positive management achievement of organizational goals (Locke et al. 1981). The model can aid coordination and synthesis of multiple conflicting views. This may be quite effective in a transfer pricing situation in which many objectives are contradictory to each other and these objectives are measured differently by a number of decision participants.

Fourth, in most multiple-criteria solution techniques, including the goal programming approach, if the decision makers are not satisfied with the optimal solution obtained by using their preferred weights of importance for the multiple objectives, then an iterative process has to be conducted for incorporating the decision makers' new preference on the weights and finding new solutions until decision makers are satisfied. However, the proposed model eliminates such a time- and cost-consuming iterative process since it already considers all possible optimal solutions with respect to the changes of parameter (γ, λ). Whenever decision makers want to change their preference on the weights, the corresponding optimal transfer prices can be immediately identified from the results like in Table 29.6. This model, in turn, allows the performance evaluation of optimal transfer prices.

Finally, the optimal transfer prices obtained by the multiple factor model have a twofold significance in terms of decision characteristics. If the problem is viewed as a deterministic decision problem, whenever the preferred weighting values of objectives and constraint levels are known, the resulting optimal solutions can be identified from the potentially optimal solutions of the model. Then, the corresponding optimal transfer prices can be adopted to handle the business situation (recall the above trade-off analysis). If the problem is viewed as a probabilistic decision problem, it involves the assessment of the likelihood of (γ, λ) to occur at the various points of the range. With proper assumptions, the uncertainty may be represented by random variables with some known probability distribution. A number of known criteria such as maximizing expected payoff, minimizing the variance of the payoff, maximin payoff, maximizing the probability of achieving a targeted payoff, stochastic dominance, probability dominance, and mean-variance dominance can be used to choose the optimal transfer prices (see Shi 1991). In summary, the multiple factor transfer pricing model fosters flexibility in designing the optimal transfer prices for the corporation to cope with all possible changes of business competition. This model is more likely to be a better aid for executives or managers to understand and deal with their current or future transfer pricing problems.

29.4.5 Conclusions

A multiple factor transfer pricing model has been developed to solve the transfer pricing problems in a multidivisional corporation. This model can provide a systematic and comprehensive scenario about all possible optimal transfer prices depending on multiple-criteria and multiple-constraint levels. The trade-offs of optimal transfer prices offer a broad basis for managers of a corporation to flexibly implement the optimal transfer pricing strategy and cope with various business situations. Furthermore, this method also aids global optimization, division autonomy, and performance evaluation.

There are some research problems remaining to be explored. From a practical point of view, the framework of this model can be applied to other accounting areas such as capital budgeting, cost allocation, audit sampling objectives, and personnel planning for an audit corporation if the decision variables and formulation are expressed appropriately. From a theoretical point of view, the decomposition algorithm of linear programming (Dantzig and Wolfe 1960) can be incorporated into the MC^2-simplex method (Seiford and Yu 1979) to sharpen the multiple factor transfer pricing model's capability of solving the large-scale transfer pricing problem. Thus, this incorporation may result in the development of more effective solution procedures. How to incorporate other trade-off techniques, such as the satisficing trade-off method (Nakayama 1994), into the MC^2 framework for designing optimal transfer pricing strategies is another interesting research problem.

29.5 Capital Budgeting with Multiple Criteria and Multiple Decision Makers

Capital budgeting is not a trivial task if a firm is to maintain competitive advantages by adopting new information or manufacturing systems. A firm may implement innovative accounting systems such as activity-based costing (ABC) or balanced scorecard to generate more useful information for better economic decision making in the ever-changing business environment. ABC can provide value-adding and non-value-adding activity information about new capital investments. Investment justification in the new manufacturing environment, however, requires a comprehensive decision-making process that involves competitive analysis, overall firm strategy, and evaluation of uncertain cash flows (Howell and Schwartz 1994). The challenge here is to measure cash flows as well as intangible benefits that these new systems will bring. Furthermore, conflicts of goals, limited resources, and uncertain risk factors may complicate the capital budgeting problem (see Hillier (1963), Lee (1993), Karanovic et al. (2010) for details). These problems of conflicts of goals among decision makers and limited resources in a typical organization support the use of multiple-criteria and multiple-constraint levels (MC^2) linear programming (Seiford and Yu 1979).

While traditional techniques such as payback or accounting rate of return are used as a secondary method, discounted cash flow (DCF) methods, including net present value (NPV) and internal rate of return (IRR), are the primary quantitative methods in capital budgeting (Kim and Farragher 1981). The payback method estimates how long it will take to recover the original investment. However, this method incorporates neither the cash flows after the payback period nor the variability of those cash flows (Boardman et al. 1982). The accounting rate of return method measures a return on the original cost of the investment. Both of the above methods ignore the time value of money. DCF methods may not be adequate to evaluate new manufacturing or information systems, because of a bias in favor of short-term investments with quantifiable benefits (Mensah and Miranti 1989).

A current trend in capital budgeting methods utilizes mathematical programming and higher discount rates to incorporate higher risk factors (Pike 1983; Palliam 2005). Hillier (1963) and Huang (2008) suggested useful ways to evaluate risky investments by estimating expected values and standard deviations of net cash flows for each alternative investment. They showed that the standard deviation of cash flows is easily obtainable. With this information, a complete description of the risky investment is possible via probability distribution of the IRR, NPV, or annual cost of the proposed investment under the assumption of the net cash flows from the investment, which are normally distributed. Similarly, Turney (1990) suggested a stochastic dynamic adjustment model to incorporate greater risk premiums when significant additional funds are required in multiple time periods. Lin (1993) also proposed a multiple-criteria capital budgeting model under risk. His model used chance constraints of uncertain cash flows and accounting earnings as risk factors. Pike (1988) empirically tested the correlation between sophisticated capital budgeting techniques and decision-making effectiveness and found that management believed that sophisticated investment techniques improve effectiveness in the evaluation and control of large capital projects. Weingartner (1963) introduced a mathematical programming approach in the capital budgeting problem. Other researchers have extended Weingartner's work with different directions (e.g., Baumol and Quandt 1965; Bernard 1969; Howe and Patterson 1985). The development of chance-constrained programming (CCP) by Charnes and Cooper (1961) also enriched with applications of mathematical programming models in the capital budgeting problem. They also developed an approximation solution method to the CCP with zero-one variables using a linear constraint.

A typical mathematical capital budgeting approach maximizes DCFs that measure a project's desirability on the basis of its expected net present value as a primary goal. DCF analysis, however, ignores strategic factors such as future growth opportunities (Cheng 1993). Furthermore, management can change their plans if operating conditions change. For example, they can change input and output mixes or abandon the project in a multi-period situation. The increasing involvement of stakeholders, other than shareholders, in a business organization supports a multiple-objective approach (Bhaskar 1979). Other empirical studies also found that firms used multiple criteria in their capital budgeting problems (e.g., Bhaskar and McNamee 1983; Thanassoulis 1985). The goal programming approach has been used to handle multiple-objective problems. It emphasizes weights of importance of the multiple objectives with respect to the decision maker's (DM) preference (Bhaskar 1979; Deckro et al. 1985). Within the framework of hierarchical goal optimization, several goal programming models have been suggested to unravel multiple-objective capital budgeting problems (e.g., Ignizio 1976; Lee and Lerro 1974). Similarly, Santhanam et al. (1989) used a zero-one goal programming approach for information system project selection, but their article lacks a multiple time horizon (learning curve) effect. Choi and Levary (1989) investigated the use of a chance-constrained goal programming approach to reflect multiple goals for a multinational capital budgeting problem. Reeves and Hedin (1993) suggested interactive goal programming (IGP) which allows more flexibility for DMs in considering trade-offs and adjusting goal target levels. However, because of the difficulties of measuring the preferences and priorities of decision makers, Reeves and Franz (1985) developed a simplified interactive multiple-objective linear programming (SIMOLP) method which uses an interactive procedure for a DM to identify a preferred solution and Gonzalez et al. (1987) applied this procedure in a capital budgeting problem (see Corner et al. (1993) and Reeves et al. (1988) for other examples). Interactive multiple-objective techniques such as SIMOLP or IGP can reduce the difficulty of the solution process. Such studies suggest that most capital budgeting problems require the analysis of multiple criteria that better reflect a real-world situation. Thanassoulis (1985) addressed multiple objectives such as the maximization of shareholder wealth, maximization of firm growth, minimization of financial risk, maximization of the firm liquidity, and minimization of environmental pollution.

However, these existing multiple-criteria approaches, including goal programming, implicitly assume that there is only one decision maker setting up the constraint (budget availability) level of a capital budgeting problem. This assumption is not realistic because in most real-world capital budgeting problems, such as constructing a major highway or shopping mall, multiple decision makers must involve the decision of the constraint levels (see Sect. 29.5.2 for detailed discussion). To remove this assumption of a single decision maker from models of capital budgeting with multiple criteria, this article attempts to incorporate multiple decision makers' preferences using (1) the analytic hierarchy process (AHP) approach and (2) MC^2 linear programming in a capital budgeting situation. Our approach shows its strength in quantifying strategic and nonfinancial factors that are important in the current competitive business environment. The problem of incommensurable units in the selection criteria, because of nonfinancial and qualitative measures, can be resolved by using the AHP approach. The MC^2 approach fosters modeling flexibility by incorporating decision makers' preferences as multiple-constraint levels.

The rest of the article is organized as follows. Sect. 29.2 introduces the AHP and MC^2 framework. Sect. 29.3 demonstrates the managerial significance and implications of capital budgeting problems by illustrating an example. Sect. 29.4 concludes the article with several future research avenues.

29.5.1 AHP and MC^2 Framework

In general, an organization has limited resources. Furthermore, each manager's risk assessments and preferences about a new project may be different. For example, a financial manager may think that his or her company has only 50 million dollars available for a new project and he will not approve this project unless he is sure about substantial financial benefits. In contrast, a production manager may think his company should have at least 60 million dollars to implement just-in-time (TIT) manufacturing. The production manager strongly believes that this new system will produce high-quality products with lower costs to maintain competitive advantages against competitors. It is clear that there is a conflict of interests between managers. Even top management's goal may be different from other managers'. To increase other managers' involvement and motivation, top management should induce other managers' inputs. There must be trade-offs between goals of different managers in this group decision-making process. AHP and MC^2 linear programming can be used to resolve this type of group dilemma.

29.5.1.1 Analytical Hierarchy Process (AHP)

AHP is a practical measurement technique that has been widely applied in modeling the human judgment process (Saaty 1980). AHP enables decision makers to structure a complex problem in the form of a hierarchy of its elements according to an organization's structure or ranks of management levels. It captures managerial decision preferences through a series of comparisons of relevant criteria. This feature of the AHP minimizes the risk of inconsistent decisions due to incommensurable units in the selection criteria. Recently, AHP has been applied to several accounting problems (e.g., capital budgeting (Liberatore et al. 1992), real estate investment (Kamath and Khaksari 1991), and municipal government capital investment (Chan 2004)).

The preference of each manager may be analyzed through the use of AHP or multiple attribute utility technique (MAUT). Both AHP and MAUT have their own strengths and weaknesses. For a recent debate regarding the two methods, readers are referred to Dyer and Forman (1991). In this article, the AHP is employed mainly due to its capability of reducing computational complexity and availability of software.

29.5.1.2 MC^2 Linear Programming

Linear programming has been applied to capital budgeting problems such as maximizing NPV with a single objective. However, the linear programming approach has limitations in handling multiple conflicting real-world goals. For instance, linear programming cannot solve the problem in which a firm ties to maximize overall profits and total market share simultaneously. This dilemma is overcome by a technique called multiple-criteria (MC) linear programming (Goicoechea et al. 1982; Steuer 1986; Yu 1985; Zeleny 1974). To extend the framework of MC linear programming, Sieford and Yu (1979) and Yu (1985) formulated a model of MC^2 linear programming. This model is based on two premises. First, from the linear system structure's perspective, the criteria and constraint levels may be interchangeable. Thus, like multiple criteria, multiple-constraint (resource availability) levels can be considered. Second, from the application's perspective, it is more realistic to consider multiple resource availability levels (discrete right-hand sides) than a single resource availability level in isolation. This recognizes that the availability of resources can fluctuate depending on the decision situation forces, such as the preferences of the different managers. The concept of multiple resource levels corresponds to the typical characteristic of capital budgeting situations where the decision-making process should reflect each manager's preference on the new project. A theoretical connection between MC linear programming and MC^2 linear programming can be found in Gyetvan and Shi (1992) and decision problems related to MC^2 linear programming have been extensively studied (see, e.g., Lee et al. (1990), Shi (1991), and Shi and Yu (1992)). Key ideas of MC^2 linear programming, as a primary theoretical foundation of this article, are described as follows.

An MC^2 linear programming problem can be formulated as

$$\max \lambda^t C_x$$

$$\text{s.t. } A_x \leq D\gamma$$

$$x \geq 0$$

where $C \in R^{qxn}$, $A \in R^{mxn}$, and $D \in R^{mxp}$ are matrices of qxn, mxn, and mxp dimensions, respectively; $x \in R^n$ are decision variables; $\lambda \in R^q$ is called the criteria parameter; and $\gamma \in R^p$ is called the constraint level parameter. Both (γ, λ) are assumed unknown.

The above MC^2 problem has q criteria (objectives) and p constraint levels. If the constraint level parameter γ is known, then the MC^2 problem reduces to an MC linear programming problem (e.g., Yu and Zeleny 1975). In addition, if the criteria parameter λ is known, it reduces to a linear programming problem (e.g., Charnes and Cooper 1961; Dantzig 1963).

Denote the index set of the basic variables $\{x_{j1}, \ldots, x_{jm}\}$ for the MC^2 problem by $J = \{j_1, \ldots, j_m\}$. Note that the basic variables may contain some slack variables. Without confusion, J is also called a basis for the MC^2 problem. Since a basic solution J depends on parameters (γ, λ), define that (1) a basic solution J is feasible for the MC^2 problem if and only if there exists a $\gamma^0 > 0$ such that J is a feasible solution for the MC^2 problem with respect to γ^0 and (2) J is potentially optimal for the MC^2 problem if and only if there exist a $\gamma^0 > 0$ and a $\lambda^0 > 0$ such that J is an optimal solution for the MC^2 problem with respect to (γ^0, λ^0). For an MC^2 problem, there may exist a number of potentially optimal solutions $\{J\}$ as parameters (γ, λ) vary depending on decision situations. Seiford and Yu (1979) derived a simplex method to systematically locate the set of all potentially optimal solutions $\{J\}$.

In summary, a model within the framework of AHP and MC^2 is proposed to formulate and solve the multiple-objective capital budgeting problems with multiple decision makers as follows. Note that because of complexity of the problem, we employed AHP to derive weights of λ and γ. This solution procedure decreases computational complexity.

29.5.1.3 A Model of MC^2 Decision-Making Capital Budgeting

The linear programming approach has limitations for solving real-world problems with multiple objectives. The goal programming or MOLP approach provides only one optimal trade-off among several objectives of the firm. None of the past mathematical models supports multiple decision makers in a capital budgeting decision-making process. As discussed before, in a real-world situation, the decision makers (DMs) can have different opinions on the same issue. These different interpretations can be represented by different constraint levels in mathematical models. Because both linear programming and goal programming presume a fixed single constraint level, they fail to mathematically describe such an organizational differentiation problem.

The MC^2 linear programming framework can resolve the above shortcomings in current capital budgeting models. The framework is flexible enough to include any objectives depending on problem situation. However, for the sake of clear presentation, we address four groups of objectives that are common in capital budgeting: (1) maximization of net present value, (2) profitability, (3) growth, and (4) flexibility of financing. Some group objectives may have multi-time periods. Net present value measures the expected net monetary gain or loss from a project by discounting all expected future cash flows to the present point in time, using the desired rate of return. If we want to incorporate risk factors, we can estimate the standard deviation of net cash flows as Hillier (1963) suggested. Profitability can be measured by return on investment (ROI). Growth can be measured by sales growth or a nonfinancial measure of market share growth, each of which focuses on the long-term success of a firm. Flexibility of financing (leverage) can be measured by the debt to equity ratio. As a firm's debt to equity ratio goes up, the firm's cost of borrowing becomes more expensive. Even though our model contains only these four specific objective groups, the flexibility of the modeling process fits in all capital budgeting problems with multiple-criteria and multiple-constraint levels.

This model integrates AHP with a mathematical model. The strengths of such an integration have been shown in several areas (e.g., advertisement media choice (Dyer and Forman 1991), R&D portfolio selection (Suh et al. 1993), and telecommunication hub design (Lee et al. 1995)). The use of both AHP and MC^2, as we will explore in this article, is the first in the decision-making literature.

In general, let i be the time period for a firm to consider capital budgeting and j be the index number of the projects that the firm can select.

Define x_j as the jth project that can be selected by the firm, $j = 1, \ldots, t$. For the coefficients of the objectives, let v_j be the net present value generated by the jth project; g_{ij} be the net sales increase generated by the jth project in the ith time period to measure sales growth, $i = 1, \ldots,$ s; r_{ij} be the ROI generated from the jth project in the ith time period; and l_{ij} be the firm's equity to debt ratio in the ith time period after jth project is selected to measure liquidity. Here, the optimum debt to equity ratio for the firm is assumed to be 1 and the inverse of debt to equity ratio is used to measure leverage. For the coefficients of the constraints, let b_{ij} be the cash outlay required by the jth project in the ith time period. For the coefficients of the constraint levels, let C_{ki} be the budget availability level believed by the kth manager (or executive) for the firm in the ith time period, $k = 1, \ldots, u$.

In this article, for illustration, if we treat $t = 10$, $i = 5$, and $u = 5$, then the model is

$$\max \sum_{j=1}^{10} v_j X_j$$

$$\max \sum_{j=1}^{10} g_{ij} X_j \ \forall i = 1, \ldots, 5$$

$$\max \sum_{j=1}^{10} r_{ij} X_j \ \forall i = 1, \ldots, 5$$

$$\max \sum_{j=1}^{10} l_{ij} X_j \ \forall i = 1, \ldots, 5$$

$$\text{subject to} \sum_{j=1}^{10} b_{ij} X_j \le (c_{li}, \ldots, c_{5i}) \ \forall i = 1, \ldots, 5$$

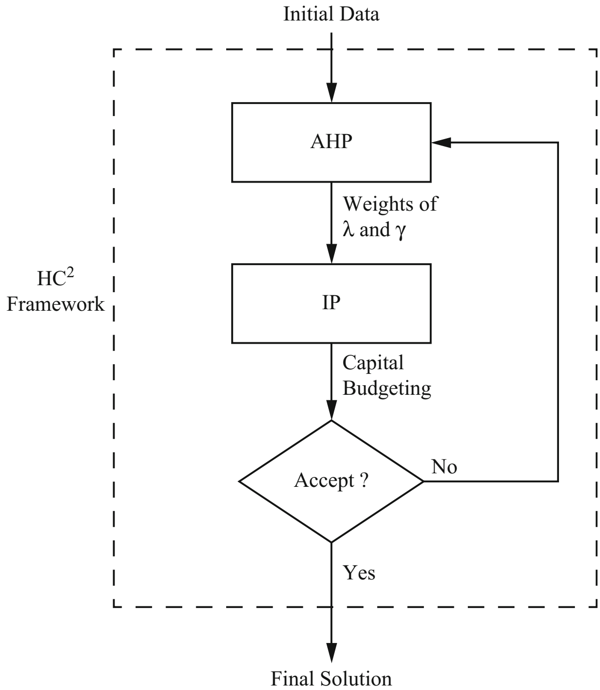
$$\text{and } X_j = \{0, 1\}$$

where $k = 1, \ldots, 5$, for c_{ki}, represents the five possible DMs (i.e., president, controller, production manager, marketing manager, and engineering manager). The weights of 16 total objectives can be expressed by $(\lambda_1, \ldots, \lambda_{16})$ with $\Sigma\lambda_q = 1$, $q = 1, \ldots, 16$, $0 < \lambda_q < 1$, and the weights of five DMs can be expressed by $(\lambda_1, \ldots, \lambda_5)$ with $\Sigma\gamma_k = 1$, $k = 1, \ldots, 5$, $0 < \gamma_k < 1$.

The above model is an integer MC^2 problem with 16 objectives and five constraint levels. Solving this problem by using a currently available solution technique (Seiford and Yu 1979) is not a trivial task because of its substantial computational complexity. To overcome this difficulty, we propose a two-phased solution procedure as depicted in Fig. 29.6.

The first phase is referred to as the AHP phase in the sense that AHP is applied to derive weights of λ and γ, which reduces the model's complexity. The computation of weights is effectively handled by Expert Choice (Forman et al. 1985), commercial software for AHP. In this phase, we first induce preferences about objectives from all five DMs involved. Note that each DM may have different preferences about the four goals. The president may think maximization of ROI is the most important goal, while the controller may think maximization of NPV is the most important goal. AHP generates relative weights for each goal. Then, multiple-constraint levels are incorporated using the MC^2 framework. For instance, DMs may estimate different budget availability levels for each time period. The controller may believe that \$40 million is available for the first year, while the production manager may believe that the company can spend \$50 million for this time period. This scenario is more realistic if we consider the characteristics of

Fig. 29.6 A two-phased capital budgeting model



a capital budgeting problem from the perspective of the different utility functions or risk factors of each DM. An example using AHP to generate relative weights for the five DMs' constraint levels (preferences) can be found in Appendix 1.

The second phase is called the integer programming (IP) phase. After applying the AHP phase to the model, the model is reduced to a linear IP problem, in which the potentially optimal projects must be selected. To solve this problem, we employ the ZOOM software (Singhal et al. 1989) that was originally introduced for solving zero-one integer linear programming problems.

29.5.2 Model Implications

29.5.2.1 Numerical Example

A prototype of our capital budgeting model demonstrates the implications for multiple criteria and multiple decision makers. We use a Lorie-Savage (1955) type of problem as follows. The Lorie-Savage Corporation has ten projects under consideration. All projects have 5-year tune periods.

The executives (president, controller, production manager, marketing manager, and engineering manager) agree on the fallowing multiple objectives:

1. Maximize net present value of each project. Net present value is computed as the sum of all the discounted, estimated future cash flows, using the desired rate of return, minus the initial investment.

Table 29.7 Net present value data for the Lorie-Savage Corporation (in millions)

		Cash outlays for each period				
Project	Net present value	Period 1	Period 2	Period 3	Period 4	Period 5
1	$20	$24	$16	$6	$6	$8
2	16	12	10	2	5	4
3	11	8	6	6	6	4
4	4	6	4	7	5	3
5	4	1	6	9	2	3
6	18	18	18	20	15	15
7	7	13	8	10	8	8
8	19	14	8	12	10	8
9	24	16	20	24	16	16
10	4	4	6	8	6	4

2. Maximize the growth of the firm due to each project. Growth can be measured by net sales increase from each project. This can measure a strategic success factor of a firm.
3. Maximize the profitability of the company. Profitability can be measured by ROI for each time period generated by each project.
4. Maximize the flexibility of financing. The flexibility of financing (leverage) can be measured by debt to equity ratio. If a firm's debt to equity ratio *is* higher than the industry average or the optimum level, the firm's cost of debt financing will become more expensive. In this example, we use the inverse of the debt to equity ratio and assume 1 as the optimum debt to equity ratio.

The data related to these objectives is given in Table 29.7. In this example, all projects are assumed to be independent. In Table 29.7, the firm's cost of capital is assumed to be known a priori and to be independent of the investment decisions. Based on these assumptions, the net present value of each project can be defined as the sum of the cash flows discounted by the cost of capital. Cash outlay is the amount of expenditure required for project $j, j = 1, 2, 3, \ldots, 10$, in each time period.

To measure growth, the net sales increase for each time period for each project is estimated. These data are provided in Table 29.8.

To measure the profitability of each project, ROI is estimated after reflecting additional income and capital expenditures from each investment for each time period. These data are provided in Table 29.9.

To measure leverage, the inverse of debt to equity ratio is used after adopting each project. Here, the optimum debt to equity ratio is assumed to be one for the Lorie-Savage Corporation. These data are provided in Table 29.10.

The five key DMs in this company (president, controller, production manager, marketing manager, and engineering manager) have different beliefs regarding resource availability. For example, for budget availability levels, each DM may have a different opinion. Of course, the president will make the final decision based on the opinions of other managers. However, the DMs' preferences of collection process should improve the quality of the final decision. Budget availability level data are provided in Table 29.11.

Table 29.8 Net sales increase data for the Lorie-Savage Corporation (in millions)

	Net sales increase for each project				
Project	Period 1	Period 2	Period 3	Period 4	Period 5
1	$120	$130	$145	$150	$170
2	100	120	140	150	160
3	80	90	95	95	100
4	40	50	50	55	60
5	40	45	50	55	60
6	110	120	140	150	165
7	60	70	65	70	80
8	110	120	100	110	120
9	150	170	180	190	200
10	35	40	40	50	50

Table 29.9 Return on investment data for the Lorie-Savage Corporation (in percentage)

	Return on investment for each project				
Project	Period 1	Period 2	Period 3	Period 4	Period 5
1	10	12	14	15	17
2	10	12	18	16	17
3	12	15	15	15	18
4	8	15	10	8	12
5	15	10	8	20	18
6	12	12	10	15	15
7	8	12	10	12	12
8	14	16	13	15	16
9	12	10	9	12	12
10	10	8	9	8	12

Table 29.10 Debt to equity data for the Lorie-Savage Corporation

	Inverse of debt to equity ratio				
Project	Period 1	Period 2	Period 3	Period 4	Period 5
1	0.85	0.90	0.98	0.95	0.95
2	0.90	0.98	0.95	0.95	0.96
3	0.96	0.97	0.98	0.92	0.92
4	0.98	0.95	0.96	0.92	0.95
5	0.90	0.95	0.95	0.98	0.95
6	0.90	0.95	0.94	0.95	0.95
7	0.96	0.98	0.98	0.98	0.98
8	0.96	0.95	0.90	0.92	0.95
9	0.90	0.88	0.85	0.95	0.95
10	0.98	0.95	0.95	0.98	0.95

Table 29.11 Budget availability level data for the Lorie-Savage Corporation (in millions)

	Estimate budget availability				
Decision maker	Period 1	Period 2	Period 3	Period 4	Period 5
President	\$50	\$30	\$30	\$35	\$30
Controller	40	45	30	30	20
Production manager	55	40	20	30	35
Marketing manager	45	30	40	45	30
Engineering	50	40	45	30	35

The parameters for budget availability levels are derived by using AHP. Here, one level of an AHP is used. The aggregated objective function also can be obtained by using AHP. The difference here is that two levels of an AHP must be used to provide preferences for the objective functions. The first level of this AHP corresponds to the four groups of objectives. The second level corresponds to time periods within each objective group. Note that the second level is not required for the objective of maximizing the net present value. Hence, four pairwise comparison matrices are required for eigenvalue computation. More detailed information about the formulation of this problem is presented in the appendix. In sum, the objective function aggregated from the 16 objectives is written as

$$\begin{aligned}\text{Maximize} \quad & 56.34x_1 + 51.74x_2 + 40.54x_3 + 25.43x_4 + 25.44x_5 \\ & + 53.63x_6 + 31.98x_7 + 50.59x_8 + 67.62x_9 + 23.17x_{10}.\end{aligned}$$

The optimum IP solution for this numerical example is selecting projects 2, 3, 4, and 8. This solution reflects the preferences of multiple DMs and satisfies budget constraints. If the DMs are not satisfied with the solution, we have to use AHP to assess and compute new values of weights γ and λ according to Fig. 29.6. This process will terminate whenever all DMs agree to the compromise solution.

29.5.2.2 Managerial Implications

The model and solution procedure proposed in this study have several managerial implications. First, the model integrates the multiple-objective capital budgeting problem with multiple decision makers. The most common problem in a capital investment decision-making situation is how to estimate cash flows and to incorporate the risk factors of decision makers (Hillier 1963; Lee 1993). In our model, we allow multiple-constraint levels to incorporate each DM's preference about budget availability. Neither linear programming nor goal programming models can handle these characteristics.

Second, our model can be applied to solve capital budgeting problems not only with the representation of a complex structure but also with motivational implications. Our method facilitates DMs' participation in order to minimize suboptimization of overall company goals (Locke et al. 1981). The model can aid coordination and synthesis of multiple objectives, some in conflict. This feature can be quite effective in a capital budgeting situation, in which many objectives are contradictory to each other and these objectives can be measured differently by a number of decision participants. This model can foster more autonomous control than any other mathematical programming models by allowing DMs to express their own priority structures and weights.

Third, adoption of the AHP reduces the solution complexity of the resulting MC^2 IP program. The MC^2 IP is a very complex problem to solve. Real-life size MC^2 IP problems are intractable by using current computational technology. Furthermore, the computing software for MC^2 IP is still under development (Shi and Lee 1992). Even though the use of AHP may lose some possible trade-offs among objectives and/or DMs' preferences, the AHP reduces MC^2 IP to a traditional zero-one IP that is much easier to be solved using currently available IF software.

Lastly, our two-phased framework (Fig. 29.6) for capital budgeting problem is highly flexible. The DMs can reach an agreement interactively. For example, alternative budgeting policy can be obtained by providing a different preference matrix of objectives and/or resource availability levels. Furthermore, the two-phased framework may be attempted iteratively until all of the DMs are satisfied with the final budgeting policy.

29.5.3 Conclusions

A decision-making process with a multiple-criteria model has been addressed to solve capital budgeting problems. This model can foster a high-quality decision in capital budgeting under multiple criteria and multiple decision makers. This decision-making strategy reflects each decision maker's preference and limits suboptimization of overall company goals. This method can also better handle real-world problems that may include uncertain factors. By incorporating information from each influential manager, our model is more likely to provide better budgeting solutions than the previous linear programming or goal programming approaches.

There are other research problems remaining to be explored. From a practical point of view, the estimation of future cash flows, determining a firm's cost of capital, measuring intangible benefits, and measuring residual value of assets are still important issues. Models like ours can reduce the chance of an ineffective decision making by incorporating multiple decision makers' preferences. This framework can be applied to other accounting problems, such as cost allocation, audit sampling objectives, and personnel planning for an audit, if the decision variables and formulation are expressed appropriately.

29.6 Conclusions

We have introduced group decision-making tools that can be applied in accounting and finance. By the nature of today's dynamic business environment, there will be more than one decision maker and business conditions keep changing. We showed

AHP and MC^2 applications in performance evaluation, banking performance evaluation, international transfer pricing, and capital budgeting. Our last paper shows the combination of AHP and MC^2 in capital budgeting to deal with multiple objectives and multiple constraints.

Our performance evaluation model shows advantages such as flexibility, feedbacks, group evaluation, and computing simplicity. A prototype was built via a personal computer so that the model can be applied to any business situations. The iterative process of getting input data in the AHP procedure helps each manager as well as employee to be aware of the importance of strategic factors of each performance measure of the bank from our second paper. Our transfer pricing model can provide a systematic and comprehensive scenario about all possible optimal transfer prices depending on multiple-criteria and multiple-constraint levels. The trade-offs of optimal transfer prices offer a broad basis for managers of a corporation to flexibly implement the optimal transfer pricing strategy and cope with various business situations. Furthermore, this method also aids global optimization, division autonomy, and performance evaluation. Our last paper shows that our capital budgeting model is more likely to provide better budgeting solutions than the previous approaches by incorporating information from each influential manager.

Appendix 1

For illustrative purposes, we show how to use AHP to induce the five DMs' preferences of budget availability and to compute the relative weights. Generally, AHP collects input judgments of DMs in the form of a matrix by pairwise comparisons of criteria (i.e., their budget availability levels). An eigenvalue method is then used to scale weights of such criteria. That is, the relative importance of each criteria is computed. The result from all of pairwise comparison is stored in an input matrix as follows:

$$\begin{array}{c} \begin{array}{ccccc} \text{President} & \text{Controller} & \begin{array}{c}\text{Production}\\ \text{Manager}\end{array} & \begin{array}{c}\text{Marketing}\\ \text{Manager}\end{array} & \begin{array}{c}\text{Engineering}\\ \text{Manager}\end{array} \end{array} \\ \begin{bmatrix} 1 & 3 & 4 & 5 & 6 \\ & 1 & 2 & 5 & 5 \\ & & 1 & 3 & 4 \\ & & & 1 & 2 \\ & & & & 1 \end{bmatrix} \end{array}$$

Applying an eigenvalue method to the above input matrix results in a vector $W_i = (0.477, 0.251, 0.154, 0.070, 0.048)$. In addition to the vector, the inconsistency ratio (γ) is obtained to estimate the degree of inconsistency in pairwise comparisons. In this example, the inconsistency ratio is 0.047. A common guideline is that if the ratio surpasses 0.1, a new input matrix must be generated. Therefore, this input matrix is acceptable.

A similar computing process can be applied for the 16 objective functions. However, two hierarchical levels are required for this case. The first level of AHP corresponds to the four groups of objectives and the second level corresponds to the time periods within each objective group.

Let x_j be the jth project that can be selected by the firm. Using data in Tables 29.7, 29.8, 29.9, and 29.10, the model for capital budgeting with multiple criteria and multiple DMs is formulated as

$$\text{maximize} \quad 20x_1 + 16x_2 + 11x_3 + 4x_4 + 4x_5 + 18x_6 + 7x_7 + 9x_8 + 24x_9 + 4x_{10}$$

$$\text{maximize} \quad 120x_1 + 100x_2 + 80x_3 + 40x_4 + 40x_5 + 110x_6 + 60x_7 + 110x_8 + 150x_9 + 35x_{10}$$

$$\text{maximize} \quad 130x_1 + 120x_2 + 90x_3 + 50x_4 + 45x_5 + 120x_6 + 70x_7 + 120x_8 + 170x_9 + 40x_{10}$$

$$\text{maximize} \quad 145x_1 + 140x_2 + 95x_3 + 50x_4 + 50x_5 + 140x_6 + 65x_7 + 100x_8 + 180x_9 + 40x_{10}$$

$$\text{maximize} \quad 150x_1 + 150x_2 + 95x_3 + 55x_4 + 55x_5 + 150x_6 + 70x_7 + 110x_8 + 190x_9 + 50x_{10}$$

$$\text{maximize} \quad 170x_1 + 160x_2 + 100x_3 + 60x_4 + 60x_5 + 165x_6 + 80x_7 + 120x_8 + 200x_9 + 50x_{10}$$

$$\text{maximize} \quad 10x_1 + 10x_2 + 12x_3 + 8x_4 + 15x_5 + 12x_6 + 8x_7 + 14x_8 + 12x_9 + 10x_{10}$$

$$\text{maximize} \quad 12x_1 + 12x_2 + 15x_3 + 15x_4 + 10x_5 + 12x_6 + 12x_7 + 16x_8 + 10x_9 + 8x_{10}$$

$$\text{maximize} \quad 14x_1 + 18x_2 + 15x_3 + 10x_4 + 8x_5 + 10x_6 + 10x_7 + 13x_8 + 9x_9 + 9x_{10}$$

$$\text{maximize} \quad 15x_1 + 16x_2 + 15x_3 + 8x_4 + 20x_5 + 15x_6 + 12x_7 + 15x_8 + 12x_9 + 8x_{10}$$

$$\text{maximize} \quad 17x_1 + 17x_2 + 8x_3 + 12x_4 + 18x_5 + 15x_6 + 12x_7 + 16x_8 + 12x_9 + 12x_{10}$$

$$\text{maximize} \quad 0.85x_1 + 0.90x_2 + 0.96x_3 + 0.98x_4 + 0.90x_5 + 0.90x_6 + 0.96x_7 + 0.96x_8 + 0.90x_9 + 0.98x_{10}$$

$$\begin{aligned}
\text{maximize}\quad & 0.90x_1 + 0.98x_2 + 0.97x_3 + 0.95x_4 + 0.95x_5 \\
& + 0.95x_6 + 0.98x_7 + 0.95x_8 + 0.88x_9 + 0.95x_{10} \\[2ex]
\text{maximize}\quad & 0.98x_1 + 0.95x_2 + 0.98x_3 + 0.96x_4 + 0.95x_5 \\
& + 0.94x_6 + 0.98x_7 + 0.90x_8 + 0.85x_9 + 0.95x_{10} \\[2ex]
\text{maximize}\quad & 0.95x_1 + 0.95x_2 + 0.92x_3 + 0.92x_4 + 0.98x_5 \\
& + 0.95x_6 + 0.98x_7 + 0.92x_8 + 0.95x_9 + 0.98x_{10} \\[2ex]
\text{maximize}\quad & 0.95x_1 + 0.96x_2 + 0.92x_3 + 0.95x_4 + 0.95x_5 \\
& + 0.95x_6 + 0.98x_7 + 0.95x_8 + 0.95x_9 + 0.95x_{10} \\[2ex]
\text{subject to}\quad & 24x_1 + 12x_2 + 8x_3 + 6x_4 + x_5 \\
& + 18x_6 + 13x_7 + 14x_8 + 16x_9 + 4x_{10} \leq 49.37 \\[2ex]
& 16x_1 + 10x_2 + 6x_3 + 4x_4 + 6x_5 \\
& + 18x_6 + 8x_7 + 8x_8 + 20x_9 + 6x_{10} \leq 35.30 \\[2ex]
& 6x_1 + 2x_2 + 6x_3 + 7x_4 + 9x_5 \\
& + 20x_6 + 10x_7 + 12x_8 + 24x_9 + 8x_{10} \leq 28.91 \\[2ex]
& 6x_1 + 5x_2 + 6x_3 + 5x_4 + 2x_5 \\
& + 15x_6 + 8x_7 + 10x_8 + 16x_9 + 6x_{10} \leq 33.44 \\[2ex]
& 8x_1 + 4x_2 + 4x_3 + 3x_4 + 3x_5 \\
& + 15x_6 + 8x_7 + 8x_8 + 16x_9 + 4x_{10} \leq 29.96 \\[2ex]
& \text{and } X_j = \{0, 1\} \text{ for } j = 1, \ldots, 10.
\end{aligned}$$

References

Abdel-khalik, A. R., & Lusk, E. J. (1974). Transfer pricing – A synthesis. *The Accounting Review*, 8–24.

Arrington, E. E., Hillison, W., & Jensen, R. E. (1984). An application of analytical hierarchy process to model expert judgment on analytic review procedures. *Journal of Accounting Research, 22*, 298–312.

Ayvaz, E., & Pehlivanli, D. (2011). The use of time driven activity based costing and analytic hierarchy process method in the balanced scorecard implementation. *International Journal of Business and Management, 6*, 146–158.

Bailey, A. D., & Boe, W. J. (1976). Goal and resource transfers in the multigoal organization. *The Accounting Review*, 559–573.

Bapna, R., Goes, P., & Gupta, A. (2005). Pricing and allocation for quality-differentiated online services. *Management Science, 51*, 1141–1150.

Baumol, W. J., & Fabian, T. (1964). Decomposition, pricing for decentralization and external economies. *Management Science, 11*, 1–32.

Baumol, W. J., & Quandt, R. E. (1965). Investment and discount rate under capital rational – A programming approach. *Economic Journal, 75*(298), 173–329.

Bernard, R. H. (1969). Mathematical programming models for capital budgeting – A survey generalization and critique. *Journal of Financial and Quantitative Analysis, 4*, 111–158.

Bhaskar, K. A. (1979). Multiple objective approach to capital budgeting. *Accounting and Business Research, 10*, 25–46.

Bhaskar, I. C., & McNamee, P. (1983). Multiple objectives accounting and finance. *Journal of Business Finance and Accounting, 10*, 595–621.

Boardman, W., Reinhart, I., & Celec, S. E. (1982). The role of payback period in the theory and application of duration to capital budgeting. *Journal of Business Finance and Accounting, 9*, 511–522.

Borkowski, S. C. (1990). Environmental and organizational factors affecting transfer pricing: A survey. *Journal of Management Accounting Research, 2*, 78–99.

Boucher, T. O., & MacStravic, E. L. (1991). Multi-attribute evaluation within a present value framework and its relation to the analytic hierarchy process. *The Engineering Economist, 37*, 1–32.

Calvert, J. (1990). Quality customer service: A strategic marketing weapon for high performance banking. *Industrial Engineering*, 54–57.

Campi, J. P. (1992). It's not as easy as ABC. *Journal of Cost Management Summer, 6*, 5–11.

Chan, Y. L. (2004). Use of capital budgeting techniques and an analytic approach to capital investment decisions in Canadian municipal governments. *Public Budgeting & Finance, 24*, 40–58.

Chan, Y. L., & Lynn, B. E. (1991). Performance evaluation and the analytic hierarchy process. *Journal of Management Accounting Research, 3*, 57–87.

Charnes, A., & Cooper, W. W. (1961). *Management models and industrial applications of linear programming* (Vols. I & II). New York: Wiley.

Cheng, J. K. (1993). Managerial flexibility in capital investment decision: Insights from the real-options. *Journal of Accounting Literature, 12*, 29–66.

Chien, I. S., Shi, Y., & Yu, P. L. (1989). *MC^2 program: A Pascal program run on PC or Vax (revised version)*. Lawrence: School of Business, University of Kansas.

Choi, T. S., & Levary, R. R. (1989). Multi-national capital budgeting using chance-constrained goal programming. *International Journal of Systems Science, 20*, 395–414.

Corner, I. L., Decicen, R. F., & Spahr, R. W. (1993). Multiple-objective linear programming in capital budgeting. *Advances in Mathematical Programming and Financial Planning, 3*, 241–264.

Curtis, S. L. (2010). Intercompany licensing rates: Implications from principal-agent theory. *International Tax Journal, 36*, 25–46.

Dantzig, G. B. (1963). *Linear programming and extensions*. New Jersey: Princeton University Press.

Dantzig, G. B., & Wolfe, P. (1960). Decomposition principles for linear programs. *Operations Research, 8*, 101–111.

Deckro, R. F., Spahr, R. W., & Herbert, J. E. (1985). Preference trade-offs in capital budgeting decisions. *IIE Transactions, 17*, 332–337.

DeWayne, L. S. (2009). Developing a lean performance score. *Strategic Finance, 91*, 34–39.

Dopuch, N., & Drake, D. F. (1964). Accounting implications of a mathematical programming approach to the transfer price problem. *Journal of Accounting Research, 2*, 10–24.

Drucker, P. F. (1993). We need to measure, not count. *Wall Street Journal*, April 13. Also *Expert Choice Voice*, July.

Dyer, R. F., & Forman, E. H. (1991). *An analytic approach to marketing decisions*. Englewood, NJ: Prentice Hall.

Eccles, R. G. (1983). Control with fairness in transfer pricing. *Harvard Business Review, 61*, 149–161.

Eccles, R. G. (1991). The performance measurement manifesto. *Harvard Business Review, 69*, 131–137.

Eccles, R. G., & Pyburn, P. J. (1992). Creating a comprehensive system to measure performance. *Management Accounting, 74*, 41–44.

Eddie, W., Cheng, L., & Li, H. (2001). Analytic hierarchy process: An approach to determine measures for business performance. *Measuring Business Excellence, 5*, 30–36.

Fisher, J. (1992). Use of nonfinancial performance measures. *Journal of Cost Management, 6*, 31–38.

Forman, E. H., Saaty, T. L., Selly, M. A., & Waldron, R. (1985). *Expert choice*. Pittsburgh: Decision Support Software.

Goicoechea, A., Hansen, D. R., & Duckstein, L. (1982). *Multi-objective decision analysis applications*. NewYork: Wiley.

Gonzalez, G., Reeves, R., & Franz, L. S. (1987). Capital budgeting decision making: An interactive multiple objective linear integer programming search procedure. *Advances in Mathematical Programming and Financial Planning, 1*, 21–44.

Gould, J. R. (1964). Internal pricing in corporations when there are costs of using an outside market. *The Journal of Business, 37*, 61–67.

Gyetvan, F., & Shi, Y. (1992). Weak duality theorem and complementary slackness theorem for linear matrix programming problems. *Operations Research Letters, 11*, 244–252.

Hardy, C., & Reeve, R. (2000). A study of the internal control structure for electronic data interchange systems using the analytic hierarchy process. *Accounting and Finance, 40*, 191–210.

Harker, P. T., & Vargas, L. G. (1987). The theory of ratio scale estimation: Saaty's analytic hierarchy process. *Management Science, 33*, 1383–1403.

Hillier, F. S. (1963). The derivation of probabilities information for the evaluation of risky investments. *Management Science*, 443–457.

Hirshleifer, J. (1956). On the economies of transfer pricing. *The Journal of Business*, 172–184.

Howe, K. M., & Patterson, I. H. (1985). Capital investment decisions under economies of scale in flotation costs. *Financial Management, 14*, 61–69.

Howell, R. A., & Schwartz, W. A. (1994). Asset deployment and investment justification, In *Handbook of cost management*. Boston: Warren Gorham Lamont.

Huang, X. (2008). Mean-variance model for fuzzy capital budgeting. *Computers & Industrial Engineering, 55*, 34.

Huang, S., Hung, W., Yen, D. C., Chang, I., & Jiang, D. (2011). Building the evaluation model of the IT general control for CPAs under enterprise risk management. *Decision Support Systems, 50*, 692.

Ignizio, J. P. (1976). An approach to the capital budgeting problem with multiple objectives. *The Engineering Economist, 21*, 259–272.

Ishizaka, A., Balkenborg, D., & Kaplan, T. (2011). Influence of aggregation and measurement scale on ranking a compromise alternative in AHP. *The Journal of the Operational Research Society, 62*, 700–711.

Kamath, R. R., & Khaksari, S. Z. (1991). Real estate investment: An application of the analytic hierarchy process. *The Journal of Financial and Strategic Decisions, 4*, 73–100.

Kaplan, R. S. (1986). Accounting lag: The obsolescence of cost accounting systems. *California Management Review, 28*, 175–199.

Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard-measures that drive performance. *Harvard Business Review, 69*, 71–79.

Karanovic, G., Baresa, S., & Bogdan, S. (2010). Techniques for managing projects risk in capital budgeting process. *UTMS Journal of Economics, 1*, 55–66.

Kim, S. H., & Farragher, E. I. (1981). Current capital budgeting practices. *Management Accounting, 6*, 26–30.

Lecraw, D. J. (1985). Some evidence on transfer pricing by multinational corporations. In A. M. Rugman & L. Eden (Eds.), *Multinationals and transfer pricing*. New York: St. Martin's Press.

Lee, C. F. (1993). *Statistics for business and financial economics*. Lexington: Heath.

Lee, H. (1993). A structured methodology for software development effort prediction using the analytic hierarchy process. *Journal of Systems and Software, 21*, 179–186.

Lee, S. M., & Lerro, A. (1974). Capital budgeting for multiple objectives. *Financial Management, 3*, 51–53.

Lee, Y. R., Shi, Y., & Yu, P. L. (1990). Linear optimal designs and optimal contingency plans. *Management Science, 36*, 1106–1119.

Lee, H., Nazem, S. M., & Shi, Y. (1995). Designing rural area telecommunication networks via hub cities. *Omega: International Journal of Management Science, 22*(3), 305–314.

Li, Y., Huang, M., Chin, K., Luo, X., & Han, Y. (2011). Integrating preference analysis and balanced scorecard to product planning house of quality. *Computers & Industrial Engineering, 60*, 256–268.

Liberatore, M. J., Mohnhan, T. F., & Stout, D. E. (1992). A framework for integrating capital budgeting analysis with strategy. *The Engineering Economist, 38*, 31–43.

Lin, T. W. (1993). Multiple-criteria capital budgeting under risk. *Advances in Mathematical Programming and Financial Planning, 3*, 231–239.

Lin, L., Lefebvre, C., & Kantor, J. (1993). Economic determinants of international transfer pricing and the related accounting issues, with particular reference to Asian Pacific countries. *The International Journal of Accounting*, 49–70.

Locke, E. A., Shaw, K. N., & Lathom, G. P. (1981). Goal setting and task performance: 1969–1980. *Psychological Bulletin, 90*, 125–152.

Lorie, J. H., & Savage, L. I. (1955). Three problems in rationing capital. *Journal of Business, 28*, 229–239.

Mensah, Y. M., & Miranti, P. J., Jr. (1989). Capital expenditure analysis and automated manufacturing systems: A review and synthesis. *Journal of Accounting Literature, 8*, 181–207.

Merville, L. J., & Petty, W. J. (1978). Transfer pricing for the multinational corporation. *The Accounting Review, 53*, 935–951.

Muralidhar, K., Santhanam, R., & Wilson, R. L. (1990). Using the analytic hierarchy process for information system project selection. *Information and Management, 18*, 87–95.

Naert, P. A. (1973). Measuring performance in a decentralized corporation with interrelated divisions: Profit center versus cost center. *The Engineering Economist*, 99–114.

Nakayama, H. (1994). Some remarks on trade-off analysis in multi-objective programming. The XIth International conference on multiple criteria decision making, Coimbra, Portugal, Aug. 1–6.

Nanni, A. J., Dixon, J. R., & Vollmann, T. E. (1990). Strategic control and performance measurement. *Journal of Cost Management, 4*, 33–42.

Natarajan, T., Balasubramanian, S. A., & Manickavasagam, S. (2010). Customer's choice amongst self-service technology (SST) channels in retail banking: A study using analytical hierarchy process (AHP). *Journal of Internet Banking and Commerce, 15*, 1–15.

Onsi, M. (1970). A transfer pricing system based on opportunity cost. *The Accounting Review, 45*, 535–543.

Palliam, R. (2005). Application of a multi-criteria model for determining risk premium. *The Journal of Risk Finance, 6*, 341–348.

Pike, R. (1983). The capital budgeting behavior and corporate characteristics of capital-constrained firms. *Journal of Business Finance and Accounting, 10*, 663–671.

Pike, R. (1988). An empirical study of the adoption of sophisticated capital budgeting practices and decision-making effectiveness. *Accounting and Business Research, 18*, 341–351.

Reeves, G. R., & Franz, L. S. (1985). A simplified interactive multiple objective linear programming procedure. *Computers and Operations Research, 12*, 589–601.

Reeves, G. A., & Hedin, S. R. (1993). A generalized interactive goal programming procedure. *Computers and Operations Research, 20*, 747–753.

Reeves, G. F., Lawrence, S. M., & Gonzalez, J. J. (1988). A multiple criteria approach to aggregate industrial capacity expansion. *Computers and Operations Research, 15*, 333–339.

Ronen, J., & McKinney, G. (1970). Transfer pricing for divisional autonomy. *Journal of Accounting Research, 8*, 99–112.

Rotch, W. (1990). Activity-based costing in service industries. *Journal of Cost Management*, 4–14.

Ruefli, T. W. (1971). A generalized goal decomposition model. *Management Science, 17*, 505–518.

Saaty, T. L. (1980). *The analytic hierarchy process*. New York: McGraw-Hill.

Santhanam, R., Muralidhar, K., & Schniederjans, M. (1989). Zero-one goal programming approach for information system project selection. *Omega: International Journal of Management Science, 17*, 583–598.

Schniederjans, M. J., & Garvin, T. (1997). Using the analytic hierarchy process and multi-objective programming for the selection of cost drivers in activity-based costing. *European Journal of Operational Research, 100*, 72–80.

Schniederjans, M. J., & Wilson, R. L. (1991). Using the analytic hierarchy process and goal programming for information system project selection. *Information and Management, 20*, 333–342.

Seiford, L., & Yu, P. L. (1979). Potential solutions of linear systems: The multi-criteria multiple constraint level program. *Journal of Mathematical Analysis and Applications, 69*, 283–303.

Shi, Y. (1991). Optimal linear production systems: Models, algorithms, and computer support systems. Ph.D. dissertation, School of Business, University of Kansas.

Shi, Y., & Lee, H. (1992). A binary integer linear programming with multi-criteria and multi-constraint levels (Working paper #92–21). School of Business, University of Nebraska at Omaha.

Shi, Y., & Yu, P. L. (1992). Selecting optimal linear production systems in multiple criteria environments. *Computer and Operations Research, 19*, 585–608.

Singhal, J., Marsten, R. E., & Morin, T. L. (1989). Fixed order brand-and-bound methods for mixed-integer programming: The ZOOM system. *ORSA Journal on Computing, 1*, 44–51.

Steuer, R. (1986). *Multiple criteria optimization: Theory. Computation and application*. New York: Wiley.

Suh, C.K., Suh, E.H., Choi, J. (1993). A two-phased DSS for R&D portfolio selection. *Proceedings of International DSI*, 351–354.

Tang, R. Y. W. (1992). Transfer pricing in the 1990s. *Management Accounting, 73*, 22–26.

Tapinos, E., Dyson, R. G., & Meadows, M. (2011). Does the balanced scorecard make a difference to the strategy development process? *The Journal of the Operational Research Society, 62*, 888–899.

Thanassoulis, E. (1985). Selecting a suitable solution method for a multi-objective programming capital budgeting problem. *Journal of Business Finance and Accounting, 12*, 453–471.

Thomas, A. L. (1980). *A behavioral analysis of joint-cost allocation and transfer pricing*. Champaign: Stipes Publishing.

Turney, S. T. (1990). Deterministic and stochastic dynamic adjustment of capital investment budgets. *Mathematical Computation and Modeling, 13*, 1–9.

Watson, D. J. H., & Baumler, J. V. (1975). Transfer pricing: A behavioral context. *The Accounting Review, 50*, 466–474.

Weingartner, H. M. (1963). *Mathematical programming and the analysis of capital budgeting problems*. Englewood Cliffs: Prentice-Hall.

Wu, C.-R., Lin, C.-T., & Tsai, P.-H. (2011). Financial service sector performance measurement model: AHP sensitivity analysis and balanced scorecard approach. *The Service Industries Journal, 31*, 695.

Yu, P. L. (1985). *Multiple criteria decision making: Concepts, techniques and extensions*. New York: Plenum.

Yu, Y. L., & Zeleny, M. (1975). The set of all nondominated solutions in the linear cases and a multicriteria simplex method. *Journal of Mathematical Analysis and Applications, 49*, 430–458.

Yunker, P. J. (1983). Survey study of subsidiary, autonomy, performance evaluation and transfer pricing in multinational corporations. *Columbia Journal of World Business, 19*, 51–63.

Zeleny, M. (1974). *Linear multi-objective programming*. Berlin: Springer.

Statistics Methods Applied in Employee Stock Options

30

Li-jiun Chen and Cheng-der Fuh

Contents

L.-j. Chen (✉)
Department of Finance, Feng Chia University, Taichung City, Taiwan
e-mail: lijiunchen@fcu.edu.tw

C.-d. Fuh
Graduate Institute of Statistics, National Central University, Zhongli City, Taiwan
e-mail: stcheng@stat.sinica.edu.tw

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_30,

Abstract

This study presents model-based and compensation-based approaches to determining the price-subjective value of employee stock options (ESOs). In the model-based approach, we consider a utility-maximizing model in which the employees allocate their wealth among company stock, a market portfolio, and risk-free bonds, and then we derive the ESO formulas, which take into account illiquidity and sentiment effects. By using the method of change of measure, the derived formulas are simply like those of the market value with altered parameters. To calculate the compensation-based subjective value, we group employees by hierarchical clustering with a K-means approach and back out the option value in an equilibrium competitive employment market.

Further, we test illiquidity and sentiment effects on ESO values by running regressions that consider the problem of standard errors in the finance panel data. Using executive stock options and compensation data paid between 1992 and 2004 for firms covered by the Compustat Executive Compensation Database, we find that subjective value is positively related to sentiment and negatively related to illiquidity in all specifications, consistent with the offsetting roles of sentiment and risk aversion. Moreover, executives value ESOs at a 48 % premium to the Black-Scholes value and ESO premiums are explained by a sentiment level of 12 % in risk-adjusted, annualized excess return, suggesting a high level of executive overconfidence.

Keywords

Employee stock option • Sentiment • Subjective value • Illiquidity • Change of measure • Hierarchical clustering with K-means approach • Standard errors in finance panel data • Exercise boundary • Jump diffusion model

30.1 Introduction

Employee stock options (ESOs) are a popular method of compensation. According to the Compustat Executive Compensation database, the number of option grants increased from roughly 0.25 billion in 1992 to 1.4 billion in 2001. Specifically, in fiscal 2001, 53 % of total pay came from granted options, compared with 33 % in 1992.[1] Moreover, executives receive averagely more than ten thousand options after 1998. ESOs can help firms retain talent and reduce agency costs (Jensen and Meckling 1976). They also mitigate risk-related incentive problems (Agrawal and Mandelker 1987; Hemmer et al. 2000) and provide an alternative to cash compensation, which is especially important for firms facing

[1]The option values are estimated by the calculation in ExecuCompustat database.

financial constraints (Core and Guay 2001). In addition, ESOs attract highly motivated and able employees (Core and Guay 2001; Oyer and Schaefer 2005). All of these factors contribute to the importance of ESOs for corporate governance and finance research.

The illiquidity problem of ESOs cannot be neglected. ESOs usually have a vesting period during which they cannot be exercised, and hence they cannot be redeemed for a fixed period of time. Furthermore, ESOs are generally not publicly traded and their sale is not permitted. Standard methods for valuing options are difficult to apply to ESOs. Because of the illiquidity of ESOs, many employees hold undiversified portfolios that include large stock options of their own firms. Because of the impossibility of full diversification, the value perceived by employees (subjective value) may be quite different from the traded option. We consider the subjective value to be what a constrained agent would pay for the ESOs and the market value to be the value perceived by an unconstrained agent. Many papers address the differences between subjective and market value, each concluding that the subjective value of ESOs should be less than the market value (Lambert et al. 1991; Hall and Murphy 2002; Ingersoll 2006).

If ESOs are generally worth less than market value, why do employees continue to accept and, indeed, sometimes prefer ESO compensation? One reasonable explanation is employee sentiment. Sentiment means positive private information or behavioral overconfidence regarding the future risk-adjusted returns of the firm. Simply, the employee believes that he possesses private information and can benefit from it. Or the employee overestimates the return of the firm and believes ESOs are valuable. Some empirical evidence supports this conjecture. Oyer and Schaefer (2005) and Bergman and Jenter (2007) posit that employees attach a sentiment premium to their stock options; firms exploit this sentiment premium to attract and retain optimistic employees. Hodge et al. (2009) provide survey evidence and find that managers subjectively value the stock option greater than its Black-Scholes value.

Three statistics methods are applied in our ESO study (Chen and Fuh 2011; Chang et al. 2013), including change of measure, hierarchical clustering with a K-means approach, and estimation of standard errors in finance panel data. We derive a solution for ESO value that is a function of both illiquidity and sentiment in a world where employees balance their wealth between the company's stock, the market portfolio, and a risk-free asset. By using the method of change of measure, we find a probability measure and then the ESO formulas are derived easily. In addition, from the ESO pricing formulas, we are able to not only estimate the subjective values but also study the exercise policies. Early exercise is a pervasive phenomenon and, importantly, the early exercise effect is critical in valuation of ESOs, especially for employees who are more risk averse and when there are more restrictions on stock holding.

Applying a comprehensive set of executive options and compensation data, this study empirically determines subjective value, grouping employees by hierarchical clustering with a K-means approach and backing out the option value in an equilibrium competitive employment market. Specifically, we group executives according to position, the firm's total market value, nonoption compensation, and the immediate exercise value of the options for each industry by hierarchical clustering with a K-means approach. We then calculate the empirical value of ESO that each executive places on their ESOs in order for total compensation to be equivalent in the same cluster. Further, we model both illiquidity and sentiment effects and test them with executive options data. We regress the empirical ESO values on the proportion of total wealth held in illiquid firm-specific holdings and sentiment, which is estimated from our pricing formula or Capital Asset Pricing Model (CAPM) risk-adjusted alpha under controlling key options pricing variables such as moneyness, time to maturity, volatility, and dividend payout. As we know, when the residuals are correlated across observations, the standard errors of estimated coefficients produced by Ordinary Least Squares (OLS) may be biased and then lead to incorrect inference. Petersen (2009) compares the different methods used in the literature and gives researchers guidance for their use. The data we collected are from multiple firms over several years. Hence, we consider the problem of standard errors in finance panel data and run the regressions, including standard errors clustered by firm, year, and both firm and year.

The remainder of this chapter is organized as follows. Section 30.2 introduces some preliminary knowledge. Section 30.3 develops our model and addresses the approaches to price-subjective value of ESOs. Section 30.4 presents the simulation results. Section 30.5 shows the empirical study; Sect. 30.6 concludes the chapter.

30.2 Preliminary Knowledge

30.2.1 Change of Measure

In our ESO study, we assume stock price follows a jump-diffusion process. Here, we briefly introduce change of measure for a compound Poisson and Brownian motion. Suppose that we have a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ on which is defined a Brownian motion W_t. Suppose that on this same probability space there is defined a compound Poisson process

$$Q_t = \sum_{i=0}^{N_t} Y_i$$

with intensity λ and jumps having density function $f(y)$. Assume that there is a single filtration $\mathcal{F}_t$, $t \geq 0$, for both the Brownian motion and the compound Poisson process. In this case, the Brownian motion and compound Poisson process must be independent (see Shreve 2008).

Let $\widetilde{\lambda}$ be a positive number, let $\widetilde{f}(y)$ be another density function with the property that $\widetilde{f}(y) = 0$ whenever $f(y) = 0$, and let $\Theta(t)$ be an adapted process. We define

$$Z_1(t) = \exp\left\{-\int_0^t \Theta(u)dW_u - \frac{1}{2}\int_0^t \Theta^2(u)du\right\},$$

$$Z_2(t) = e^{\left(\lambda-\widetilde{\lambda}\right)t}\prod_{i=0}^{N_t}\frac{\widetilde{\lambda}\widetilde{f}(Y_i)}{\lambda f(Y_i)},$$

$$Z(t) = Z_1(t)Z_2(t).$$

It can be verified that the process $Z(t)$ is a martingale.

Lemma 30.1 *Define* $\widetilde{\mathbb{P}}(A) = \int_A Z(T)d\mathbb{P}$ *for all* $A \in \mathcal{F}$. *Under the probability measure* $\mathbb{P}$, *the process*

$$\widetilde{W}_t = W_t + \int_0^t \Theta(s)ds$$

is a Brownian motion, Q_t *is a compound Poisson process with intensity* $\widetilde{\lambda}$ *and independent, identically distributed jump sizes having density* $\widetilde{f}(y)$, *and the processes* $\widetilde{W}_t$ *and* Q_t *are independent.*

The proof of Lemma 30.1 is presented by Shreve (2008). This lemma is useful for us to derive the ESO formula in Sect. 30.3.

30.2.2 Hierarchical Clustering with K-Means Approach

The main assumption of the compensation-based subjective value is that all executives within the same group receive the same total compensation. For each executive in the group, the implied subjective value is derived by comparing the difference between nonoption compensation and the average compensation. Grouping executives appropriately is essential in the compensation-based approach.

Clustering is an unsupervised technique for analyzing data and dividing patterns (observations, data items, or feature vectors) into groups (clusters). There are two kind of clustering algorithms: hierarchical and partitional approaches. Hierarchical methods produce a nested structure of partitions, whereas partitional methods produce only one partition (Jain et al. 1999).

Hierarchical clustering algorithms repeat the cycle of either merging smaller clusters into larger ones or dividing larger clusters to smaller ones. An agglomerative clustering strategy uses the bottom-up approach of merging clusters into larger ones, whereas divisive clustering strategy uses the top-down approach of splitting larger clusters into smaller ones. Typically, the greedy approach is used in deciding which larger/smaller clusters are used for merging/dividing. Euclidean distance, Manhattan distance, and cosine similarity are some of the most commonly used metrics of similarity for numeric data. For non-numeric data, metrics such as the Hamming distance are used. It is important to note that the actual observations (instances) are not needed for hierarchical clustering, because only the matrix of distances is sufficient. The user can obtain different clustering depending on the level that is cut.

A partitional clustering algorithm obtains a single partition of the data instead of a clustering structure. A problem accompanying the use of a partitional algorithm is the choice of the number of desired output clusters (usually called k). One of the most commonly used partitional clustering algorithms is the K-means clustering algorithm. It starts with a random initial partition and keeps reestimating cluster centers and reassigning the patterns to clusters based on the similarity between the pattern and the cluster centers. These two steps are repeated until a certain intra-cluster similarity objective function and inter-cluster dissimilarity objective function are optimized. Therefore, sensible initialization of centers is an important factor in obtaining quality results from partitional clustering algorithms.

Hierarchical and partitional approaches each have their advantages and disadvantages. Therefore, we apply a hybrid approach combining hierarchical and partitional approaches in this study.

30.2.3 Standard Errors in Finance Panel Data

In finance panel data, the residuals may be correlated across firms or across time, and OLS standard errors can be biased. Petersen (2009) compares the different methods used in the literature and provides guidance to researchers as to which method should be used.

The standard regression for a panel data is

$$Y_{it} = X_{it}\beta + \epsilon_{it} \quad i = 1, \ldots, N; \;\; t = 1, \ldots, T$$

where there are observations on firms i across years t. X and ϵ are assumed to be independent of each other and to have finite variance. OLS standard errors are unbiased when the residuals are independent and identically distributed. However, it may result in incorrect inference when the residuals are correlated across observations.

In finance study, there are two general forms of dependence: time-series dependence (firm effect), in which the residuals of a given firm may be correlated across years for a given firm, and cross-sectional dependence (time effect), in which the residuals of a given year may be correlated across different firms. Considering the firm effect, the residuals and independent variable are specified as

$$\epsilon_{it} = \gamma_i + n_{it}; \; X_{it} = \mu_i + v_{it}$$

Both the independent variable and the residual are correlated across observations of the same firm but are independent across firms. Petersen (2009) shows that OLS, Fama-MacBeth, and Newey-West standard errors are biased, and only clustered standard errors are unbiased as they account for the residual dependence created by the firm effect.

Considering the time effect, the residuals and independent variable are specified as

$$\epsilon_{it} = \delta_t + \eta_{it};\;\; X_{it} = \zeta_t + v_{it}.$$

OLS standard errors still underestimate the true standard errors. The clustered standard errors are much more accurate, but unlike the results with the firm effect, they underestimate the true standard error. However, the bias in the clustered standard error estimates declines with the number of clusters. Because the Fama-MacBeth procedure is designed to address a time effect, the Fama-MacBeth standard errors are unbiased.

In both firm and time effect, the residuals and independent variable are specified as

$$\epsilon_{it} = \gamma_i + \delta_t + \eta_{it};\;\; X_{it} = \mu_i + \zeta_t + v_{it}.$$

Standard errors clustered by only one dimension are biased downward. Clustering by two dimensions produces less biased standard errors. However, clustering by firm and time does not always yield unbiased estimates. We only introduce the method we use in this study. Readers wanting to know more methods for estimation of standard error in financial panel data should refer to Petersen (2009).

30.3 Employee Stock Options

This section introduces two methods to price the subjective value of ESOs. We call them the model-based approach and the compensation-based approach. To derive the model-based subjective value, we use the technique of change of measure and find a probability measure P^* such that the option value can be generated simply. To calculate compensation-based subjective value, we group employees by hierarchical clustering with a K-means approach and back out the option value in an equilibrium-competitive employment market.

30.3.1 Model-Based Approach to Subjective Value

From Chen and Fuh (2011), we have a three asset economy in which the employee allocates their wealth among three assets: company stock S, market portfolio M, and risk-free bond B. Because of the illiquidity of ESO, the employee is constrained to allocate a fixed fraction α of their wealth to company stock (via some form of ESO). Define the jump-diffusion processes for the three assets as follows:

$$\begin{cases} \dfrac{dS}{S} = \mu_s dt + \sigma_s dW_m + v dW_s + d\sum_{i=0}^{N_t}(Y_i - 1), \\ \dfrac{dM}{M} = (\mu_m - d_m)dt + \sigma_m dW_m, \\ \dfrac{dB}{B} = rdt, \end{cases} \tag{30.1}$$

where $\mu_s = \mu - d - \lambda k$. μ, μ_m, r are instantaneous expected rates of return for the stock, market portfolio, and risk-free bond, respectively. d and d_m are dividends for the stock and market portfolio, respectively. The Brownian motion process W_m represents the normal systematic risk of the market portfolio. The Brownian motion process W_s and jump process N_t are the idiosyncratic risk of the company stock, where N_t captures the jump risk of company stock and follows a Poisson distribution with average frequency λ. $Y_i - 1$ represents the percentage of stock variation when the ith jump occurs. Denote $E(Y_i - 1) = k$ and $E(Y_i - 1)^2 = k_2$ for all i. σ_s and σ_m are the normal systematic portions of total volatility for the stock and the market portfolio, respectively, whereas v is the normal unsystematic volatility of the stock. The two Brownian motions and the jump process are presumed independent. For simplicity, we assume that CAPM holds so that the efficient portfolio is the market.

Following the idea of Ingersoll (2006), we solve the constrained optimization problem and determine the employee's indirect utility function. The marginal utility is then used as a subjective state price density to value compensation. We assume that the employee's utility function $U(\cdot)$ is set as $U(C) = C^\gamma/\gamma$ with a coefficient of relative risk aversion, $1 - \gamma$. The process of the employee's marginal utility or the pricing kernel can be derived as[2]:

$$\begin{aligned} \frac{dJ_W}{J_W} = & -\hat{r}\, dt - \hat{\sigma}\, dW_m - (1-\gamma)\alpha v dW_s \\ & + d\sum_{i=0}^{N_t}\left\{[\alpha(Y_i - 1) + 1]^{\gamma-1} - 1\right\}, \end{aligned} \tag{30.2}$$

where $J_W = \frac{\partial J[W(t), t]}{\partial W(t)}$ is the marginal utility, $J[W(t), t]$ and $W(t)$ are the employee's total utility and wealth at time t, $\hat{r} = r - (1-\gamma)\left(\alpha^2 v^2 + \frac{1}{2}\alpha^2 \gamma \lambda + \alpha \lambda k\right)$, and $\hat{\sigma} = \frac{\mu_m - r}{\sigma_m}$.

The rational equilibrium value of the ESO at time t, $F(S_t, t)$, satisfies the Euler equation

$$F(S_t, t) = \frac{E_t\{J_W[W(T), T]F(S_T, T)\}}{J_W[W(t), t]}, \tag{30.3}$$

where $F(S_T, T)$ is the payoff at the maturity T. To easily calculate the ESO value, we find a probability measure P^* by using change of measure method and then the second equality in the following Eq. (30.4) is satisfied.

[2]Because the process can be similarly derived from Chang et al. (2008), it is omitted.

$$
\begin{aligned}
F(S_t,t) &= \frac{E_t\{J_W[W(T),T]F(S_T,T)\}}{J_W[W(t),t]} \\
&= e^{-r*(T-t)}E_t^*[F(S_T,T)],
\end{aligned}
\tag{30.4}
$$

where $\frac{dP*}{dP} = \frac{Z(T)}{Z(t)}$, $Z(t) = e^{r^*t}J_W[W(t), t]$, r^* is subjective bond yield, and E_t^* is the expectation under P^* and information at time t. Under P^*, the stock process can be expressed as

$$
\frac{dS}{S} = (r^* - d^*)dt + \sigma_N dW_t^* + d\sum_{i=0}^{N_t}(Y_i - 1),
$$

where

$$
\begin{aligned}
r^* &= r - (1-\gamma)\left(\alpha\lambda k + \frac{1}{2}\gamma\lambda k_2\alpha^2 + \alpha^2 v^2\right) \\
&\quad - \lambda(\xi - 1), \\
d^* &= d - (1-\gamma)\left[\alpha\lambda k + \frac{1}{2}\gamma\lambda k_2\alpha^2 - (1-\alpha)\alpha v^2\right] \\
&\quad - \lambda(\xi - 1) + \lambda k, \\
\sigma_N^2 &= \sigma_s^2 + v^2, \lambda^* = \lambda\xi, \xi = E[\alpha(Y_i - 1) + 1]^{\gamma-1},
\end{aligned}
$$

W_t^* is the standard Brownian motion and N_t is a Poisson process with rate λ^*. A detailed explanation is given in Appendix 1.

30.3.1.1 European ESO

First, we consider the simple ESO contract, the European ESO. The price formula is presented in Theorem 30.1.

Theorem 30.1 *The value of the European ESO with strike price K and time to maturity τ, written on the jump-diffusion process in Eq.* (30.1), *is as follows:*

$$
C_E(S_t,\tau) = \sum_{j=0}^{\infty}\frac{(\lambda*\tau)^j e^{-\lambda*\tau}}{j!}C(j) \tag{30.5}
$$

where

$$
\begin{aligned}
C(j) &= \left\{S_t e^{-d^*\tau}E^*\left[\prod_{i=0}^{j}Y_i\Phi\left(d_1^*\right)\right]\right. \\
&\quad \left. -Ke^{-r^*\tau}E^*\left[\Phi\left(d_2^*\right)\right]\right\} \\
d_1^* &= \frac{\ln\left(S_t\prod_{i=0}^{j}Y_i/K\right) + \left(r^* - d^* + \frac{1}{2}\sigma_N^2\right)\tau}{\sigma_N\sqrt{\tau}}, \\
d_2^* &= d_1^* - \sigma_N\sqrt{\tau}.
\end{aligned}
$$

The proof of Theorem 30.1 is in Appendix 2.

30.3.1.2 American ESO

Suppose that the option can be exercised at n time instants. These time instants are assumed to be regularly spaced at intervals of Δt, and denoted by t_i, $0 \leq i \leq n$, where $t_0 = 0$, $t_n = T$, and $t_{i+1} - t_i = \Delta t$ for all i. Denote C_A as the value of the American call option, C_E as the value of the European call option, K as the strike price, and $S_i = S_{t_i}$. The critical price at these time points is denoted by S_i^*, $0 \leq i \leq n$, and is the price at which the agent is indifferent between holding the option and exercising. Denote E_i^* as the expectation under P^* and information at time t_i.

Theorem 30.2 (Chen and Fuh 2011) *The value of the American ESO exercisable at n time instants, when the ESO is not exercised, written on the jump-diffusion process in Eq.* (30.1) *is as follows:*

$$\begin{aligned}
&C_A(S_0,T)\\
&= C_E(S_0,T) + \sum_{\ell=1}^{n-1} e^{-r*\ell\Delta t} E_0^*\Big\{ [S_\ell\big(1-e^{-d*\Delta t}\big)\\
&\quad - K\big(1-e^{-r*\Delta t}\big)]I_{\{S_\ell \geq S_\ell^*\}}\Big\}\\
&\quad - \sum_{j=2}^{n} e^{-r*j\Delta t} E_0^*\Big\{[C_A(S_j,(n-j)\Delta t) - \big(S_j - K\big)]I_{\{S_{j-1}\geq S_{j-1}^*\}} I_{\{S_j<S_j^*\}}\Big\}.
\end{aligned} \tag{30.6}$$

The critical price S_i^* at time t_i for $i = 1, \cdots, n$ is defined as the solution to the following equation:

$$\begin{aligned}
&S_i^* - K\\
&\quad = C_E\big(S_i^*,(n-i)\Delta t\big)\\
&\qquad + \sum_{\ell=1}^{n-i-1} e^{-r^*\ell\Delta t} E_i^*\Big\{\big[S_{i+\ell}\big(1-e^{-d^*\Delta t}\big)\\
&\qquad - K\big(1-e^{-r^*\Delta t}\big)\big]I_{\{S_{i+\ell}\geq S_{i+\ell}^*\}}\Big\}\\
&\qquad - \sum_{j=2}^{n-i} e^{-r^*j\Delta t} E_i^*\Big\{\big[C_A\big(S_{i+j},(n-i-j)\Delta t\big)\\
&\qquad - \big(S_{i+j}-K\big)\big]I_{\{S_{i+j-1}\geq S_{i+j-1}^*\}} I_{\{S_{i+j}<S_{i+j}^*\}}\Big\},
\end{aligned}$$

where $C_E(S_0, T)$ and $C_E(S_i^*, (n-i)\Delta t)$ are calculated in Theorem 30.1.

The value of the American call option, when exercise is allowed at any time before maturity, is obtained by taking the limit as Δt tends to zero in Eq. (30.6).

30.3.1.3 Sentiment Effect

Often, the manager awarded an incentive option may have different beliefs about the company's prospects than the investing public does. The manager believes that they possess private information and can benefit from it. Or they have behavioral overconfidence regarding future risk-adjusted return of their firm and believe ESOs are valuable. We consider the impact of sentiment on ESO values and the exercise decision. Now, the drift term of stock process in Eq. 30.1 becomes $\mu_s = \mu + s - d - \lambda k$, where sentiment level is denoted by s. In other words, the employee overestimates or rationally adjusts the risk-adjusted return of the company owing to inside information by s, then the same analysis in Theorem 30.1 and 30.2 are valid with a simple adjustment in parameters. The adjusted interest rate and dividend yield used in pricing are

$$r^* = r + \alpha s - (1-\gamma)\left(\alpha\lambda k + \frac{1}{2}\gamma\lambda k_2\alpha^2 + \alpha^2 v^2\right) - \lambda(\xi - 1),$$

$$d^* = d - (1-\alpha)s - (1-\gamma)\left[\alpha\lambda k + \frac{1}{2}\gamma\lambda k_2\alpha^2 - (1-\alpha)\alpha v^2\right] - \lambda(\xi - 1) + \lambda k.$$

30.3.2 Compensation-Based Approach to Subjective Value

The subjective value of ESOs implied by total compensation packages is calculated using a hierarchical clustering with K-means methodology to split executives into like groups based on industry, rank, year, the firm market value,[3] nonoption compensation, and the immediate exercise value of the options.[4] The number of groups is decided by a cubic clustering criterion and the average total compensation is calculated. Then, assuming that all executives within the same cluster receive the same total compensation, for each executive in this cluster, the implied subjective value is derived by comparing the difference between nonoption compensation and the average compensation. We then set all negative implied ESO values equal to zero and recalculate average compensation in each cluster with these subjective values, repeating until the relative sum of changes in subjective values in a given cluster is less than 0.01. This eliminates some negative subjective values such that the final number of negative or zero values is about 5.7 % of our dataset. Worth noting is the observation that, even in the first iteration of the process, after grouping, only about 7.9 % of our data has options with a negative or zero value, lending credence to the stability of our groupings.

To illustrate the intuition, presume that all executives within the same cluster receive the same compensation on average, where any differences in salaries, bonuses, and other income should be accounted for by options. If CEOs' average total annual compensation in a given year is three million, a particular CEO who receives one

[3]Gabaix and Landier (2008) finds that total market value as a proxy for firm size has the strongest predictive power on compensation. We, however, redo all tests using number of employees as the size proxy and find qualitatively identical results.

[4]The details of hierarchical clustering with a K-means approach and its performance are presented in Appendix 2.

million in nonoption compensation must then value their options at two million in order to agree to continued employment. Importantly, it may be the case that the market value of these options is only $100,000, but the CEO subjectively values them at $500,000 because they believe the market to have undervalued the options. Although this method of calculation is clearly not perfectly precise, we tried numerous robustness checks by using different grouping criteria, all of which arrived at qualitatively identical results. Some intangible sources of value, such as training, learning opportunities, and advantageous work environments, are not controlled here but may be relatively unimportant given that this is an executive database of listed firms.

30.4 Simulation Results

Section 30.3 provided a pricing formula for ESOs that includes illiquidity of the options and sentiment of the employees. Moreover, from this ESO pricing formula, we can not only estimate the subjective values but also study the exercise policies. The exercise boundary is endogenously derived by finding the minimum stock price such that the option value equals its intrinsic value for each time. In other words, the employee exercises the option when the stock price is above the exercise boundary. To illustrate our model, this section discusses factors that affect ESO values and exercise decisions including stock holding constraint, sentiment, level of risk aversion, moneyness, dividend, time to maturity, total volatility, and normal unsystematic volatility. We also empirically test these effects in next section.

According to the collected data from Compustat, the model parameters stock price S, strike price K, total volatility σ, dividend yield d, interest free rate r, and time to maturity τ are set to 25, 25, 0.3, 2 %, 5 %, 10, respectively. Normal unsystematic volatility v is two-thirds of the total volatility following calibrations applied by Compustat and the majority of papers in the area.[5] We employ the common parameterization for the coefficient of relative risk aversion $1 - \gamma = 2$ and three jump size models: double exponential, bivariate jump, and $Y = 0$ (no residual value).[6] Additionally, default intensity $\lambda = 0.01$, following Duffee (1999) and Fruhwirth and Sogner (2006), which use US and German bond data, respectively.

30.4.1 Exercise Behavior

Employees exercising their ESOs earlier is a pervasive phenomenon. Considering the exercise policies is necessary when studying American ESOs. This is an essential departure from Chang et al. (2008), which considers European-type ESOs. A number

[5]See Bettis et al. (2005), Aggarwal and Samwick (2003), Ingersoll (2006), and Bryan et al. (2000).

[6]The parameters of double exponential are estimated by daily return data from 1992 to 2004. A jump occurs if return goes beyond ± 10 %, which relates to an approximately three-standard deviation daily return during this period.

of papers link early exercise behavior to under-diversification of employees (Hemmer et al. 1996; Core and Guay 2001; Bettis et al. 2005). The problem of valuing ESOs with early exercise is often approximated in practice by simply using the expected time until exercise in place of the actual time to maturity (Hull and White 2004; Bettis et al. 2005). The expected time until exercise is estimated from past experience. However, Ingersoll (2006) mentions that even using an unbiased estimate of the expected time until exercise will not give a correct estimate of the option's value. And this method cannot be used to determine the subjective value because it will be smaller due to the extra discounting required to compensate the lack of diversification.

A proper calculation must recognize that the decision to exercise is endogenous. Liao and Lyuu (2009) incorporate the exercise pattern instead of using the expected time until exercise technique in valuation of ESOs, to which the exercise patterns are under Chi-square distribution assumption and not derived endogenously. Ingersoll (2006) derives the exercise boundaries endogenously, while the exercise policies are restricted constant in time. We extend the method developed in Gukhal (2001), with a modification to include that an agent faces a constrained portfolio problem, and derive the time varying exercise policies endogenously.

Which factors cause employees to exercise their options early? Figure 30.1 compares the exercise boundaries for some factors. Note that exercise boundaries are decreasing function of time in all cases, which are different from the constant exercise policies in Ingersoll (2006). The more restrictions on the stock holding or the more risk averse the employee, the lower the exercise boundary. In other words, because of the impossibility of full diversification, employees who are more restricted on the stock holding or more risk averse prefer early exercise of their options. Employees with high sentiment will postpone the exercise timing due to the brightening prospect of the company. The employees who receive the money-type options also tend to exercise early. In addition, larger dividends induce employees to exercise their options sooner. Options with shorter lifetime are more quickly exercised. Employees do not have much time value in these options and tend to exercise their options earlier. Employees exercise volatile options early to balance their portfolio risk, especially for idiosyncratic risk increases. Indeed, our model findings are consistent with several empirical studies. For instance, Hemmer et al. (1996), Huddart and Lang (1996), and Bettis et al. (2005) show that early exercise is a pervasive phenomenon owing to risk aversion and undiversification of employees. Huddart and Lang (1996) find that exercise is negatively related to the time to maturity and positively correlated with the market-to-strike ratio and with the stock price volatility. Hemmer et al. (1996) and Bettis et al. (2005) also find that stock price volatility has a significant effect on exercise decisions. In high-volatility firms, employees exercise options much earlier than in low-volatility firms.

30.4.2 Factors' Effects on ESOs

Understanding the factors that affect ESO values and the exercise decision is important for firms designing stock option programs. As we mentioned before,

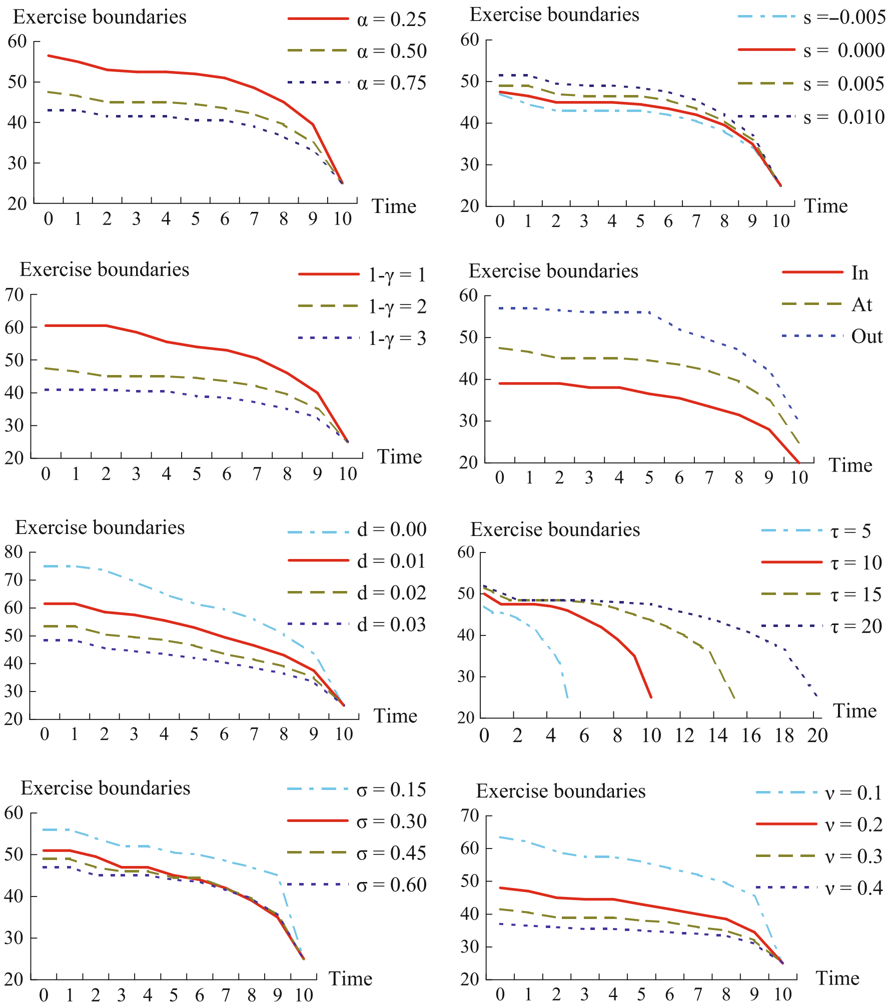


Fig. 30.1 Exercise boundaries. This figure presents the exercise boundaries according to stock holding constraint α, sentiment s, level of risk aversion $1-\gamma$, moneyness In: $K = 20$, At: $K = 25$, Out: $K = 30$, where K is exercise price, dividend yield d, time to maturity τ, total volatility σ, and idiosyncratic risk ν, respectively

ESO values and exercise decisions are closely related. Factors affecting the employees exercise policies will directly influence the valuation of ESOs. This section discusses the impact of factors on ESOs. The results are shown in Table 30.1, which presents the studied factors effect on ESO value, discount ratio, and early exercise premium, where ESO value is calculated by formula (30.6), discount ratio is defined as one minus the ratio of subjective to market value, and early exercise premium is the difference between American and European ESO value.

Table 30.1 Factors effect on employee stock options

Panel A: ESO values and discount ratios and early exercise premiums

	CA	CD	Premium		CA	CD	Premium
$\alpha = 0.00$	9.6487	0.0000	0.1839	$s = -0.005$	6.1495	0.3120	0.9385
$\alpha = 0.25$	7.5859	0.2110	0.5796	$s = 0.000$	6.4611	0.3279	0.8592
$\alpha = 0.50$	6.4628	0.3278	0.8592	$s = 0.005$	6.7632	0.3469	0.7504
$\alpha = 0.75$	5.7706	0.3998	0.9366	$s = 0.010$	7.1101	0.3655	0.6661
$1 - \gamma = 1$	8.5661	0.1122	0.4010	$sok = 25/30$	6.4228	0.2321	0.4413
$1 - \gamma = 2$	7.5859	0.2110	0.5796	$sok = 25/25$	7.5859	0.2110	0.5796
$1 - \gamma = 3$	6.9702	0.2811	0.9958	$sok = 25/20$	9.3061	0.1733	1.0268
$\mathrm{d} = 0.00$	8.4788	0.3557	0.3372	$\tau = 5$	5.4222	0.2545	0.4084
$\mathrm{d} = 0.01$	7.5724	0.3260	0.8031	$\tau = 10$	6.7636	0.3018	1.1608
$\mathrm{d} = 0.02$	6.7704	0.3054	1.1729	$\tau = 15$	7.4605	0.3257	1.9932
$\mathrm{d} = 0.03$	6.0673	0.2941	1.4674	$\tau = 20$	7.8992	0.3363	2.8481
$\sigma = 0.15$	5.4571	0.1800	0.0677	$v = 0.1$	7.1077	0.2607	0.2400
$\sigma = 0.30$	6.5440	0.3216	0.9413	$v = 0.2$	6.2785	0.3470	1.0146
$\sigma = 0.45$	7.0701	0.4378	2.1005	$v = 0.3$	5.6321	0.4142	2.1113
$\sigma = 0.60$	7.2776	0.5184	3.4397	$v = 0.4$	5.1440	0.4650	3.0637

Panel B: Vesting effect

	CA			CD			Premium		
	VP = 0	VP = 2	VP = 4	VP = 0	VP = 2	VP = 4	VP = 0	VP = 2	VP = 4
$\alpha = 0.25$	7.5859	7.5660	7.4923	0.2110	0.2125	0.2200	0.5796	0.5651	0.4898
$\alpha = 0.50$	6.4628	6.4029	6.2639	0.3278	0.3335	0.3479	0.8592	0.8041	0.6635
$\alpha = 0.75$	5.7706	5.6867	5.5299	0.3998	0.4081	0.4243	0.9366	0.8570	0.6989

This table presents the impact of factors on ESOs, including stock holding constraint α, sentiment s, level of risk aversion $1 - \gamma$, moneyness sok, dividend yield d, time to maturity τ, total volatility σ, idiosyncratic risk v, and vesting period VP. CA, CD and Premium are the ESO value, discount ratio (1-subjective/market) and early exercise premium, respectively

Unlike traditional options, ESOs usually have a vesting period during which they cannot be exercised and employees are not permitted to sell their ESOs. In this situation, employees receive the ESOs in a very illiquid market. Table 30.1 shows that subjective values ($\alpha \neq 0$) are uniformly smaller than the market values ($\alpha = 0$). These results are consistent with Lambert et al. (1991) and Hall and Murphy (2002), in which the subjective value is lower than market value due to the constrained fixed holding in the underlying stock. The more risk averse the employee (more positive $1 - \gamma$) and more restrictions on the stock holding (larger α), the more they lean to depreciate the option values and incur the higher early exercise premium. Note that early exercise effect on ESO values cannot be ignored in these situations. Because of the restriction of ESOs, many employees have undiversified portfolios with large stock options for their own firms. Therefore, a risk-averse employee discounts the ESO values. Discount ratios increase with stock holding constraint and the degree of risk aversion. In other words, employees who are more risk averse and more restricted on the stock holdings need to compensate more risk premium.

Employee sentiment enhances the option value and reduces the early exercise premium. One would expect for options with high sentiment having higher discount that the option value declines sharply when employees face an undiversification problem. The moneyness of each option *Sok* is the stock price at issuance divided by strike price. If the option is in (out of) the money, *Sok* is greater (less) than 1. In the money options have higher values, lower discount and higher early exercise premiums. Interestingly, even in the money options having less discount than out of the money, employees still tend more to early exercise in the money options to diversify their wealth portfolio risk. Larger dividends depreciate the option values and induce employees to exercise their options sooner, even they have lower discount ratios. More interestingly, the early exercise premium is not zero when no dividends are paid. This is a departure from traditional option theory, although it is consistent with the phenomenon that ESOs are exercised substantially before maturity date, even ESOs not paying dividends, because of the lack of diversification. Options with longer lifetime have more value; at the same time, they have higher discount ratios and early exercise premiums. Although not reported in the table, the lifetime of option may be negatively related to European ESO value. This is due to the longer one must wait and the greater risk caused by undiversification affecting the ESO value.

Whereas options may provide incentives for employees to work harder, they can also induce suboptimal risk-taking behavior. General option pricing results show that value should increase with risk while employees need to compensate more risk premium at the same time. It is not necessarily the case that subjective value is positive related to risk, as is the traditional result.[7] We have the usual finding that total volatility increases the option value; however, on the contrary, with respect to normal unsystematic volatility, the subjective value decreases with it. In the Black-Scholes framework, this risk is eliminated under a risk-neutral measure. However, in our model, the employee has an illiquid holding and full diversification is impossible. Hence, a risk-averse employee depreciates the ESO values. The discount and early exercise premium increasing with the volatility risk also can be found in Table 30.1. This is intuitive, because the more volatile the stock price, the higher the opportunity cost of not being able to exercise. Therefore, employees have more incentives to early exercise volatile options.[8]

Further we examine the effect of vesting on subjective value. Panel B compares the ESOs that vest immediately, after 2, and 4 years, respectively. Vesting obviously reduces the ESO values since it restricts the exercise timing. Discount ratio increasing with vesting period implies that market value is affected less than the subjective value. Because the constrained ESOs are usually exercised much earlier than unconstrained ESOs and more tend to fall afoul of the vesting rule. While

[7]Nohel and Todd (2005), Ryan and Wiggins (2001), and others show that option values increase with risk, however, they do not study the impact of increased idiosyncratic risk. Carpenter (2000) presents examples where convex incentive structures do not imply that the manager is more willing to take risks. The model used in Chang et al. (2008) is able to capture this result.

[8]While Panel A shows the results of options that vest immediately, we can also consider the vesting effect and there is no significant qualitative difference.

vesting has negative effect on the American ESOs, it has no effect on the European ESOs, therefore, early exercise premium decreases with vesting.

30.4.3 Offsetting Effect of Sentiment and Risk on ESO

The level of sentiment is estimated from two perspectives. First, we consider the sentiment effect on ESO value (SenV), and then the estimated sentiment level can be calculated whereby subjective value with sentiment is equal to market value. Secondly, we estimate sentiment level from the early exercise perspective (SenE), that is, what value of sentiment such that employees exercise their options at the time that unconstrained investors do. The sentiment level of European ESOs (SenVE) is also calculated. Due to the limitation of European options (they are not allowed to early exercise), the sentiment level can only be estimated from value perspective.

Estimation of sentiment is shown in Table 30.2. Here, we only list the estimated sentiment level at the money option because there is no obvious relationship between sentiment level and moneyness. Table 30.2 shows that sentiment estimated from the American and European ESO formulas have similar patterns. The more risk averse the employee and more restricted on the stock holding, the higher the sentiment level is needed. SenVE is slight higher than SenV because of more restrictions in the European contract.

30.5 Empirical Study

Applying a comprehensive set of executive options and compensation data, this study empirically prices both subjective value discount created by stock holding constraint and the risk-adjusted excess returns necessary for employees to offset the ESO risk premium, that is, the sentiment effect.

30.5.1 Data

Data for this study are collected from the Compustat Executive Compensation (Execucomp) database. From this database, all executive stock options issued between 1992 and 2004 are collected with stock price at issuance S, strike price K, maturity date T, implied volatility Vol, and dividend yield Div. While the median option is issued at the money, the mean is in the money ($Sok = 1.012$). Note that virtually all options are issued at the money ($Sok = 1$). Indeed, this is true for about 90 % of our dataset. Average values of time to maturity, implied volatility, and dividend yield are 9.3 years,[9] 0.43 % and 1.37 %, respectively.

[9]For some issues for which there is no time stamp, we assume an issuance date of July 1, inasmuch asthis would be the middle of the fiscal year for the vast majority of firms.

Table 30.2 Estimation of sentiment

	SenV			SenVE			SenE		
	$R=1$	$R=2$	$R=3$	$R=1$	$R=2$	$R=3$	$R=1$	$R=2$	$R=3$
$\alpha=0.25$	0.0100	0.0200	0.0300	0.0100	0.0200	0.0300	0.0097	0.0180	0.0291
$\alpha=0.50$	0.0199	0.0399	0.0599	0.0200	0.0400	0.0599	0.0194	0.0369	0.0585
$\alpha=0.75$	0.0298	0.0598	0.0896	0.0299	0.0598	0.0897	0.0289	0.0579	0.0866

This table presents the sentiment levels necessary to offset the ESO risk premium. Sentiment levels SenV and SenVE are calculated while the subjective value with sentiment is equal to market value for the American and European options, respectively. SenE is the value of sentiment such that an employee exercises their options at the time that unconstrained investors do. α and $R = 1 - \gamma$ represent the stock holding constraint and level of risk aversion

In addition to options data, we collect total compensation data from the Execucomp database, which includes salary, bonus, restricted stock, option, long-term incentive pay, and other income earned by executives each year. As can be seen in Table 30.3, the mean total annual compensation for executives in this dataset is a bit over \$2 million, with a median of just over \$1 million. The mean and median ESO compensation numbers are roughly \$1.2 and \$0.4 million, respectively. Not surprisingly, chief executives who were also board members received the highest compensation (\$4.2 million), but options are a substantial portion of that compensation (\$2.4 million). Indeed, options compensation generally substantially outweighs all other forms of compensation.

Following Dittmann and Maug (2007), we further define the net cash inflow (*NCash*) for each year as follows:

$$\begin{aligned}
&NCash\\
&= Fixed\ salary\ (after\ tax)\\
&\quad + \ Dividend\ income\ from\ shares\ held\ in\ own\ company\ (after\ tax)\\
&\quad + \ Value\ of\ restricted\ stock\ granted\\
&\quad - \ Personal\ taxes\ on\ restricted\ stock\ that\ vest\ during\ the\ year\\
&\quad + \ Net\ value\ realized\ from\ exercising\ options\ (after\ tax)\\
&\quad - \ Cash\ paid\ for\ purchasing\ additional\ stock
\end{aligned}$$

Fixed salary is the sum of five Compustat data types: Salary, Bonus, Other Annual, All Other Total, and long-term incentive pay (LTIP).[10] The year when the executive enters the database is denoted by t_E. The executive's cumulative wealth for year t is then

$$W(t) = NCash_t + \sum_{\ell=t_E}^{t-1} NCash_\ell \prod_{s=\ell+1}^{t} \left(1 + r_f^s\right).$$

[10] For cash paid for purchasing additional stock, where direct data is unavailable, we use the change in stock holdings times the year-end stock price to calculate this value.

Table 30.3 Compensation summary statistics

	Aggregate			Mean by title			
	Mean	Median	Std Dev	B&C	B&NC	NB&C	NB&NC
Salary	365	300	234	556	481	335	286
Bonus	336	151	816	650	479	289	222
Other annual	24	0	179	44	35	22	16
All other total	70	11	540	94	131	50	45
LTIP	77	0	442	127	128	72	48
Restricted stock	163	0	803	366	220	184	101
Options	1,178	378	3,407	2,382	1,683	1,264	748
Total	2,214	1,074	4,262	4,219	3,158	2,217	1,465

This table presents summary statistics for compensation data for four categories of executives: board & CEO (*B&C*), board & not CEO (*B&NC*), not board & CEO (*NB&C*), and not board & not CEO (*NB&NC*). Numbers are reported in 1,000s and *LTIP* represents the long-term incentive pay

In other words, assume that the executive has no wealth before entering the firm; all $NCash_t$ are realized at the end of the fiscal year and invested at the risk-free rate r_f^{t+1} during the next fiscal year. Then, α is the sum of all illiquid firm-specific holdings, including unvested restricted stocks and options, divided by total cumulative wealth. Average α is about 35 %, implying that the illiquid firm-specific holdings account for more than one-third of executive total wealth.[11] We also calculate α by an iterated approach that synchronizes α and subjective value simultaneously. Qualitative findings with respect to sentiment are identical. Several proxies of sentiment are used in an empirical study including previous year risk-adjusted CAPM alpha and *Fortune* magazine's list of top 100 firms to work for, and estimated from our ESO pricing formula.

30.5.2 Preliminary Findings

Substituting the subjective value implied by compensation data into our model along with the options variables given in our dataset, we are able to back out sentiment levels, *Sen*. Results are presented in Table 30.4. There are about 105,000 options issued by each firm (AvgIss) over the test period, with a total of nearly 2,700 firms and 82,000 total observations accounted for. Industry breakdowns, while exhibiting some fluctuations in point estimates, show that results across industries are qualitatively similar. While the mean Black-Scholes value of options, *BSOPM*, is about \$13.09, with some variation across industries, the mean subjective value, *Sub*, is more than \$19.38, reflecting a 48 % premium. That is, although virtually all of the theoretical literature implies a subjective value discount, empirical data show that executives generally value ESOs more highly than their Black-Scholes values. Though not reported in the table, t-tests show that subjective values

[11]Holland and Elder (2006) find that rank-and-file employees exhibit an α close to 10 % and concur that subjective value is decreasing in α because of risk aversion and under-diversification.

Table 30.4 Summary statistics for subjective value and sentiment by industry

Sector		*BSOPM*	*Sub*	*Sen*	AvgCom	AvgIss	Obs
10	Energy	12.589	16.239	0.089	1,774.08	78.46	4,307
15	Materials	10.822	17.613	0.076	1,399.90	61.89	6,412
20	Industrials	12.569	20.305	0.115	1,591.53	70.64	12,134
25	Con. Dis.	12.177	19.799	0.106	2,030.63	98.64	15,925
30	Con. Sta.	12.053	18.121	0.076	2,405.29	111.40	4,347
35	Health care	16.290	21.098	0.115	2,338.68	105.61	8,883
40	Financials	13.440	21.770	0.066	2,892.28	99.37	10,441
45	Inf. Tec.	15.853	18.499	0.229	2,748.47	164.53	14,614
50	Tel. Ser.	12.768	22.941	0.162	5,310.29	272.27	1,324
55	Utilities	5.475	14.633	0.063	1,370.11	67.66	3,931
	Others	9.487	9.583	0.208	1,360.34	111.24	56
	Total	13.088	19.385	0.120	2,213.73	105.13	82,374

This table presents, by industry: Black-Scholes value *BSOPM*, subjective value *Sub*, sentiment level *Sen*, average total compensation AvgCom, number of options issued AvgIss, and number of observations by individual Obs. AvgCom and AvgIss are reported in thousands. *Sen* is calculated using the European ESO formula (30.5), where the distribution of jump size follows $y = 0$. Con. Dis., Con. Sta., Inf. Tec., and Tel. Ser. refer to Consumer Discretionary, Consumer Staples, Information Technology and Telecommunication Services, respectively

are statistically significantly higher than Black-Scholes values at the 1 % level for almost all industries and in aggregate. The only exception is the others industry, where *Sen* is still significantly positive but *Sub* is about equal to *BSOPM* owing to a particularly high α in this industry.

Given the large proportion of executive income that is attributed to illiquid, firm-specific options holdings, this finding suggests substantial overconfidence or positive inside information regarding their firm's future prospects. Indeed, the data show that the average executive prices ESOs such that the firm should outperform the market's expectations by an average of 12 % per annum (*Sen*). T-tests show that these values are significantly different from zero at the 1 % level in all industries and in aggregate.

Table 30.5 shows the mean and median values of R_t and *Sub* in each subsample, where R_t is the CAPM alpha. *Top* is a dummy variable taking value 1 if the executive works for a firm listed in *Fortune* magazine's top 100 companies to work for. Results show that firms with higher previous-year return tend to have significantly higher subjective values. This is true of both the mean and median value. Interestingly, subsequent return momentum is not consistently present in this data, at least as regards mean values. Firms listed in the top 100 in fact make significantly lower risk-adjusted returns in the year in which they are so listed. However, they enjoy substantially higher subjective value. This indicates that sentiment may generally be independent of performance but does significantly affect subjective value.

Figure 30.2 shows that relative subjective values are greater than one but relatively stable over time. In contrast, the number of issuances generally increases.

Table 30.5 Difference tests for subjective value and sentiment

	$R_{t-1} > 0$	$R_{t-1} < 0$	P-value	$Top = 1$	$Top = 0$	P-value
mean(R_t)	0.00042	0.00041	0.4493	0.00038	0.00059	<.0001
median(R_t)	0.00031	0.00028	0.0176	0.00030	0.00044	<.0001
mean(Sub)	21.4131	15.6336	<.0001	24.7420	17.6169	0.0003
median(Sub)	14.1350	10.9487	<.0001	17.4147	10.9794	<.0001

This table shows the mean and median values of R_t and Sub in each subsample, where R_t is the CAPM alpha at time t. Top equals 1 if the firm is listed as a top 100 firm by *Fortune* magazine in a given year. The p-values measure the significance of difference tests

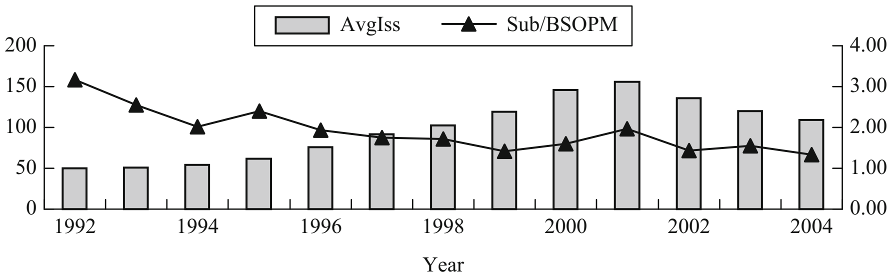


Fig. 30.2 Summary AvgIss and Sub/BSOPM by year. AvgIss and Sub/BSOPM for each year are graphed in this figure. AvgIss, BSOPM, and Sub are number of options issued, Black-Scholes value, and subjective value, respectively. The y-axis of the histogram is on the *left* and that of the line chart is on the *right*

The industry with the second highest subjective values (Financials) has a below-average number of issuances. These observations highlight the importance of looking at pricing, rather than issuance alone, as high subjective values do not imply that ESOs will be a more popular financing tool.

30.5.3 Implications of Regression Results and Variable Sensitivities

We now shift our attention to the testable implications of our model, namely confirming the relations between key options' variables and subjective value. Specifically, we apply the following regression equation:

$$\begin{aligned} Sub = Int + \beta_{\alpha}\alpha + \beta_{Sen}Sen + \beta_{Sok}Sok \\ +\beta_{\tau}\tau + \beta_{Vol}Vol + \beta_{Div}Div + \varepsilon, \end{aligned} \tag{30.7}$$

where *Int* is the intercept term and all variables are defined as before. Note that for all results presented here, the calculation of significance is via clustered standard errors by firm, though OLS results are nearly identical.

Table 30.6 Regression results for subjective value

	Int	α	*Sen*	*Sok*	τ	*Vol*	*Div*	Obs
$Sen = R_{t-1}$								
Coefficient	1.8988	−0.2126	0.0356	0.0361	−0.2813	−0.3983	−0.0783	56,602
(*p*-value)	(<.0001)	(<.0001)	(<.0001)	(0.2476)	(0.0489)	(<.0001)	(<.0001)	
Insider								
Coefficient	1.8650	−0.3001	0.0628	0.0860	−0.2883	−0.3409	−0.0844	23,826
(*p*-value)	(<.0001)	(<.0001)	(<.0001)	(0.0401)	(0.1712)	(0.0004)	(0.0017)	
True sentiment								
Coefficient	1.9071	−0.1252	0.0100	−0.0039	−0.2631	−0.4344	−0.0904	21,333
(*p*-value)	(<.0001)	(<.0001)	(<.0001)	(0.8583)	(0.2858)	(<.0001)	(<.0001)	
$Sen = Top$								
Coefficient	1.5014	−0.2244	0.0114	0.0802	−0.0560	−0.2318	−0.0807	49,090
(*p*-value)	(<.0001)	(<.0001)	(0.0169)	(0.6641)	(0.7216)	(<.0001)	(<.0001)	
$Y = 0$								
Coefficient	2.2013	−0.4364	0.0076	0.0424	−0.3265	−0.3872	−0.1012	82,374
(*p*-value)	(<.0001)	(<.0001)	(<.0001)	(0.1194)	(0.3041)	(<.0001)	(<.0001)	
Double exp								
Coefficient	2.1671	−0.4270	0.0017	0.0429	−0.3056	−0.3786	−0.1005	82,374
(*p*-value)	(<.0001)	(<.0001)	(<.0001)	(0.1155)	(0.3310)	(<.0001)	(<.0001)	
Bivariate con								
Coefficient	2.1799	−0.4267	0.0021	0.0428	−0.3075	−0.3883	−0.1023	82,374
(*p*-value)	(<.0001)	(<.0001)	(<.0001)	(0.1162)	(0.3284)	(<.0001)	(<.0001)	

This table presents the estimated coefficients from the following regressions:

$$Sub = Int + \beta_\alpha \alpha + \beta_{Sen} Sen + \beta_{Sok} Sok + \beta_\tau \tau + \beta_{Vol} Vol + \beta_{Div} Div + \varepsilon$$

where *Sub*, *Int*, α, *Sen*, *Sok*, τ, *Vol*, and *Div* refer to the subjective value, intercept term, proportion of total wealth held in illiquid firm-specific holdings, sentiment, ratio of stock price to exercise price, time to maturity, implied volatility, and dividend payout, respectively. In the first three tests, $Sen = R_{t-1}$, the CAPM alpha. We split the data into two groups according to the sign of the product of R_t and *Sen*. When *Sen* correctly forecasts the sign of the CAPM alpha for a given year, this is denoted as an "insider." When *Sen* and R_t do not match in sign, we denote this as "true sentiment." In the fourth test, *Sen* is a dummy variable that takes value 1 if the firm is in *Fortune's* top 100 and 0 otherwise. In the next three tests, *Sen* is calculated from European ESO formula (30.5), with the distribution of jump size following $y = 0$, a double exponential, and a bivariate constant jump model, respectively

First, we apply gross subjective value *Sub* as the dependent variable. The first three tests in Table 30.6 use CAPM risk-adjusted alpha from the year prior to option issuance R_{t-1} as a proxy for sentiment under the conjecture that those stocks that performed better in the previous year generate more positive sentiment prior to options being issued. Note that our model implies that only the risk-adjusted excess return should be priced because the market portion of the firm's return is eliminated via the risk-neutral measure. Bergman and Jenter (2007), in contrast, test the gross prior year return. Because a year's worth of data is required to calculate these alphas, the dataset is reduced to about 57,000 observations. We find that α is significantly negatively related to subjective value. This matches our intuition that the larger the proportion of one's portfolio held in options, the less diversified the portfolio, and the less valuable the ESO. *Sen*, on the other hand, is positively related, and significantly so. In other words, positive sentiment is associated with higher subjective value. Note that these results control for the usual options' pricing factors. Whereas *Div* is significantly negative related as expected, *Sok* and τ are not consistently significantly related, and *Vol* is negatively related. As explored more fully later, this last negative relation is quite telling and is consistent with our model as the sensitivity of subjective value to idiosyncratic risk is negative.

Further, the data are split into two groups according to the sign of the product of R_t and *Sen*. A positive (negative) sign implies that the positive sentiment measure is (not) accompanied by strong performance. The positive case (insider), then, can be explained by nonsentiment-related factors. The executive may have private inside information and hence be able to forecast future returns. They also have the ability to affect future returns so that optimism may be self-fulfilling. The negative case (true sentiment), on the other hand, has not such a concern inasmuch as it would imply that positive (negative) sentiment is followed by poor (good) performance. As it turns out, similar results are obtained in both cases: sentiment is positively related to subjective value while α is negatively related, both significantly so. As a result, it is not likely that insider information explains whole sentiment effect on subjective value.

Next, the *Top* dummy is selected as a proxy for sentiment. Once again, sentiment is significantly positively related to subjective value while α is significantly negatively related. All other relations are as above.

We also back *Sen* out of the European ESO formula (30.5) under the aforementioned three different jump size assumptions.[12] Because our model itself determines the relation between subjective value and *Sen*, the purpose of these tests is simply to observe the other variable relations as well as the stability of the model to the specification of the jump process. Results are quite consistent across the three processes tested here. All other coefficients remain qualitatively as before with the coefficient of α, importantly, remaining significantly negative in all cases.

Finally, in order to more clearly test the difference in impact of sentiment for insider versus true sentiment events, we interact the event identification dummy with our sentiment proxy as follows:

$$\begin{aligned} Sub &= Int + \beta_{\alpha}\alpha + \beta_{InSen}D_{In}Sen + \beta_{TSen}D_{T}Sen + \beta_{Sok}Sok + \beta_{\tau}\tau \\ &\quad + \beta_{Vol}Vol + \beta_{Div}Div + \varepsilon. \end{aligned} \tag{30.8}$$

All variables are defined as before and *Sen* is the previous-period CAPM alpha, also as before. D_{In} is a dummy variable that takes value 1 if the event is insider and

[12]Here, sentiment is estimated from the European option formula. It can also be calculated from the American option formula but with more exhaustive computations. As we mentioned before, sentiment estimated from the European and American ESO formulas have similar patterns. It may not affect the regression results much.

Table 30.7 Regression results for insider versus true sentiment events

	Int	α	$D_{In}Sen$	$D_T Sen$	*Sok*	τ	*Vol*	*Div*
$Sen = R_{t-1}$								
Coefficient	1.9070	−0.2214	0.0372	0.0044	0.0343	−0.2866	−0.3872	−0.0876
(p-value)	(<.0001)	(<.0001)	(<.0001)	(<.0001)	(0.3769)	(0.0876)	(<.0001)	(<.0001)

This table presents regression results for insider versus true sentiment events, where D_{In} is a dummy variable taking value 1 if the event is insider and 0 otherwise, D_T takes value 1 if it is true sentiment and 0 otherwise, and *Sen* is again defined as the CAPM alpha

0 otherwise. By analogy, D_T takes value 1 if the event is true sentiment and 0 otherwise. The results appear as in Table 30.7. Note that, while sentiment increases subjectively value significantly in both cases, the impact of sentiment when the event is likely to be an insider event is much larger. In other words, when strong prior performance reveals real information regarding future performance that may be known to managers, the impact on subjective value is strong. When prior performance proves not to be informative, the impact on subjective value is small. However, the impact is positive and significant in both cases.

30.5.4 Subjective Value and Risk

We now turn our attention to the sensitivity of subjective value to risk. While we note that our model implies a positive relation between total risk and subjective value, it further dictates that the sensitivity of subjective value to idiosyncratic risk is negative, a notion supported by our empirical findings. This indicates that increased levels of risk may negatively affect subjective value owing to the inability of executives to fully diversify their holdings. In contrast, the Black-Scholes as well as the majority of options pricing models prescribe no role to idiosyncratic risk, that is, the sensitivity should be zero, and are generally not be able to capture the empirical finding that subjective value is negatively related to risk.

In applying the empirical data to the formulae for the sensitivities of subjective value to various forms of risk, our model does indeed generate a negative relation between firm-specific risk and subjective value, a finding that is consistent also with the empirical observations of Meulbroek (2001). This finding is particularly important as managers can easily affect the firm's idiosyncratic risk level through various moral hazard-related activities.

In Table 30.8, risk sensitivities are calculated, vegas, for all options issues in our dataset assuming there are no illiquid holdings (UV), that is, $\alpha = 0$, and using our default value for α (V), with and without consideration of sentiment. The first two columns find as expected that the sensitivity with respect to total risk is positive, for both UV and V, regardless of whether sentiment is considered or not. This is true of all jump specifications. In every case, the sensitivity is higher when sentiment is not considered. Looking at the vegas with respect to jump frequency, UV(freq) can be either positive or negative depending on the jump specification, while V(freq) is always negative. Interestingly,

Table 30.8 Summary statistics for vega

Panel A: Y = 0						
	UV(total)	V(total)	UV(freq)	V(freq)	V(idio)	V(size)
Without sentiment						
Mean	11.237	13.715	12.787	−0.739	−24.212	−0.162
Median	9.173	11.684	10.159	−0.143	−19.611	−0.077
With sentiment						
Mean	5.968	7.646	16.486	−11.222	−41.638	−0.415
Median	3.007	4.906	13.191	−2.857	−31.063	−0.144
Panel B: Double exponential jump model						
	UV(total)	V(total)	UV(freq)	V(freq)	V(idio)	V(size)
Without sentiment						
Mean	13.456	16.791	−1.039	−0.564	−25.877	−0.017
Median	10.875	14.129	−0.871	−0.449	−20.396	−0.008
With sentiment						
Mean	7.397	10.063	−14.217	−1.213	−44.233	−0.042
Median	4.173	7.041	−1.823	−0.806	−32.466	−0.014
Panel C: Bivariate constant jump model						
	UV(total)	V(total)	UV(freq)	V(freq)	V(idio)	V(size)
Without sentiment						
Mean	13.366	16.792	−1.194	−0.677	−25.860	−0.017
Median	10.879	14.131	−1.015	−0.542	−20.382	−0.008
With sentiment						
Mean	7.396	10.060	−15.984	−1.407	−44.231	−0.042
Median	4.173	7.037	−2.103	−0.957	−32.468	−0.014

This table presents test results for vega. In Panels A, B, and C, the distribution of jump sizes are zero jump, double exponential jump, and bivariate constant jump, respectively. UV(total) and UV (freq) are total risk and jump frequency risk vegas under our model when all holdings are liquid. V(total), V(idio), V(freq), and V(size) refer to total risk vega, idiosyncratic risk vega, jump frequency risk vega, and jump size risk vega, respectively

UV is positive for the constant jump model but negative for the other two models, pointing out the importance of jump specification when liquidity is not also considered. The magnitude of UV is always smaller than that of V.

Perhaps the most interesting factor affecting our subjective value in our model is idiosyncratic risk, for which the estimate is always negative and is significantly larger in magnitude than the other vegas. While the jump size vega also plays a role and is likewise always negative, the magnitude of this effect is much smaller. This finding highlights the role of idiosyncratic risk in our model and explains why the empirical sensitivity of subjective value to volatility is found to be negative, contrary to generally accepted moral hazard models that dictate that option compensation encourages risk taking. If agents are sufficiently under-diversified, the risk premium from taking on excess idiosyncratic risk offsets gains from convexity and discourages risk-taking behavior. The corresponding UVs for idiosyncratic and jump size risk are both zero as these do not play a role in determining market value when there are no under-diversified holdings. Also, the Vs are substantially more negative when sentiment is introduced, pointing out the sharply offsetting effects of positive sentiment and risk aversion in this model. Which piece dominates then depends on the risk aversion parameter and α of the employee.

30.6 Conclusion

This chapter applies three statistics methods in ESO study, including change of measure, hierarchical clustering with a K-means approach, and estimation of standard errors in finance panel data. We use a model for employee stock options that incorporates illiquidity of the options, a jump diffusion for the stock price evolution that includes various jump processes, and the potential roles of employee sentiment and insider information in a world where employees balance their wealth among the company's stock, the market portfolio, and a risk-free asset. Our option contract is American type and the optimal exercise boundary is derived endogenously. From the ESO pricing formula, we are able to not only estimate the subjective values but also study the exercise policies.

The subjective value placed on ESOs implied by compensation data is calculated by applying empirical data. Specifically, using data provided by Compustat, executives are grouped by a hierarchical clustering with a K-means approach based on a number of firm and individual criteria. By assuming that all executives in the same cluster receive the same total compensation, a notion that relies on the existence of competitive labor markets, we then back out the valuation placed by each executive on their respective ESO. These groups include consideration of nonoption compensation, rank, industry, year, firm size, and immediate exercise value. Though the extant literature predicts that employees should discount the value of their options, we find that executives in fact value their options more highly than implied by Black-Scholes, applying an average premium of 48 %. As such, the cost of issuance for the firm is vastly lower than the benefit perceived by employees, suggesting that ESO compensation should be an even larger part of executive compensation. We then relate subjective value to sentiment levels and generate the novel finding that executives must expect their firm's risk-adjusted returns to outpace that predicted by the market by 12 % in order to justify the subjective value placed on ESOs. This expectation may be the result of private information regarding the growth prospects of the firm.

Testing subjective value and its relation to pertinent variables, subjective value is negatively related to the proportion of wealth held in illiquid firm-specific holdings and positively related sentiment. In other words, the larger the illiquid ESO position is, the larger the discount risk aversion prescribed and the lower the subjective value implied in the compensation package. On the other hand, the more positive the employee's view of future risk-adjusted returns, the more valuable the ESO. Interestingly, subjective value may be negatively related to risk as the inability of executives to fully diversify their holdings may lead to risk premia that outweigh the value placed on risk by the convexity of options payouts. Note that this relation is particularly negative with regard to idiosyncratic risk and is empirically also negative for risk associated with both jump frequency and size. Because these aspects of return are precisely those that may be most directly controlled by executives, traditional moral hazard arguments relating solely to the convexity of the options payout may not hold.

Firms increasingly grant nontraditional employee stock options to link stock price performance and managerial wealth and provide greater incentives to employees. While this study focuses on the traditional employee stock option, the main intuition can be involved in nontraditional ESOs. Premium stock option, performance-vested stock option, repriceable stock option, purchased stock option, reload stock option, and index stock option are the objects of future study. We can derive the option formulas and compare the value, incentive effect, and cost per unit of subjective incentive across the nontraditional ESOs and the traditional ones. This future study provides a firm a proper compensation vehicle according to its characteristics.

Appendix 1: Derivation of Risk-Neutral Probability by Change of Measure

Here we introduce the idea for deriving the ESO formulas and define using probability measure P^* by the technique of change of measure.

The process of an employee's marginal utility or the pricing kernel is:

$$\frac{dJ_W}{J_W} = -\hat{r}dt - \hat{\sigma}dW_m + \hat{v}dW_s + \left(\hat{Y} - 1\right)dN_t,$$

where

$$\begin{aligned}
\hat{r} &= r - (1-\gamma)\left(\alpha^2 v^2 + \frac{1}{2}\alpha^2\gamma\lambda k_2 + \alpha\lambda k\right),\\
\hat{\sigma} &= \frac{\mu_m - r}{\sigma_m},\\
\hat{v} &= -(1-\gamma)\alpha v,\\
\hat{Y} &= [\alpha(Y-1)+1]^{\gamma-1},
\end{aligned}$$

then

$$J_W[W(t),t] = J_W[W(0),0]\exp\left\{\left(-\hat{r} - \frac{1}{2}\hat{\sigma}^2 - \frac{1}{2}\hat{v}^2\right)t +, \hat{\sigma}W_m(t) + \hat{v}W_s(t)\right\}\prod_{i=0}^{N_t}\hat{Y}_i.$$

Let $B(t, T)$ be the price of a zero coupon bond with maturity T, then

$$\begin{aligned}
B(t,T) &= E\left\{\frac{J_W[W(T),T]}{J_W[W(t),t]}B(T,T)|\mathcal{F}_t\right\}\\
&= E\left\{\exp\left\{\left(-\hat{r} - \frac{1}{2}\hat{\sigma}^2 - \frac{1}{2}\hat{v}^2\right)\tau + \hat{\sigma}W_m(\tau) + \hat{v}W_s(\tau)\right\}\prod_{i=0}^{N_\tau}\hat{Y}_i\right\}\\
&= \exp\{[-\hat{r} + \lambda(\xi - 1)]\tau\},
\end{aligned}$$

where

$$\tau = T - t, \;\; \xi = E\left(\hat{Y}\right) = E\left\{[\alpha(Y-1)+1]^{\gamma-1}\right\}.$$

The bond yield

$$\begin{aligned} r^* &\equiv -\frac{1}{T-t}\ln B(t,T) \\ &= r - (1-\gamma)\left(\alpha\lambda k + \frac{1}{2}\gamma\lambda k_2\alpha^2 + \alpha^2 v^2\right) - \lambda(\xi - 1). \end{aligned}$$

Let

$$Z(t) \equiv e^{r*t}J_W = \exp\left\{\left[-\frac{1}{2}\hat{\sigma}^2 - \frac{1}{2}\hat{v}^2 - \lambda(\xi-1)t + \hat{\sigma}W_m(t)\right] + \hat{v}W_s(t)\right\}\prod_{i=0}^{N_t}\hat{Y}_i,$$

then $Z(t)$ is a martingale under P, and we have

$$\frac{J_W[W(T),T]}{J_W[W(t),t]} = e^{-r*(T-t)}\frac{Z(T)}{Z(t)}.$$

The rational equilibrium value of the ESO at time t, $F(S_t, t)$, satisfies the Euler equation,

$$\begin{aligned} F(S_t,t) &= E_t\left\{\frac{J_W[W(T),T]}{J_W[W(t),t]}F(S_T,T)\right\} \\ &= e^{-r*(T-t)}E_t^*[F(S_T,T)], \end{aligned}$$

where $\frac{dP^*}{dP} = \frac{Z(t)}{Z(0)}$, and E_t^* is the expectation under P^* and information at time t. Under the probability measure P^*, the processes $W_m^* = W_m - \hat{\sigma}t$ and $W_s^* = W_s - \hat{v}t$ are Brownian motions, N_t is a Poisson process with intensity $\lambda^* = \lambda\xi$, and the jump sizes follow density $f_Y^*(y)$,

$$f_Y^*(y) = \frac{1}{\xi}[\alpha(y-1)+1]^{\gamma-1}f_Y(y).$$

Therefore,

$$\begin{aligned} \frac{dS}{S} &= \mu_s dt + \sigma_s dW_m + v dW_s + (Y-1)dN_t \\ &= \left[r - d - (1-\gamma)\alpha v^2 - \lambda k\right]dt \\ &\quad + \sigma_s dW_m^* + v dW_s^* + (Y-1)dN_t \\ &\equiv (r^* - d^*)dt + \sigma_N dW_t^* + \left(Y-1\right)dN_t, \end{aligned}$$

where

$$
\begin{aligned}
d^* &= d - (1-\gamma)\left[\alpha\lambda k + \frac{1}{2}\gamma\lambda k_2\alpha^2 - (1-\alpha)\alpha v^2\right] \\
&\quad - \lambda(\xi - 1) + \lambda k, \\
\sigma_N^2 &= \sigma_s^2 + v^2, \\
\sigma_N W_t^* &= \sigma_s W_m^* + v W_s^*.
\end{aligned}
$$

In other words,

$$
S_t = S_0 \exp\left\{\left(r^* - d^* - \frac{1}{2}\sigma_N^2\right)t + \sigma_N W_t^*\right\}\prod_{i=0}^{N_t} Y_i.
$$

Appendix 2: Valuation of European ESOs

The option price at time t is

$$
\begin{aligned}
C_E(S_t,\tau) &= e^{-r*(T-t)}E_t^*\left\{[S_T - K]^+\right\} \\
&= e^{-r*\tau}E_t^*\left\{\left\{S_t\exp\left[\left(r^* - d^* - \tfrac{1}{2}\sigma_N^2\right)\tau + \sigma_N\left(W_T^* - W_t^*\right)\right]\times\prod_{i=N_t}^{N_T} Y_i - K\right\}^+\right\} \\
&= S_t E^*\left\{\exp\left[\left(-d^* - \tfrac{1}{2}\sigma_N^2\right)\tau + \sigma_N W_\tau^*\right]\right. \\
&\quad \left.\times\prod_{i=0}^{N_\tau} Y_i I_{\left(W_\tau^* \geq a_1\right)}\right\} - Ke^{-r*\tau}E^*\left\{I_{\left(W_\tau^* \geq a_1\right)}\right\} \\
&= S_t e^{-d*\tau}E^*\left\{E^*\left\{\exp\left[-\tfrac{1}{2}\sigma_N^2\tau + \sigma_N W_\tau^*\right]\times\prod_{i=0}^{N_\tau} Y_i I_{\left(W_\tau^* \geq a_1\right)}\middle|\prod_{i=0}^{N_\tau} Y_i\right\}\right\} \\
&\quad - Ke^{-r*\tau}E^*\left\{E^*\left[I_{\left(W_\tau^* \geq a_1\right)}\middle|\prod_{i=0}^{N_\tau} Y_i\right]\right\} \\
&= S_t e^{-d^*\tau}E^*\left\{\prod_{i=0}^{N_\tau} Y_i\Phi\left(-\frac{a_1 - \sigma_N\tau}{\sqrt{\tau}}\right)\right\} - Ke^{-r*\tau}E^*\left\{\Phi\left(-\frac{a_1}{\sqrt{\tau}}\right)\right\} \\
&= \sum_{j=0}^{\infty}\frac{(\lambda\xi\tau)^j e^{-\lambda\xi\tau}}{j!}\left\{S_t e^{-d^*\tau}E^*\left[\prod_{i=0}^{j} Y_i\Phi\left(d_1^*\right)\right] - Ke^{-r*\tau}E^*\left[\Phi\left(d_2^*\right)\right]\right\}
\end{aligned}
$$

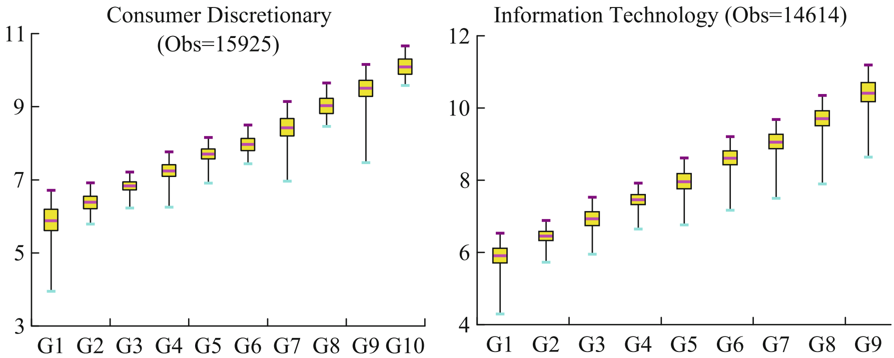


Fig. 30.3 Natural log of total compensation. The box plots show the natural log of total compensation for the two largest industries in our sample. Executives are grouped according to position, the firm's total market value, nonoption compensation, and the immediate exercise value of the options for each industry by using a hierarchical clustering with K-means approach

where

$$d_1^* = \frac{\ln\left[\left(S_t \prod_{i=0}^{j} Y_i\right)/K\right] + \left(r^* - d^* + \frac{1}{2}\sigma_N^2\right)\tau}{\sigma_N\sqrt{\tau}},$$
$$d_2^* = d_1^* - \sigma_N\sqrt{\tau}, a_1 = -d_2^*\sqrt{\tau}.$$

Appendix 3: Hierarchical Clustering with a K-Means Approach

The compensation-based subjective value is calculated by assuming that all executives within the same cluster receive the same total compensation. For each executive in this cluster, the implied subjective value is derived by comparing the difference between nonoption compensation and the average compensation. It is very important to group executives appropriately. We use cubic clustering criterion to generate number of groups and split executives according to position, the firm's total market value, nonoption compensation, and the immediate exercise value of the options for each industry by hierarchical clustering with a K-means approach. For each group, we calculate the average total compensation for all executives. If all executives receive the same compensation, on average, any differences in salaries, bonuses, and other income should be accounted for by options.

Figure 30.3 presents box plots of the natural log of total compensation for the two largest industries in our sample: Consumer Discretionary and Information Technology. With the exception of some outliers, which are subsequently removed in our main tests, the boxed areas generally do not overlap from cluster to cluster, demonstrating the relative homogeneity of firms within each cluster and generally distinctly separated from other clusters. As a result, we believe that compensation characteristics within each cluster should be quite comparable, lending a measure of credence to our method of calculating subjective value.

References

Aggarwal, R. K., & Samwick, A. A. (2003). Performance incentives within firms: The effect of managerial responsibility. *Journal of Finance, 58*, 1613–1649.

Agrawal, A., & Mandelker, G. N. (1987). Managerial incentives and corporate investment and financing decisions. *Journal of Finance, 42*, 823–837.

Bergman, N. K., & Jenter, D. (2007). Employee sentiment and stock option compensation. *Journal of Financial Economics, 84*, 667–712.

Bettis, C. J., Bizjak, J. M., & Lemmon, M. L. (2005). Exercise behavior, valuation, and the incentive effects of employee stock options. *Journal of Financial Economics, 76*, 445–470.

Bryan, S., Hwang, L., & Lilien, S. (2000). CEO stock-based compensation: An empirical analysis of incentive-intensity, relative mix and economic determinants. *Journal of Business, 73*, 661–693.

Carpenter, J. N. (2000). Does option compensation increase managerial risk appetite? *Journal of Finance, 55*, 2311–2331.

Chang, C., Fuh, C. D., & Hsu, Y. H. (2008). ESO compensation: The role of default risk, employee sentiment, and insider information. *Journal of Corporate Finance, 14*, 630–641.

Chang, C., Chen, L. J., & Fuh, C. D. (2013). The pricing of risk and sentiment: A study of executive stock options. *Financial Management, 42*, 79–99.

Chen, L. J., & Fuh, C. D. (2011). A study of employee stock options and the exercise decision. *International Research Journal of Finance and Economics, 67*, 142–159.

Core, J. E., & Guay, W. (2001). Stock option plans for non-executive employees. *Journal of Financial Economics, 61*, 253–287.

Dittmann, I., & Maug, E. (2007). Lower salaries and no options? On the optimal structure of executive pay. *Journal of Finance, 62*, 303–343.

Duffee, G. R. (1999). Estimating the price of default risk. *The Review of Financial Studies, 12*, 197–226.

Fruhwirth, M., & Sogner, L. (2006). The Jarrow/Turnbull default risk model-evidence from the German market. *European Journal of Finance, 12*, 107–135.

Gabaix, X., & Landier, A. (2008). Why has CEO pay increased so much? *Quarterly Journal of Economics, 123*, 49–100.

Gukhal, C. R. (2001). Analytical valuation of American options on jump-diffusion processes. *Mathematical Finance, 11*, 97–115.

Hall, B. J., & Murphy, K. J. (2002). Stock options for undiversified executives. *Journal of Accounting and Economics, 33*, 3–42.

Hemmer, T., Matsunaga, S., & Shevlin, T. (1996). The influence of risk diversification on the early exercise of employee stock options by executive officers. *Journal of Accounting and Economics, 21*, 45–68.

Hemmer, T., Kim, O., & Verrechia, R. K. (2000). Introducing convexity into optimal compensation contracts. *Journal of Accounting and Economics, 28*, 307–327.

Hodge, F. D., Rajgopal, S., & Shevlin, T. (2009). Do managers value stock options and restricted stock consistent with economic theory? *Contemporary Accounting Research, 26*, 899–932.

Holland, L. C., & Elder, E. M. (2006). Employee stock options in compensation agreements: A financing explanation. *Journal of Corporate Finance, 12*, 367–379.

Huddart, S., & Lang, M. (1996). Employee stock option exercises: An empirical analysis. *Journal of Accounting and Economics, 21*, 5–43.
Hull, J., & White, A. (2004). How to value employee stock options. *Financial Analysts Journal, 60*, 114–119.
Ingersoll, J. E. (2006). The subjective and objective evaluation of incentive stock options. *The Journal of Business, 79*, 453–487.
Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: A review. *ACM Computing Surveys, 31*, 264–323.
Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs, and ownership structure. *Journal of Financial Economics, 3*, 305–360.
Lambert, R. A., Larcker, D. F., & Verrecchia, R. E. (1991). Portfolio considerations in valuing executive compensation. *Journal of Accounting Research, 29*, 129–149.
Liao, F. Y., & Lyuu, Y. D. (2009). An expanded model for the valuation of employee stock options. *Journal of Futures Markets, 29*, 713–735.
Meulbroek, L. K. (2001). The efficiency of equity-linked compensation: Understanding the full cost of awarding executive stock options. *Financial Management, 30*, 5–44.
Nohel, T., & Todd, S. (2005). Compensation for managers with career concerns: The role of stock options in optimal contracts. *Journal of Corporate Finance, 11*, 229–251.
Oyer, P., & Schaefer, S. (2005). Why do some firms give stock options to all employees? An empirical analysis of alternative theories. *Journal of Financial Economics, 76*, 99–133.
Petersen, M. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies, 22*, 435–480.
Ryan, H. E., & Wiggins, R. A. (2001). The influence of firm- and manager-specific characteristics on the structure of executive compensation. *Journal of Corporate Finance, 7*, 101–123.
Shreve, S. E. (2008). *Stochastic calculus for finance II: Continuous-time models*. New York: Springer.

Structural Change and Monitoring Tests 31

Cindy Shin-Huei Wang and Yi Meng Xie

Contents

Abstract

This chapter focuses on various structural change and monitoring tests for a class of widely used time series models in economics and finance, including $I(0)$, $I(1)$,

C.S.-H. Wang (✉)
CORE, Université Catholique de Louvain and FUNDP, Academie Louvain, Louvain-la-Neuve, Belgium

Department of Quantitative Finance, National TsingHwa University, Hsinchu City, Taiwan
e-mail: cindywang9177@gmail.com

Y.M. Xie
School of Business and Administration, Beijing Normal University, Beijing, China

Department of Economics, University of Southern California, Los Angeles, CA, USA
e-mail: yimengxie69@gmail.com

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_31,

$I(d)$ processes and the cointegration relationship. A structural break appears a change in endogenous relationships. This break could be caused by a shift in mean, variance, or a persistent change in the data property. In general, structural change tests can be categorized into two types: one is the classical approach to testing for structural change, which employs retrospective tests using a historical data set of a given length; the other one is the fluctuation-type test in a monitoring scheme, which means for given a history period for which a regression relationship is known to be stable, we then test whether incoming data are consistent with the previously established relationship. Several structural changes such as CUSUM squared tests, the QLR test, the prediction test, the multiple break test, bubble tests, cointegration breakdown tests, and the monitoring fluctuation test are discussed in this chapter, and we further illustrate all details and usefulness of these tests.

Keywords
Cointegration breakdown test • Structural break • Long-memory process • Monitoring fluctuation test • Boundary function • CUSUM squared test • Prediction test • Bubble test • Unit root time series • Persistent change

31.1 Introduction

The structural break refers to the phenomenon illustrating the time series comes across unanticipated and significantly influential shifts. Generally, most macroeconomic and financial time series are subject to occasional structural breaks. Ignoring the presence of structural breaks can lead to seriously biased estimates and forecasts and the unreliability of models (see Clements and Hendry 1998, 1999). Since economic activities often experience rather long period, there exist plausible possibilities that structural breaks could occur. Hence, it is very important to monitor and modify the changes of the model characteristics. If those model characteristics were not corrected in *real time*, it may lead to wrong decisions on the policy making and investment plans for policy makers and practitioners. Consequently, those wrong decisions could result in the fluctuations or crises in economy and finance.

Examples in the real world are a large sum of financial crises ever since before the Industrial Revolution. For instance, ten enormous bubbles took place ahead of the twenty-first century documented by Kindleberger (2005), and the latest subprime crisis occurred in 2008. All crises crushed the wealth of the whole world again and again. After the first reported financial crisis – Dutch tulip bulb bubble in 1636 – all meltdowns in the economy systems were obviously characterized by the deviation from steadily development.

It implies that structural breaks could happen prior to the bubbles or explosions. Accordingly, an accurate approach to detecting structural breaks correctly can be treated as a measurement to notice the possible incoming financial crisis or to monitor market activities. On the other hand, developing a suitable structural break-testing mechanism is crucial to regulators and investors to modify decisions in order to avoid a potential incoming crash.

The first seminal paper of the issue regarding to the structural break tests was proposed by Chow (1960). Nowadays, structural change tests have been explored far more than the original Chow test and developed in many different dimensions. In this chapter, we survey various structural changes and monitoring tests for a class of widely used time series models in economics and finance, including $I(0)$, $I(1)$, $I(d)$ processes, the cointegration breakdown test, and bubble tests. The conventional structural change tests are categorized into three types: the mean shift, the persistent change, and the parameter change. Since the issues of break numbers and properties of the data-generating process (DGP) would be the main factors to affect the performances of the structural tests, it is essential to design the structural tests with respect to different scenarios.

The rest of this chapter is organized as follows. Section 31.2 illustrates the changes in mean and coefficients when DGP are stationary $I(0)$ and stationary long-memory $I(d)$ processes. Section 31.3 focuses on persistent change systematically. Tests of bubbles and the cointegration breakdown are discussed in Sect. 31.4. Section 31.5 provides the monitoring tests in *real time*.

31.2 Structural Breaks in Parameters

Structural break has been an intriguing issue since Chow (1960) test resolves how to detect a single structural change with known break date, and more ensuing literatures expand that in different dimensions. Break numbers can be multiple while break dates are no longer required as necessitates. Moreover, econometricians apply research closely to reality, adjusting these models from classical stationary process to nonstationary and long memory, which are found more likely to reflect economic and financial data. Now, we will offer systematical analysis of these marked developments.

31.2.1 Stationary Processes

It is the Chow (1960) test that firstly detect a potential structural break for a given break date. Suppose the model is

$$y_t = \beta_t' X_{t-1} + \varepsilon_t', \tag{31.1}$$

where our null hypothesis is $\beta_t = \beta$ for all t, ε_t is a martingale difference sequence with respect to the $\sigma-$ fields, and X_t is a $k \times 1$ vector of regressor, which are here assumed to be constant and/or I(0) with $EX_tX_t' = \Sigma_X$ and, possibly, a nonzero mean. Particularly, X_{t-1} can include lagged dependent variables as long as they are I(0) under the null. When the break date is unknown, a natural solution proposed by Quandt (1960) calculates break date, and it is extended by Davies (1977) to general models with parameters unidentified under the null. And Andrews (1993) reports asymptotic critical values of this QLR statistics. in the form of Quandt likelihood ratio (QLR) statistic.

To the contrast of the null, the alternative hypothesis of a single break in some or all of the coefficients is

$$\beta_t = \beta, \quad t \leq r \quad \text{and} \quad \beta_t = \beta + \gamma, \;\; t > r, \tag{31.2}$$

where r, $k + 1 < r < T$, is the "break date" (or "change point") and $\gamma \neq 0$.

When the potential break date is known, a natural test for change in β is the Chow (1960) test, which can be implemented in asymptotically equivalent Wald, Lagrange multiplier (LM), and LR forms. In the Wald form, the test for a break at a fraction r/T through the sample is

$$F_T(r/T) = \frac{SSR_{1,T} - \left(SSR_{1,r} + SSR_{r+1,T}\right)}{\left(SSR_{1,r} + SSR_{r+1,T}\right)/(T - 2k)}, \tag{31.3}$$

where $SSR_{1,\ r}$ is the sum of squared residuals from the estimation of Eq. 31.1 on observation 1,. . ., r, etc. For fixed r/T, $F_T(r/T)$ has an asymptotic χ_k^2 distribution under the null. However, when break date is unknown, the situation can be more complicated. An expectable idea is to estimate the break date and then compute Eq. 31.3 for that break. But since the change point is selected by virtue of an apparent break at the point, the null distribution of the resulting test is not the same as if the break date were chosen without regard to the data. Thus means of determining r/T must be further specified before the distribution of the resulting test can be obtained. Quandt (1960) and Davies (1977) come up with another F_T statistics called Quandt likelihood ratio, thus QLR statistic:

$$QLR = \underset{r=r_0,\ldots,r_1}{\text{MAX}} \quad F_T(r) \tag{31.4}$$

Intuitively, it has power against a change in β even without a known break date. Although it confuses econometricians for years to calculate its null asymptotic distribution, Andrews (1993, Table I) reports computable critical values for it. For more details, please see Stock (1994).

31.2.2 Multiple Breaks Case

This issue of detecting structural breaks becomes even richer after the possible multiple breaks are considered no matter the numbers and dates of breaks are known or not. Bai and Perron (1998) proposed a comprehensive treatment for estimating linear models with multiple structural breaks. They set the multiple linear regression with m breaks (m + 1 regimes) as follows:

$$y_t = x_t'\beta + z'\delta_j + u_t, \tag{31.5}$$

where $j = 1, \cdots, m + 1, T_0 = 0$, and $T_{m+1} = T$; y_t is the observed independent variable, $x_t(p \times 1)$, and $z_t(q \times 1)$ are vectors of covariates, and β and $\delta_j (j = 1, \cdots, m{+}1)$ are the corresponding vectors of coefficients; u_t is the disturbance. Note this is a partial structural change model in the sense that β is not subject to shifts and is effectively estimated using the entire sample. When $p = 0$, however, it is translated to a pure structural change. Now rewrite the liner regression system in Eq. 31.5 in matrix form as $Y = X\beta + \overline{Z} + U$, where $Y = (y_1, \ldots, y_T)'$, $X = (x_1, \ldots, x_T)'$, $U = (u_1, \ldots, u_T)'$, $\delta = (\delta_1', \delta_2', \ldots, \delta_{m+1}')'$, and $\overline{Z}$ is the matrix which diagonally partitions Z at the m-partition $(T_1, \ldots, T_m)$s, i.e., $\overline{Z} = \text{diag}(Z_1, \ldots, Z_{m+1})$ with $Z_i = \left(Z_{T(i-1)+1}, \ldots, Z_{T_i}\right)'$. Following this, apply associated least-squares estimates of β, and δ are obtained by minimizing the sum of squared residuals. And let $\hat{\lambda} = \left(\hat{\lambda}_1, \ldots, \hat{\lambda}_m\right) = \left(\hat{T}_1/T\hat{T}_m/T\right)$ with corresponding true value $\lambda^0 = (\lambda_1^0, \ldots, \lambda_m^0)$. Thus take the sup F type test of no structural break $(m = 0)$ versus the alternative hypothesis that there are $m = k$ breaks. Let $(T_1, \ldots, T_k)$ be a partition such that $T_i = [T\lambda_i](i = 1, \ldots, k)$. Operationally, to test a null of no break versus some fixed number of breaks, define

$$F_T(\lambda_1, \ldots, \lambda_k;\ q) = \frac{(T - (k+1)q - p)}{kq} \frac{\hat{\delta}' R' \left(R\left(\overline{Z}' M_x \overline{Z}\right)^{-1} R'\right)^{-1} R\hat{\delta}}{SSR_k} \tag{31.6}$$

where R is the conventional matrix such that $(R\delta)' = (\delta_1' - \delta_2', \ldots, \delta_k' - \delta_{k+1}')$ and $M_x = I - X(X'\ X)'\ X'$. Here SSR_k is the sum of squared residuals under the alternative hypothesis, which depends on $(T_1, \ldots, T_k)$. Define the following set for some arbitrary small positive number ϵ: $\Lambda_\epsilon = \{(\lambda_1, \ldots, \lambda_k);\ |\lambda(i+1) - \lambda_i| \geq \epsilon, \lambda_1 \geq \epsilon, \lambda_k \leq 1 - \epsilon\}$. The reason for this is to restrict each break date to be asymptotically distinct. And apparently, it is the generation of the sup F test considered by Andrews (1993) and others for the case $k = 1$.

Similarly, estimated break points by testing for l versus $l + 1$ breaks are also obtained by a global minimization of the sum of squared residuals. Comparing the difference between the sum squared residuals obtained with l breaks and that obtained with $l + 1$ breaks, it is advised to test each $(l + 1)$ segment for the presence of an additional break. This requires that the magnitude of shifts is fixed (nonshrinking). More precisely, the test is defined by

$$F_T\left(l+1|l\right) = S_T\left(\hat{T}_1, \ldots, \hat{T}_l\right) - \min_{1 \leq i \leq l+1} \inf_{\tau \in \Lambda_{i,\eta}} S_t\left(\hat{T}_1, \ldots, \hat{T}_{i-1}, \tau, \hat{T}_i, \ldots, \hat{T}_l\right)/\hat{\sigma}^2, \tag{31.7}$$

where

$$\Lambda_{i,\eta} = \left\{\tau; \hat{T}_{i-1} + \left(\hat{T}_i - \hat{T}_{i-1}\right)\eta \leq \tau \leq \hat{T}_i - \left(\hat{T}_i - \hat{T}_{i-1}\right)\eta\right\} \tag{31.8}$$

and $\hat{\sigma}^2$ is a consistent estimate of σ^2 under the null hypothesis. Note that for $i = 1$, $S_T\left(\hat{T}_1, \ldots, \hat{T}(i-1), \tau, \hat{T}_i, \ldots, \hat{T}_l\right)$ is understood as $S_T\left(\tau, \hat{T}_1, \ldots, \hat{T}_l\right)$ and for $i = l + 1$ as $S_T\left(\hat{T}_1, \ldots, \hat{T}_l, \tau\right)$.

As well, Bai and Perron (1996) also provide a valid proof for consistency of the estimated break fractions along with its limiting distributions. They also apply their model to the situations where autocorrelation could happen and compare it with sequential estimation.

31.2.3 Prediction Tests

Prediction tests which are often applied to forecast stock and exchange rate also lose power when potential structural change is incorporated without reasonable adjustment. While numerical literatures explore suitable models to fit the reality, their efforts still fail for rather long period or large forecasting sample as proven in Welch and Goyal (2008). This means tests for structural stability are equivalently vital in forecasting. For time series satisfying characters of stationarily I(0), prediction tests are operated on the difference between observed variables and predictions and only when it is before the suspected structural change could such tests be valid. Lutkepohl (1989) proposes a prediction tests for structural stability of univariate times series and compares its efficacy to the univariate time series investigated by Lutkepohl (1988).

Lutkepohl (1989) assumes a multiple time series generated by a stationary vector stochastic process $y_t = (y_{1t}, \ldots, y_{kt})$ in autoregressive (AR) representation

$$y_t = v + \sum_{i=0}^{\infty} A_i y_{t-i} + u_i \tag{31.9}$$

and moving average (MA) representation

$$y_t = \mu + \sum_{i=0}^{\infty} \Phi_i u_{t-i}, \Phi_0 = I_k, \tag{31.10}$$

Noted that I_k is the (K $\times$ K) identity matrix, v is a K-dimensional vector of constant terms, μ is the mean vector of y_t, the A_i and Φ_1 are (K $\times$ K) coefficient matrices, and $u_t = (u_{1t}, \ldots, u_{Kt})'$ is K-dimensional white noise. In other words, the $\widetilde{t}$ is i.i.d. multivariate normal, $u_t \sim N(0, \sum_u)$. Although the two present models focus on infinity, it is permitted that they can both be finite-order models. Thus the mean squared error (MSE) h-step predictor at origin is

$$y_t(h) = \mu = \sum_{i=h}^{\infty} \Phi_i u_{t+h-i},$$

and its MSE matrix is

$$\Sigma(h) = \sum_{i=0}^{h-1} \Phi_i \Sigma_u \Phi'.$$

Since y_t is Gaussian, the forecast error vector is normally distributed,

$$e(h) = y_{t+h} - y_t(h)(0, \Sigma(h)). \tag{31.11}$$

Therefore, the rest statistics

$$T(h) = e(h)^{'}\Sigma(h)^{-1}e(h) \sim \chi^2(K) \tag{31.12}$$

for $h = 1, 2, \ldots$, can be used for testing whether a structural change has occurred after period t. As usual $\chi^2(K)$ denotes the chi-squared distribution with Kdf.

Besides the univariate description of this system, another test relevant for multivariate systems is also designed by Lutkepohl (1989). Since the vector of one-step forecast errors to h-step forecast errors is also normally distributed,

$$e(h) = [e(1)(h)] \sim N(0, \Sigma(h)).$$

Here

$$\Sigma(h) = \Phi_h(I_h \otimes \Sigma_u)\Phi_h 6', \tag{31.13}$$

where $\otimes$ denotes the Kronecker product and

$$\Phi_h = \left[I_k\dot{0}\Phi_1 I_k\dot{0}\Phi_{h-1}\Phi_{h-2}\dot{I}_k\right] \tag{31.14}$$

Hence,

$$L(h) = e(h)'\Sigma(h)^{-1}e(h) \sim \chi^2(hK) \tag{31.15}$$

for $h = 1, 2, \ldots$, is another sequence of statistics that can be used for stability tests. This $L(h)$ statistics are particularly easy to compute and according to Lutkepohl (1989), while univariate time series tests can be relevant for multivariate systems, a structural change in a multivariate system will not be necessarily be reflected in the univariate models for the individual variables of the system. And Lutkepohl (1989) points that since the coefficients Φ_i, A_i, and Σ_u are usually unknown so that the statistics cannot be computed, the replacement by estimators is needed. Thus the estimated processes $\hat{T}(h) \to \chi^2(K)$ and $\hat{L}(h) \to \chi^2(hK)$ can match the null hypothesis of structural change. Generally, Lutkepohl (1989) proposes a prediction test for checking the structural stability of a system of time series. Despite the fact it lacks power in modelling the form of the change, it serves as a useful tool without imposing severe restrictions on the data. Another merit is that very few data are required after the time period or the time point at which a structural change is suspected.

31.2.4 Structural Changes in Long-Memory Processes

So far long memory has been recognized as one of most common features shared by macroeconomic and financial time series. Let X_t be a stationary process with autocovariances $r(k) = \mathrm{cov}(X_t, X_{t+k})$, then X_t is said to have long memory if as $|k| \to \infty$,

$$r(k) \sim L_1(k)|k|^{2H-2}, \ \ H \in \left(\frac{1}{2}, 1\right)$$

where $L_1(k)$ is a slowly varying function as $|k| \to \infty$. That is $L_1(ta)/L_1(t) \to 1$ as $t \to \infty$ for any $a > 0$. This property implies that the correlations are not summable, and the spectral density has a pole at zero. Under suitable conditions on $L_1(\cdot)$, the spectral density

$$f(x) = \frac{1}{2\pi} \sum_{k=-\infty}^{\infty} r(k) e^{ikx} L_2(x) |x|^{1/2H}$$

is $|x| \to 0$ for some $L_2(\cdot)$ slowly varying the origin. The explanation for this phenomenon is not elusive since these economic data always persist for a rather long period and dependence between time thus has a lasting influence. Yet traditional wisdoms usually appear stuck in analysis of long memory, and thus need for structural break detection requires econometricians to develop singular new estimations for use. Among these manufactory efforts attempt to deal with long dependence either invariant to time or matching the time future. Hereafter more methods will be discussed.

31.2.4.1 Stochastic Volatility Model with Long-Memory Property

Macroeconomic and financial data often display long-run dependence between different economic individuals. Researchers and practitioners prefer the synthesized form to describing these issues, and thus data aggregation becomes another focus in econometrics literature. Interestingly such time series often turn out distinct from the originals, for example, the singular stock price return is normally stationary I(0), while the index often displays long memory. The Standard Poor's 500 index would be a supporting evidence for it. Furthermore Hsiao and Robinson (1978) and Granger (1980) have pointed that it makes sense for contemporaneous aggregation with stationary heterogenous autoregressive-moving average processes. Portfolio arrangement in this sense also regards to connect its volatility with the long-memory feature. As discussed in Zaffaroni (2000) and Kazakevicius et al. (2004), converse results also exist in the generalized autoregressive conditional heteroskedasticity (GARCH). Thus whether a long memory can be obtained by aggregation in volatility models has attracted a lot of attentions.

Zaffaroni (2006) investigates a large class of volatility models for the memory implications. With this aim, he selects the class of square root stochastic autoregressive volatility (SR-SARV), which nests both GARCH and stochastic volatility (SV) models. So the exponential SV model of Taylor (1986) and the nonlinear moving average model (nonlinear MA) of Robinson and Zaffaroni (1998) would be the referents. The result shows long memory is ruled out for the former but is permitted for the latter. To put it more evidently, set a stationary square integrable process x_t with increasing filtration J_t if x_t is J_t-adapted, $E(x_t|J_{t-1}) = 0$, and $\mathrm{var}(x_t|J_{t-1}) =: f_{t-1}$. And the $SR - SARV(1)$ process then satisfies

$$f_t = \omega + \gamma_{t-1} + v_t, \tag{31.16}$$

where the sequence v_t is assumed as $E(v_t | J_{t-1}) = 0$. ω and γ are constant nonnegative coefficients with $\gamma < 1$. The implication means $E((x)_t^2) < \infty$. It is also assumed that ω_i, γ_i are independent identically distributed (i.i.d.), randomly drawn from a joint distribution such that $\gamma_i < 1$. At each point n heterogeneous units $x_{i,t}(1)$ are observed as SR-SARV(1).

31.2.4.2 Testing for a Change in the Long-Memory Parameters

The best known models representing long dependence are fractional Gaussian noise (Mandelbrot van and Ness 1968) and fractional ARIMA (Granger and Joyeux 1980). These models are stationary with a constant long-memory parameter H. But this assumption could not hold for some time series since the long dependence structure might change over time. It makes sense in macroeconomic data for some economic events persist over long period or has a lasting influence. Note that changes of H are relevant particularly for that the rate of convergence of confidence intervals of constants and for parameter estimates in regression with certain classes of design matrices, e.g., polynomial regression would change if H changes. More details are referred to Yajima (1988) and Beran (1991). There are huge sums of literature focused on this issue, among which Beran and Terrin (1996) offer a simple estimation by exact or approximate maximum likelihood. The method testing such structural break consists of the alternative that H is not constant. They make use of functional central limit theorem to compute the quadratic forms and offer more test statistics.

Define the spectral density by a finite dimensional parameter vector $\theta = (\tau, \eta) = (\tau, H, \eta_2, \ldots, \eta_m)$ so that

$$f(x;\theta) = \tau f\{x;\ (1,\eta)\}, \int_{-\pi}^{\pi} \log\ f\{x;\ (1,\eta)\} d_x = 0.$$

By definition, τ is the expected mean squared error of the best linear prediction of X_t given X_s, $s \le t - 1$. The long-memory behavior is characterized by H, the additional parameters $\eta_2, \ldots, \eta_m$ allow for flexible modelling of short-term features. Define

$$\alpha_k(\eta) = \int_{-\pi}^{\pi} e^{ikx} f^{-1}\{x;\ (1,\eta)\} dx.$$

Given $X_1, X_2, \ldots, X_N$, let $\hat{\eta}$ be the value of η that minimizes

$$Q(\eta) = \sum_{i,j=1}^{N} \alpha_{i-j}(\eta) \left(X_i - \overline{X}\right) \left(X_j - \overline{X}\right).$$

For Gaussian processes, the asymptotic covariance matrix is the same as for the exact maximum likelihood estimator. As proved by Beran and Terrin (1996), for $0 < t < 1$,

$$Q_1(t;\eta) = \sum_{i,j=1}^{\lfloor Nt \rfloor} \alpha_{i-j}(\eta)\left(X_i - \overline{X}\right)\left(X_j - \overline{X}\right),$$
$$Q_2(t;\eta) = \sum_{\lfloor Nt \rfloor+1}^{N} \alpha_{i-j}(\eta)\left(X_i - \overline{X}\right)\left(X_j - \overline{X}\right).$$

Let $\hat{\eta}^{(1)}(t)$ and $\hat{\eta}^{(2)}(t)\,(j=1,2)$ be defined by

$$\hat{\eta}^{(j)}(t) = \text{argmin}\ Q_j\left(t; \hat{\eta}^{(j)}\right)(j=1,2),$$

and denote by $\hat{H}^j(t) = \hat{\eta}_1^j(t)$ the corresponding estimates of H. Moreover, define $\kappa^2 = 2D_{11}^{-1}(\eta)$, where D is the $k \times k$ matrix with elements

$$D_{ij} = (2\pi)^{-1}\int_{-\pi}^{\pi} \frac{\partial}{\partial \eta_i} logf\{x;\ \ (1,\eta)\}dx.$$

Then the process

$$\widetilde{Z}_N(t) := N^{\frac{1}{2}}\kappa^{-1}\{t(1-t)\}^{\frac{1}{2}}\left\{\hat{H}^{(1)}(t) - \hat{H}^{(2)}(t)\right\}$$

converges in the Skorokhod topology on $D[0,\ 1]$ to the Gaussian process $Z(t)$ defined by

$$Z(t) = \{t(1-t)\}^{\frac{1}{2}}\left\{\frac{1}{t}B_1(t) - \frac{1}{1-t}B_2(1-t)\right\} \tag{31.17}$$

where B_1 and B_2 are two independent standard Brownian motions. Based on this distribution, the test statistics is suggested as

$$T_N = \sup_{\delta<t<1-\delta}\left|\widetilde{Z}_N(t)\right| \tag{31.18}$$

for some $0 < \delta < 1$. Also it is implied that T_N in distribution, where

$$Y = \sup_{\delta<t<1-\delta}|Z(t)| \tag{31.19}$$

and $Z(t)$ is defined by Eq. 31.17. Note that, due to the standardization by $\{t(1-t)\}^{\frac{1}{2}}$, $Z(t)$ is a standard normal random variable for each fixed t.

Comparatively, tests for a change in parameter values at a given time point are proposed in linear regression models with long-memory errors. Hidalgo and Robinson (1996) design a structural change test specially for I(d) data. They derive the new test in terms of stochastic and non-stochastic regressors. Generally, their model of form is $y_t = \beta\left(\frac{t}{n}\right)' x_t + u_t$, where x_t is a $K \times 1$ vector of observable regressors, the prime indicates transposition, and $\beta(s)$ is a $K \times 1$ vector such that

$$\beta(s) = \beta_A, 0 \le s \le \tau$$
$$= \beta_B, \tau < s < 1,$$

for a known number τ between zero and one, where the unobservable error u_t has auto-covariances featuring long memory. Thus the null hypothesis is $H_0 : \beta_A = \beta_B$ against the alternative $H_1 : \beta_A \neq \beta_B$.

For the case of non-stochastic regressors, a type of Wald testing procedure is applied. For a given τ, define $h = [\tau n]$, where $[\cdot]$ indicated integer part. Correspondingly, let $X_1 = (x_1, \ldots, x_h)'$, $X_2 = (x_{h+1}, \ldots, x_n)'$, $Y_1 = (y_1, \ldots, y_h)'$, $Y_2 = (y_{h+1}, \ldots, y_n)'$, and then estimate β_A and B by

$$\hat{\beta}_A = \left(X_1'X_1\right)^{-1}X_1'Y_1,$$
$$\hat{\beta}_B = \left(X_2'X2\right)^{-1}X_2'Y_2$$

Put

$$W = \left(\left(X_1'X_1\right)^{-1}X_1', -\left(X_2'X_2\right)^{-1}X_2'\right)',$$
$$u = (u_1, \ldots, u_n)',$$

Set w_t the tth column of W'. Then H_0 is equivalent to

$$\hat{\beta}_A - \hat{\beta}_B = W'u.$$

Assume u_t is Gaussian with zero mean, then

$$\hat{\beta}_A - \hat{\beta}_B \sim N\left(0, W'\right) \tag{31.20}$$

where Γ is the $n \times n$ Toeplitz matrix ($\gamma(s - t)$).

When x_t is stochastic but independent of u_t, however, Eq. 31.20 can be translated as

$$\left(W'\Gamma W\right)^{-1/2}\left(\hat{\beta}_A - \hat{\beta}_B\right) \sim N(0, \ I_k).$$

Hidalgo and Robinson (1996) also allowed for a degree of generality in the specification of the error structure, and the test is mainly designed for liner regression model by applying a semiparametric but a parametric model. Besides, the change point is assumed to be already known. Yet Kuan and Hsu (1998) point that such a test would incorporate size distortion that could lead to a change point while there is none. This is not peculiar since according to Kuan and Hsu (1998) when the memory parameter is in (0, 0.5), many well-known structural changes are supposed to suggest a nonexistent change. They also indicate a spurious change which might arise from stationary data with long memory while not responsible to nonstationary. They thus infer distinguishing between long-memory series with a change would not be reliable sometimes. More specifically, more and more financial literature illustrate that the realized volatility, which plays a major role in forming the portfolios, displays long-memory properties; thus, in order to avoid the loss caused by the market fluctuations, we could use those structural break tests for long-memory processes to detect the turning points of the realized volatility and further adjust the portfolio.

31.3 Persistent Change in Time Series

Permanent property shifts of the time series are characters of many key macroeconomic and financial variables in developed economies. Hence, correctly characterizing the time series whether stationary or nonstationary would be very helpful to build the accurate models. This issue has been investigated in terms of the data property change from I(0) to I(1) or the vice versa. The change might be responsible for portfolio adjustment especially for financial crises. Furthermore, recent studies indicate that other than those classical structural break types, there could also exist persistent changes in processes characterized by long memory, for example, the long-memory parameter changes from d_1 to d_2 or the vice versa, where d is the fractional differencing parameter.

31.3.1 Tests for a Change in Persistence for I(0) or I(1) Process

A normal question in testing for the persistent change is about what the process used to be and what direction it changes to. To solve that systematically, we could set the data-generating process (DGP) as following:

$$y_t = d_t + v_t, \tag{31.21}$$

$$v_t = \rho_t v_{t-1} + \varepsilon_t, t = 1, \ldots T. \tag{31.22}$$

Eq. 31.21 consists of the deterministic kernel d_t which is a constant either plus linear time trend or nothing. The error term ε_t is stationary.

Within the model (31.21), there exist four possibilities. The first is that y_t is all the time I(0) as $\rho_t = 1$ for all t, which can be defined as $H_{1,}$ and the second is denoted as $H_0 1$, reflecting Y_t changing from I(0) to I(1) at time $\lfloor \tau^* \rfloor$, where $\lfloor \cdot \rfloor$ represents the integer part. In this situation, $\rho_t = \rho$, $|\rho| < 1$ $t \leq \lfloor \tau^* \rfloor$ and $\rho_t = 1$ for $t > \lfloor \tau^* \rfloor$ in the context of (i). Noted that the change-point proportion τ^* is assumed to be unknown while set in $\Lambda = [\Lambda_L, \Lambda_U]$, which is an interval symmetric around 0.5 and between zero and one. The third denoted as $H_1 0$ is that y_t is form I(1) to I(0) at time τ^*.

Extant tests to detect the above structural changes are CUSUM of squares-based tests and regression-based tests. While the former has been improved by Leybourne et al. (2006) as displaying no tendency to spurious over-rejection, modified tests from regression angle by LKSN still lack the power for special structural change, taking large sample, for instance, where tests reject H_1 with probability one even though there is no change. Owing to these reasons, we here provide the CUSUM test according to Leybourne et al. (2006) as the ideal way in detecting this kind of persistent change.

In order to test the null hypothesis of constant I(0) behavior H_1 against a change in persistence from I(0) to I(1), H_{01}, a standardized CUSUM of squared subsample OLS residuals is needed:

$$K^f(\tau) = \frac{[\tau T]^{-2}\sum_{t=1}^{[\tau T]} \hat{v}_{t,\tau}^2}{\hat{w}_f^2(\tau)} \tag{31.23}$$

where

$$\hat{v}_{t,\tau} = y_t - \overline{y}(\tau) \quad \text{with} \quad \overline{y}(\tau) = [\tau T]^{-1}\sum_{t=1}^{[\tau T]} y_t.$$

And similarly $H(10)$ equivalent to change from $I(0)$ to $I(1)$ in the reversed series, let $x_t \equiv y(T - t + 1)$, occurring at time $(T - [\tau^* T])$ and the reversed series can be obtained as

$$K^r(\tau) = \frac{(T - [\tau T])^{-2}\sum_{t=1}^{T-[\tau T]} \widetilde{v}_{t,\tau}^2}{\hat{\omega}_r^2(\tau)} \tag{31.24}$$

where

$$\widetilde{v}_{t,\tau} = x_t - \overline{x}(1-\tau) \quad \text{with} \quad \overline{x}(1-\tau) = (T - [\tau T])^{-1}\sum_{t=1}^{T-\lfloor \tau T \rfloor} x_t.$$

Notice for both forward and reversed series, the long-run variance ω is replaced by the estimator

$$\hat{\omega}_f^2(\tau) = \hat{\gamma}_0 + 2\sum_{s=1}^{m} w_{s,m}\hat{\gamma}_s, \tag{31.25}$$

$$\hat{\gamma}_s = \lfloor \tau T \rfloor^{-1}\sum_{t=1}^{\lfloor \tau T \rfloor} \Delta\hat{v}_{t,\tau}\Delta\hat{v}_{t-s,\tau}, \tag{31.26}$$

Therefore, a test against persistent change could be based on the ratio R of the forward and reverse CUSUMs of squared statistics, i.e.,

$$R(\tau^*) = \frac{inf_{\tau \in \Lambda} K^f(\tau^*)}{inf_{\tau \in \Lambda} K^r(\tau^*)} =: \frac{N}{D}$$

where Λ is as described before and τ is assumed to be unknown. Furthermore, denote the demeaned, written as $z_t = 1$, and de-trend, $z_t = (1,t)'$, respectively. As $T \to \infty$, $l^{2/T} \to 0$,

$$R \Rightarrow \frac{inf_{\tau \in \Lambda} L_\xi^f(\tau)}{inf_{\tau \in \Lambda} L_\xi^r(\tau)}. \quad (31.27)$$

As shown in Leybourne et al. (2006), this limiting distribution of R is not dependable on the long-run variance, thus R could still hold if it is formed by the unstandardized ratio:

$$R = \frac{inf_{\tau \in \Lambda} \overline{K}^+}{inf_{\tau \in \Lambda} \overline{K}^r(\tau)} \quad (31.28)$$

where

$$\overline{K}^+(\tau) = [\tau T]^{-2} \sum_{t=1}^{[\tau T]} \hat{v}_{t,\tau}^2 \quad \text{and} \quad \overline{K}^r(\tau) = (T - [\tau T])^{-2} \sum_{t=1}^{T-[\tau T]} \widetilde{v}_{t,}^2.$$

And the large sample behavior of R, N, and D under both H_{01} and H_{10} will be provided in their Theorem 2 along with consistent estimators of τ^*. Additionally, the results of R in their Theorem 2 imply a consistent test of H_1 against H_{01} or H_{10}, respectively, by the left-tail and the right-tail distribution of R, while the unstandardized form yields a consistent estimate of break date τ^*. What's more, even the direction of the change is unknown, such a test with two tails can be appropriate against either H_{01} and H_{10}. As for the limiting distribution of R under H_0, it remains available for both the demeaned and de-trend cases. Thus, we could further know that under H_0, $R \to^p 1$.

Although Leybourne et al. (2006) rule out the size distortion when applied to process displaying constant persistence, they neglect such a test for long-memory case could cause over-rejection, which is often very serious. Sibbertsen and Kruse (2009) provide a solution based on simulation, by which they find despite that critical values may need modified, estimators for break dates still hold consistently. In this sense, the y_t generated from DGP is assumed to follow an ARFIMA(p,d,q) process, which means only the memory parameter d determine the degree of integration of y_t. Then the hypothesis should be written as

$$\begin{aligned} H_0 &: d = d_0 \quad \text{for all} \quad t, \\ H_1 &: d = d_1 \quad \text{for} \quad t = 1, \ldots, [\tau T] \\ & d = d_2 \quad \text{for} \quad t = [\tau T] + 1, \ldots, T \end{aligned}$$

Typically under H_0 the memory parameter d_0 is restricted to [0, 3/2), while $d_1 \in$ [0, 1/2) and $d_2 \in$ (1/2, 3/2]. It is equivalent to changes from stationary to nonstationary long memory at some unknown breakpoint with the vice versa also valid. Similarly, parts constituting R statistics are thus defined by

$$K^f(\tau) = [\tau T]^{-2d_0} \sum_{t=1}^{[\tau T]} \hat{v}_{t,\tau}, {}^r(\tau) = (T - [\tau T])^{-2d_0} \sum_{t=1}^{T-[\tau T]} \hat{v}_{t,\tau}.$$

However, Sibbertsen and Kruse (2009) find that asymptotic property has changed since the limiting distribution depends strongly on the memory parameter, which still leads to heavy size distortion. To deal with this problem, they simulate the adjusted critical values which are dependent on d. Although in this way practitioners benefit as such sample size will not influence critical values as significantly much as memory parameter, they have to decide the exact d for this simulation.

31.3.2 Unit Root with Persistent Change

Conventional unit root tests usually seem problematic when they are applied to nonstationary volatility, although they are proved to perform well in other situations. Breaks following a stationary Markov switching process and time-varying conditional variances are proved not to cause significant size distortion and impact on unit root and cointegration tests. Similarly, stationary time-varying conditional variances including (ARCH) are also known to have no impact on unit root and cointegration tests. Conversely, it can be different for permanent changes in volatility (so that volatility is nonstationary), since it is easy to greatly affect unit root inference. Hamori and Tokihisa (1997) show that a single abrupt increase in the innovation variance increases the size of augmented Dickey and Fuller (1979, 1981) (ADF) tests when no deterministic components are present. The similar situation happens to Kim et al. (2002) who report severe oversizing in the constant-corrected ADF tests in the presence of an early single abrupt decrease in the innovation variance. Cavaliere (2004) proposes a more general framework for investigating the effects of permanent changes in volatility by showing that the limiting null distributions of the Phillips and Perron (1988) (PP) test statistics depend on a particular function of the underlying volatility process which can lead to either oversizing or undersizing in the tests.

Cavaliere and Taylor (2006) consider both smooth volatility changes and multiple volatility shifts and construct a numerical solution to obtain approximate quantiles from the asymptotic null distributions of the standard unit root statistics. They focus on M unit root tests by Perron and Ng (1996) which is still robust when applied to autocorrelation. Set the following model to generate the time series process

$$\left(X_t - \gamma_t^{'}\right) = \alpha\left(X_{t-1} - \gamma^{'} Z_{t-1}\right) + u_t \quad t = 1, 2, \ldots, T \tag{31.29}$$

where z_t is a vector of deterministic components, thus set as $z_t = (1, t, \ldots, t^p)'$, while $\varepsilon_t = \sigma_t e_t$, $e_t \sim iid\,(0, 1)$ which is responsible for the heterogeneity. And besides X_0 is of $O_p(1)$. Furthermore, the M statistics for a given sample $\{X_t\}_0^T$ is defined as

$$MZ_\alpha := \frac{T^{-1}\hat{X}_T^2 - T^{-1}\hat{X}_0^2 - s_{AR}^2(k)}{2T^{-2}\sum_{t=1}^{T}\hat{X}_{t-1}^2}.$$

where $\hat{X}^t$ are the OLS residuals from the regression of X_t on z_t, $t = 0, \ldots, T$ and $s_{AR}^2(k)$ is an autoregressive estimator of the (non-normalized) spectral density at frequency zero of $\{u_t\}$. Specifically,

$$s_{AR}^2(k) := \hat{\sigma}^2 / \left(1 - \hat{\beta}(1)\right)^2, \quad \hat{\beta}(1) := \sum_{i=1}^{k} \hat{b}_i,$$

where $\hat{b}_i$, $i = 1, \ldots, k$, and $\hat{\sigma}^2$ are, respectively, the OLS slope and variance estimators from the regression equation $\Delta\hat{X}_t = \hat{X}_{t-1} + \sum_{i=1}^{k} b_i \Delta\hat{X}_{t-i} + \varepsilon_{t.k}$, where the lag truncation parameter satisfies the following assumption. Details about the asymptotic distribution are referred to Cavaliere and Taylor (2006). Moreover, the Cavaliere and Taylor method performs as a general way to deal with the persistent change in volatility. However, in real world it usually cannot be identified easily. In other words, it smooths off structural change and hold for any process, providing another conveniently operative way for this issue.

31.4 Tests for Special Structural Breaks

There also exist several types of structural breaks that have exclusive features distinctly from common economic phenomena. For example, a financial bubble being originated or bursting indicates a significant and catastrophical structural break, which is capable of wiping out the great wealth of people as much as possible. Additionally, the cointegration, recognized as a common relationship between economic individuals especially in international finance and macroeconomics, could be with breaks. Thus, testing a break in cointegrating relationship correctly could be a solution of financial crises.

31.4.1 Bubble Tests

31.4.1.1 Definition of the Bubble

We first present the definition of the bubble. A bubble is always characterized as an explosive asset price deviation from its fundamentals regardless of whether such a deviation is positive or negative. However, investors and regulators often ignore bubbles rather easily, because there exists no efficient methodology to monitor and detect those. In general, a bubble could be triggered by the irrational behaviors. Shiller (2000) conducts this conclusion by observing investors who switch their attentions from the specific financial asset to another and points out these behaviors take no consideration of deviations in fundamentals. More importantly, most historical events coincided with Shiller's findings; thus those opinions become more convincing. A famous example is the remark given by Greenspan, the ex-chairman of the Federal Reserve Board, which illustrates that it would be fairly difficult to know when irrational exuberance has unduly escalated asset values and thus the USA experienced the bubble collapse.

Furthermore, those claims have already been proved. Kirman and Teyssiere (2005) rebuked that the deviations from fundamental assets are not the outcome of irrational herding but the significant shifting compositions of expectations, which result from opinion diffusion processes. This view is backed up by an analysis of two types of agents in financial market, which are fundamentalists and chartists. The former denotes agents who believe asset price P_t is related to underlying fundamental $\overline{P}_t$, i.e., a constant $\overline{P}$. The latter depends on the P_t which is cumulated by historical prices. Moreover, there also exists heterogeneity between these agents. Thus for the former,

$$E^f(P_{t+1}|I_t) = \overline{P}_t + \sum_{j=1}^{M^f} v_j\left(P_{t-j+1} - \overline{P}_{t-j}\right) \tag{31.30}$$

where $v_j, j = 1, \ldots, M^f$ are positive constants and M^f is the memory of the fundamentalists. And in return for the latter

$$E^c(P_{t+1}|I_t) = \sum_{j=0}^{M^c} h_j P_{t-j} \tag{31.31}$$

where $h_j, j = 1, \ldots, M^c$ are constants, M^c is the memory of the chartists. And based on the two opinions, the market view of prices are synthetic by the weighted average of these forecast

$$E^m\left(P_{t+1}|I_t\right) = w_t E^{\mathrm{f}}(P_{t+1}|I_t) + (1 - w_t)E^c(P_{t+1}|I_t) \tag{31.32}$$

where w_t is the proportion of fundamentalists. Such a market view is dynamic since contacts or meetings happen randomly between agents and often help build their next investment plans. In this sense, consider k_t as the number of fundamentalists at time t and the remaining $N - k_t$ as chartists. Allow some fixed number M of meetings to take place at each time. Set $q_t = k_t/N$ for each agents as the observation to decide the opinion chosen by majority. And thus

$$q_{i,t} = q_t + \varepsilon_{i,t}, \varepsilon \sim N\left(0,\ \sigma_q^2\right),\ \ q_{i,t} \in [0, 1].$$

And it can be easily understood that w_t functioned in Eq. 31.32 originates from the following mechanism,

$$w_t = N^{-1} \sum_{i=1}^{N} \sharp\left\{i : q_{i,t} \geq \frac{1}{2}\right\}.$$

Apparently w_t varies with time as the q_t process, and it is believed to be dependable on k_t reciprocally. Given the two types of agents, the market price can be written as

$$P_t = cE^m(P_{t+1}|I_t) + Z_t \tag{31.33}$$

where c is a constant and Z_t is an index of a vector of fundamental variables. Obviously, the time-varying W_t decides the market view of the prices in the future and thus is determinant in the real price. As theoretically proved by Kirman and Teyssiere (2005), W_t switching from zero to one and vice versa could lead to varying prices around the fundamental level where a change-point process in the conditional mean definitely exist. It also helps explain that such process never herds on the extreme and sometimes moves gradually.

Therefore the above two given opinions are agreeable in the aspect that when prices explode there may be a change point. Accordingly, the implicit change point is another version of structural break; thus testing for bubbles efficiently could be viewed as a way to detect the explosive behavior in pricing.

31.4.1.2 Tests for Bubbles

According to the above explanation, many literatures propose suitable testing mechanism for rational bubbles. As among these tests, a basic solution is based on a unit root test. Since rational bubbles always manifest explosive characteristics in prices, these tests such as augmented Dickey-Fuller (ADF) could find this distinction with stationary process and unit root. Commonly regarded as expectations of prices and dividends in the future, the current price is determined sly by the fundamental which refers to dividends if there is no bubble. This means pricing follow the variation of dividends, and often there is cointegration relationship between them. Conversely, if a bubble occurs in the prices, explosive behaviors are expected regardless of what character dividends own. This implication motivated Diba and Grossman (1988) to firstly look for the presence of bubble behavior by applying unit root tests to $_t$. However, these tests are criticized by Evans (1991) who questions the validity of the empirical tests since none of them have much power to detect periodically collapsing bubbles. He supports this criticism by explaining that a periodically collapsing bubble process can behave much like an I(1) process or even like a stationary linear autoregressive process, as a result of which the standard unit root and cointegration tests in this context are not reliable.

Phillips et al. (2009) design a new synthetical unit root test and avoid such problem. Such a refinery ADF test against the alternative of an explosive root(the right-tailed) is conducted in an autoregressive specification as

$$x_t = \mu_x + {}_{t-1} + \sum_{j=1}^{J} \phi_{jt-j} + \varepsilon_{x,t}, \varepsilon_{x,t} \sim NID\left(0, \sigma_x^2\right) \tag{31.34}$$

where x_t stands for log stock or log dividend and the certain order of J is suggested by Campbell and Perron (1991), while NID denotes independent and normal distribution. Thus the unit root null hypothesis is $H_0 : \delta = 1$, and the right-tailed alternative is $H_1 : \delta > 1$.

After this, a series of forward recursive regressions are applied to Eq. 31.34 repeatedly with sequential subsets of the sample data incremented by one observation at each pass. Set number of observations in the first regression as $\tau_0 = [nr_0]$ and subsequent regressions employ the originating data set supplied by successive observations giving a sample of size $\tau = [nr]$ for $r_0 \le r \le 1$. And the corresponding t-statistic by ADF_r and ADF_1 corresponds to the full sample. Under the null,

$$ADF_r \Rightarrow \frac{\int_0^r WdW}{\left(\int_0^r W^2\right)^{1/2}},$$

and

$$\sup_{r\in[r_0,1]} ADF \Rightarrow \sup_{r\in[r_0,1]} \frac{\int_0^r WdW}{\left(\int_0^r W^2\right)^{1/2}}$$

where W is the standard Brownian motion. Comparing $\sup_{r\in[r_0,1]}$ ADF with the right-tailed critical values with $\sup_{r\in[r_0,1]}\left(\int_0^r WdW\right)\Big/\left(\left(\int_0^r W^2\right)^{1/2}\right)$ will lead to a unit root test for against explosiveness. And to pinpoint the structural break, the recursive test statistics ADF_r are needed against the right-tailed critical values of the asymptotic of the standard Dickey-Fuller t-statistic. In particular, if r_e denotes the origination date and r_f is the collapse date of explosive behavior in the data, estimates of these dates are as follows:

$$\hat{r}_e = \underset{s\ge r_0}{\text{INF}}\, s : ADF_s > cv_{\beta_n}^{adf}, \quad \hat{r}_f = \inf_{s\ge \hat{r}_e} \inf\left\{s : ADF_s < cv_{\beta_n}^{adf}\right\} \tag{31.35}$$

where $cv_{\beta_n}^{adf}(s)$ denotes the right-sided critical value of ADF_s corresponding to a significant level of β_n. Notice the consistent estimation of the bubble period $\hat{r}_e, \hat{r}_f$ requires the significance level β_n approaching zero asymptotically and correspondingly $cv_{\beta_n}^{adf}(s)$ diverging to infinity in order to eliminate type I error as $n \to \infty$. This can be implemented in a convenient way employing $cv_{\beta_n}^{adf}(s) = \log(\log(ns))/100$. Particularly, it replaces the explicit setting for β_n which is more complicated. This test referred by Phillips et al. (2009) is advantageous for the reason that it doesn't allow for the possibility of periodically collapsing bubbles, which are often observed in practical economic and financial applications.

What's more, there are more other attempts that could avoid potential setbacks in performance of testing for bubbles. Among them, large sums of progressing improvement are from unit root tests for explosive characters. For example, bootstrapping is applied to improve the power of tests, of which Paparoditis and Politis (2003) make use for the heavy-tail problem. For DGP in Eq. 31.34, it is assumed that

$$\hat{\varepsilon}_t = \theta \hat{\varepsilon}_{t-1} + e_t \tag{31.36}$$

For a test against $|\theta| < 1$, the residual block bootstrap method is carried as follows: for each block bootstrap series satisfying H_0, set the sample k, $k = 1, \ldots, B$, and the integer $b = b(T) < T$ is employed such that $1/b + b/\sqrt{T} \to 0$ as $T \to \infty$. Define $\kappa = [T(-1)/b]$ and draw with replacement κ integers $i_0, \ldots, i_{\kappa-1}$ from the set $1, \ldots, T - b$, and build the bootstrap samples $\varepsilon j(k)$ as follows:

$$\varepsilon_1(k) = \hat{\varepsilon}_1, \varepsilon_j(k) = \varepsilon_{j-1} + \overline{\hat{e}}_{i_m+s}, j = 2, \ldots, l, l = \kappa b + 1,$$
$$m = [(j-2)/b], s = j - mb - 1, \quad k = 1, \ldots, B.$$

Let $\hat{\theta}_T$ be the least-squares (LS) estimator of θ in Eq. 31.36 and $\widetilde{\theta}_l(k)$ be the LS estimator of the regression of $\varepsilon_j(k)$ on $\varepsilon_{j-1}(k)$. Thus H_0 would be rejected if $T\left(\hat{\theta}_T - 1\right) < q_{l,B}(\gamma)$, where $q_{l,B}(\gamma)$ is the γth quantile of the distribution of $l\left(\widetilde{\theta}\right)_l(\cdot) - 1\big)$ when the size for testing is γ. Since some financial time series might have heavy tails, the tail index α would make $Pr(|e_t| > x) \sim x^{-\alpha}$ as $x \to \infty$ for some $\alpha > 0$. Yet this modification of standard asymptotic theory for unit root tests is still valid.

Horvath and Kokoszka (2003) and Jach and Kokoszka (2004) make the mild hypothesis that the $E(e_t) = 0$ and that the e_t are in the domain of attraction of an α-stable law with $\alpha \in (1, 2)$, while Kokoszka and Parfionovas (2004) consider the more general case $\alpha \in (1, 2]$ which then includes the Gaussian case $\alpha = 2$. Chan and Tran (1989) have shown that under the null hypothesis H_0,

$$T\left(\hat{\theta}_T - 1\right) \xrightarrow{d} \xi := \frac{\int_0^1 L_\alpha(\tau-) dl_\alpha(\tau)}{\int_0^1 L_\alpha^2(\tau) d\tau}.$$

The purpose of the unit root and subsampling tests here is to approximate the distribution of the unit root statistics ξ without knowledge of the tail index *alpha* which is difficult to estimate.

Another method employing subsampling is advocated by Jach and Kokoszka (2004) consists in constructing $T - b$ processes which satisfies H_0,

$$\varepsilon_1(k) = \overline{\hat{e}}_k, \ldots, \varepsilon_b(k) = \overline{\hat{e}}_k + \cdots + \overline{\hat{e}}_{k+b-1}, k = 2, \ldots, T - b + 1,$$

where b is the size of the subsampling blocks. Let $\widetilde{\theta}_b(k)$ be the LS estimator of the regression of $\varepsilon_j(k)$ on $\varepsilon_{j-1}(k)$. Then the distribution of ξ would be estimated by $b\left(\widetilde{\theta}_b(\cdot) - 1\right)$.

Moreover Wu and Xiao (2008) have proposed a procedure similar to cointegration tests by Xiao and Phillips (2002) to help detect collapsible bubbles against which Evans (1991) pointed out that unit root tests would lose the power.

Their procedure is based on the magnitude of variation of the partial sum processes $S_k = \sum_{t=1}^{k} \hat{\varepsilon}$ of the residuals of regression (31.36). If there is no bubble, the magnitude of fluctuation of the process S_k is proportional to $k^{1/2}$, while the presence of a bubble makes the process S_k diverging to ∞. Their statistic denoted by $\underline{R}$ enjoys the advantage of no influence from serial correlation and the correlation between the residuals and the fundamentals under the null hypothesis. As showed below, $\underline{R}$ converges to the supremum of a functional of Brownian motions, i.e.,

$$\underline{R} := \max_{1\leq k\leq T} \frac{k}{\hat{\omega}_{\varepsilon,d}\sqrt{T}} \left| k^{-1} S_k^{+} - T^{-1} S_T^{+} \right| \xrightarrow{d} d\sup_{0\leq\tau\leq 1} \left| \left(\widetilde{V}\right)(\tau) \right|$$

where $\widetilde{V}(\tau) = \underline{W}_d(\tau) - \tau \underline{W}_d(1)$, $\underline{W}_d = W_1(\tau) - \left[\int_0^1 dW_1 S'\right] \left[\int_0^1 SS'\right]^{-1} \int_0$, $S(\tau)' = (1, W_2(\tau))$, $W_1(\tau)$ and $W_2(\tau)$ are Brownian motions that are independent of each other; $\hat{\omega}_{t,d}$ is a nonparametric long-run variance estimator.

31.4.2 Cointegration Breakdown Tests

In addition to detecting the structural breaks in a time series, tests for a breakdown in the cointegration relationship between two nonstationary I(1) process are also of interest. For example, correlation of global financial market becomes more obvious as a result of rapid international capital flows, which feature I(1) processes and are generally the linear combinations of nonstationary time series. As shown by most empirical evidences, the existing cointegration relationship often comes to the end prior to the finite crisis. Such a structural break is vital since it is an indicator for finance crash. Traditional wisdom formulates the cointegration breakdown tests always from the assumption that the post-breakdown period is relatively long, and this is often strong and unrealistic especially for practitioners, see Hansen (1992) and Quintos and Phillips (1993). Andrews and Kim (2003) on the contrary abandon this assumption and introduce tests for cointegration breakdown with fixed post-breakdown time length m, under the condition that the sample $T + m$ goes to infinity. Clearly, their implementation concentrating on end-of-sample could be conveniently extended to breakdown tests occurring at the beginning or in the middle of the sample.

The data-generating process is as follows:

$$y_t = \begin{cases} x'\beta_0 + u_t & \text{for} \quad t = 1, \ldots, T \\ x'\beta_t + u_t & \text{for} \quad t = T+1, \ldots, T+m \end{cases}$$

where y_t, $x_t \in R$ and x_t, β_0, $\beta_t \in R^k$ and the regressors for all time periods are I(1) processes with potential deterministic and stochastic trend or other stationary random variables. Thus the null and the alternative hypotheses are

$$H_0 : \begin{cases} \beta_t = \beta_0 \quad \text{for} \quad \text{all} \quad t = T+1, \ldots, T+m \quad \text{and} \\ u_t : t = 1, \ldots, T+m \quad \text{are stationary and ergodic} \end{cases}$$
$$H_1 : \begin{cases} \beta_t \neq \beta_0 \ \text{for some} \quad t = T+1, \ldots, T+m \quad \text{and/or} \\ \text{the distribution of} \quad u_{T+1}, \ldots, u_{T+m} \quad \text{differs from} \\ \text{the distribution of} \quad u_1, \ldots, u_m \end{cases} .$$

The alternative hypothesis H_1 represents a break in cointegrating relationship in systematical perspectives including (i) a shift in the vector β_0 to β_t and (ii) a shift in the distribution of u_t from being stationary to being a unit root random variable. They consider a test statistics in the quadratic form of the "post-breakdown" residuals $\hat{u}_t : t = T+1, \ldots, T+m$. The critical value of the test statistics is determined via a parametric subsampling method, which if the test statistics exceeds then the test rejects the null hypothesis.

For any $1 \le r \le s \le T+m$, let $Y_{r-s}\overline{(}y_r, \ldots, y_s)', X_{r-s}\overline{(}x_r, \ldots, x_s)', U_{r-s}\overline{(}u_r, \ldots, u_s)'$. And the quadratic form will be

$$P_j(\beta, \omega) = \left((Y_{j-(j+m-1)} - X_{j-(j+m-1)}\beta\right)' \omega\left((Y_{j-(j+m-1)} - X_{j-(j+m-1)}\beta\right)$$

and for $j = 1, \ldots, T + 1$. For the P tests in Andrews and Kim (2003), ω is some nonsingular m matrix and I_m denotes the m dimensional identity matrix. Naturally, an estimator of β_0 denoting as $\hat{\beta}$ is based on the least squares with observations $t = r, \ldots, s$ for $1 + m$ as following:

$$\hat{\beta}_{r-s} = \left(X'_{r-s}X_{r-s}\right)^{-1} X'_{r-s}Y_{r-s}$$

Although other estimators like the fully modified estimator of Phillips and Hansen (1990) and the ML estimator of Johansen can also be applied, the priority for explanation is given to LS estimator.

Then the first test statistics, P_a is defined as

$$P_a = P_{T+1}\left(\hat{\beta}_{1-T}\right) = \sum_{t=T+1}^{T+m} \left(y_t - x_t'\hat{\beta}_{1-T}\right)^2.$$

Referred to as a predictive statistic, P_a is the post-breakdown sum of squared residuals. The motivation for considering this is equivalent to the F statistics employed to test a single change in the regression just like Chow Test (1969). Set $P_j(\beta)$ at a "leave-m-out" estimator, $\hat{\beta}_{(j)}$ as to mirror that $\hat{\beta}_{1-T}$ is not dependent on observations after point T, then

$$\hat{\beta}_{(j)} = \begin{cases} \text{estimator of } \beta \text{ using observations indexed by } t = 1, \ldots, T \text{ with} \\ t, \ldots, j+m-1. \end{cases}$$

Clearly, $\hat{\beta}_{(j)}$ by the estimators mentioned above is consistent for β_0 under suitable assumptions. Define

$$P_{a,j} = P_j\left(\hat{\beta}_{(j)}\right) \text{ for } j = 1, \ldots, T - m + 1.$$

And the empirical df of $P_{a,j} : j = 1, \ldots, T - m + 1$ is

$$\hat{F}_{P_{(a)},T}(x) = 1 - m + 1 \sum_{t=1}^{T-m+1} l\left(P_{a,j}\right).$$

Then define the test statistic P_a to be the $1 - \alpha$ sample quantile, $\hat{q}p_a$, $1 - \alpha$ and

$$\hat{q}p_a, \;\; 1 - \alpha = \text{INF } \; x \in: \hat{F}_{P_a,T}(x) \geq 1 - \alpha.$$

Thus one can reject H_0 if $P_a > \hat{q}p_a$, $1 - \alpha$.

However, P_a test's simulation performance is not satisfying since it often over-rejects the null hypothesis. Instead, P_b test could have better finite-sample properties as defined by

$$P_b = P_{T+1}\left(\hat{\beta}_{1-(T+\lceil m/2 \rceil}\right) \text{ and } P_{b,j} = P_j\left(\hat{\beta}_{(j)}\right) \text{ for } j = 1, \ldots, T - m + 1.$$

The P_b test supersedes the P_a for it is less variable as the estimator $\hat{\beta}_{1-(T+m/2\mathrm{e})}$ depends on the observation indexed by $t = T + 1, \ldots, T + 2\rceil$.

Moreover, a naturally less variable statistic P_c is dependent on the complete sample estimator $\hat{\beta}_{1-(T+m)}$:

$$P_c = P_{T+1}\left(\hat{\beta}_{1-(T+m)}\right) P_{c,j} = P\left(\hat{\beta}_{2(j)}\right) \text{ for } j = 1, \ldots, T - m + 1.$$

where

$$\hat{\beta}_{2(j)} = \begin{cases} \text{estimator of } \beta \text{ using observations indexed by } t = 1, \ldots, T \text{ with} \\ t \neq j, \ldots, j + 2\rceil - 1 \end{cases}$$

for $j = 1, \ldots, T - m + 1$. The P tests are suitable for models where errors are uncorrelated, and including weights in the statistics based on an estimator of the error covariance matrix will be advantageous if the errors are correlated. Tests with the weights are the same as $P_a - P_c$ except that $\omega = I_m$ is replaced by

$$\hat{\omega}_{1-(T+m)} = \left(1 + 1 \sum_{j=1}^{T+1} \hat{U}_{j,j+m-1} \hat{U}'_{j,j+m-1}\right)^{-1}$$

where

$$\hat{U}_{j,j+m-1} = Y_{j,j+m-1} - X_{j,j+m-1}\hat{\beta}_{1-(T+m)},$$

The estimator $\hat{\omega}_1 - (T+m)$ is an estimator of the inverse of the m covariance matrix of the errors $\omega_0^{-1} = (EU_{1-m}U'_{1-m})^{-1}$.

As for the presence of unit root errors from $t = T + 1$ to $t = T + m$ in a linear regression model with i.i.d. normal errors, the locally best invariant (LBI) test statistic is applied. According to Andrews and Kim (2003), set $A_m(k, l) = \min\{k, l\}$ for $k, l = 1, \ldots, m$ then the R tests which are aimed to solve the breakdown will be defined as

Andrews and Kim's (2003) tests are not consistent because m is fixed as $T \to \infty$ while they are simultaneously asymptotically unbiased. The power of the tests is determined by the magnitude of the breakdown and m. The former includes the magnitude of the parameter shift and the magnitude of the unit root error variance, and the larger is m, the greater is the power. That means failure to reject the null hypothesis should not be interpreted as strong evidence in favor of stable cointegration.

31.5 Monitoring Structural Breaks

Foregoing tests almost deal with in-sample data, and it may be tempting to apply them to real-time data. Under this distinct condition, however, traditional method for detecting structural breaks mainly within a historical data set normally generates comparatively large chance of mistaken instability, which leads to variant over-rejection. Typically, such a test will signal a nonexistent break with probability one. Thus due to the law of the iterated logarithm, no matter how perfectly these methods perform as one-shot type detection, their sound performance could not be translated to monitor out-of-sample stability. Conversely, monitoring test could detect the break on time without these concerns. Focusing on monitoring structural break, this section will explain how this relatively advantageous test works better timely. Plus, it is also to provide evidence to support that the good performance still holds even in long memory, as articulated before to some degree an ideal description of economy issues.

31.5.1 Monitoring in Comparably Ideal Conditions

For a linear regression $Y_t = X'\beta_t + \varepsilon_t$, $t = 1, 2, \ldots$, usually a real-time structural break test is bounded with the assumption: $\beta_t = \beta_0$ for $t = 1, 2, \ldots, m$. And the null of the test is $\beta_t = \beta_0$ for some $t \geq m + 1$. To solve this problem, Chu et al. (1996) come up with a CUSUM monitoring procedure based on the behavior of recursive residuals and a fluctuation monitoring procedure based on recursive estimates of parameters. They use sequential testing to develop tests of structural stability for real-time economic systems and widen the class of boundary functions of previous monitoring tests, which is outstanding as it matches economic research more exactly.

Let $\{S_n\}$ be the partial sum process constructed from ε_i and assume the process ε_i follows the FCLT (functional central limiting theorem) as $\lambda \to \gamma_0^{-1} m^{-1/2} S_{[m\lambda]} \Rightarrow \lambda(\lambda)$, $\lambda \in [0, \infty)$, where $S_n = \sum_{i=1}^{n} \varepsilon_i$ and $\sigma_0^2 \to n^{-1} E(S_n^2) < \infty$. Chu et al. (1996) extend Robinson and Siegmund's (1970) limiting relation as their central mechanism

$$\lim_{m\to\infty} P\left\{S_n \geq \sqrt{m} g((n/m),\ \text{forsome}\ n \geq 1\right\} = P\{W(t) \geq g(t),\ \text{for some}\ t \geq 0\}$$

where $S_n = \sum_{t=1}^{n} \varepsilon_t$, W denotes a standard Brownian motion and g is a stopping boundary satisfying some regularity conditions. This suggests a stopping function for the monitoring test. Another limiting result similar to Eq. 31.37 is

$$\lim_{m\to\infty} P\left\{|S_n| \ \geq \sqrt{m} g((n/m),\ \text{for some}\ n \geq m\right\} = P\{|W(t)|(t), t \geq 1\} \tag{31.37}$$

Robinson and Siegmund (1970) give their proof for an i.i.d. ε_i in Eq. 31.38 could hold for a certain class of continuous functions $g(t)$ that satisfy $t^{-1/2}$ $g(t)$ is ultimately nondecreasing as $t \to \infty$ and other assumptions. Though this is useful, yet such assumption is not applicable all the time for research. Segen and Sanderson (1980) provide a partial resolution for they show that under the boundary function $t^{-1/2}$ $g(t)$ is nondecreasing, stochastic sequence S_n that satisfies the FCLT continues to hold. Yet such condition turns out more restrictive than that of Robbins and Siegmund. Chu et al. (1996) propose a new limiting similar relation without imposing these restrictions. Typically, the boundary function $g(t)$ considered to be pertinent to economy is given as

$$P\left\{|W(t)| \geq \left[t\left(a^2\right) + \ln t\right)\right]^{1/2},\ \text{for some} \quad t \geq 1\right\} = 2[1 - \Phi(a) + a\phi(a)] \tag{31.38}$$

$$P\left\{|W(t)| \geq (t+1)^{1/2}\left[a^2 + \ln(t+1)\right]^{1/2},\ \text{some}\ t > 0]\right\} = \exp\left(-a^2\right) \tag{31.39}$$

where Φ and ϕ stand for the cdf and pdf, respectively, of a standard normal random variable.

A monitoring is a stopping time, determined by a detecting statistic (detector) Γ_n and a threshold g(m,n), according to $\tau_g(\Gamma_n)n$, $\Gamma_n > g(m, n)$. Firstly consider the situation where the detector is in form of CUSUM. Let $\hat{\beta} = \left(\sum_{i=1}^{n} X_i X_i^{'}\right)^{-1} \left(\sum_{i=1}^{n} X_i Y_i\right)$ be the OLS estimator at time n. Also define recursive residuals as $w_k = 0$ and $w_n = \hat{\varepsilon}/v_n^{1/2}$, $v_n = 1 + X_n^{'}(\sum_{i=1}^{n-1} X_i X_i^{'})^{-1} X_n$, $\hat{\varepsilon}_n = Y_n - X_n^{'} \hat{\beta}_{n-1}$, $n = k + 1, \ldots m, \ldots$. Thus the nth-cumulated sum of recursive residuals is $Q_t^m = \hat{\sigma}^{-1} \sum_{i=k}^{n} \omega_i = \hat{\sigma}^{-1} \sum_{i=k}^{k+\left[\widetilde{m}\right)t]} \omega_i$, for $(n-k)/\widetilde{m} < (n-k+1)/\widetilde{m}$, where $\hat{\sigma}$ is a consistent if σ, $\widetilde{m} = (m-k)$ is the integer of $(\widetilde{m})t$. Thus under H_0,

$$t \to \widetilde{m}^{-1/2} Q_t^m, \ \ t \in [0, \infty) \Rightarrow t(t), \ \ t \in (0, \infty) \tag{31.40}$$

where "⇒" denotes the weak convergence of the associated probability measures. According to Chu et al. (1996), as the monitoring starts as $m + 1$, define $\widetilde{Q}_t^m = \hat{\sigma} \sum_{i=m+1}^{k+[\widetilde{m}(1+t)]} \omega_i$, $t \in [0, \infty)$. In particular, for $n/(m - k) < (n + 1)/(m - k)$, $\widetilde{Q}_n^m = \hat{\sigma}_{-1} \left(\sum_{i=k}^{m+n} \omega_i - \sum_{i=k}^{m} \omega_i \right)$, $n \geq 1$. It follows that $t \to \widetilde{m}^{-1/2} \widetilde{Q}_n^m \Rightarrow t \to [W(t+1) - W(t)]$, $t \in [0, \infty)$. Since this is a Brownian motion, it can be written in the form of Eq. 31.36

$$\begin{aligned} &\lim_{m\to\infty} P\left\{ |Q_n^m| \ \geq \sqrt{m-k} g\left(\frac{n}{m-k}\right), \ \text{for some} \ \ n \geq 1 \right\} \\ &= P\left\{ |W(t)| \ \geq g(t), \ \text{for some} \ t \geq 0 \right\} \end{aligned}.$$

Its mechanism can be explained as that once the path of $\left| \widetilde{Q}_n^m \right|$ crosses the boundary $(m - k)^{1/2}$ $g(n/m - k)$, it will reject the null hypothesis and imply that the model suitable for historical period is no longer reliable. Moreover, an extension of this can be applied to FL monitoring procedure as

$$\begin{aligned} &\lim_{m\to\infty} P\left\{ |Q_n^m| \geq \sqrt{m-k} g\left(\frac{n}{m-k}\right), \ \text{for some} \ n \geq m \right\} \\ &= P\left\{ |W(t)| \geq g(t), \ \text{for some} \ t \geq 1 \right\} \end{aligned} \tag{31.41}$$

It needs notice that CUSUM algorithms, except those recursive residuals, also hold for such mechanism, e.g., LR statistics for dependent sequences. Yet LR detector sometimes could be superfluously complex in general and leads to a complicated computation. According to the above statement, a CUSUM monitoring can be implemented as follows:

Suppose (i) $Y_t = X_t' \beta_0 + \varepsilon_t$, $t = 1, \ldots, m + 1, \ldots$, where X_t is a $k \times 1$ random vector such that $m^{-1} \sum_{t=1}^{m} X_t$ and $m^{-1} \sum_{t=1}^{m} X_t X_t'$ converge in probability to b, a non-stochastic $k \times 1$ vector and M, a k matrix of full rank, respectively; (ii) ε_t is a martingale difference sequence with respect to a sequence of σ-algebra F_t such that $E(\varepsilon^2) < \infty$ and $E(\varepsilon_t^2 | F_{t-1}) = \sigma_0^2$ for all t, where F_t is generated by…, (Y_{t-2}, X_{t-1}'), (Y_{-1}, X_t'); (iii) the sequence $X_t \varepsilon_t$ obeys the functional central limit theorem, then

$$\begin{aligned} &\lim_{m\to\infty} P\left\{ \left| \widetilde{Q}_n^m \right| \geq \sqrt{n+m-k} \left[a^2 + ln\left(\frac{n+m-k}{m-k} \right) \right]^{1/2}, \ \text{for some,} \ \ n \geq 1 \right\} \\ &= \exp(-a^2/2); \end{aligned} \tag{31.42}$$

$$\begin{aligned} &\lim_{m\to\infty} P\left\{ \left| \widetilde{Q}_n^m \right| \geq \sqrt{m-k} \left(\frac{n}{m-k} \right)^{1/2} \left[a^2 + ln\left(\frac{n}{m-k} \right) \right]^{1/2}, \ \text{some} \ n \right\} \\ &= 2[1 - \Phi[a] + a\phi(a)] \end{aligned} \tag{31.43}$$

Additionally, by the right-handed side of Eqs. 31.42 and 31.43, the asymptotic size control of the CUSUM monitoring is fairly effortless. For CUSUM monitoring based on $\widetilde{Q}_n^m$, the 10 % and 5 % asymptotic size correspond to a^2 = 4.6 and 6, respectively. Equally handily, we obtain the 10 % and 5 % asymptotical size of the CUSUM monitoring based on Q_m by setting $a_2 = 6.25$ and 7.78 in Eq. 31.43, respectively.

Furthermore, from Eq. 31.36 we also can construct the fluctuation monitoring of sequential parameter estimates. The key condition is that $X_t\varepsilon_t$ obeys the multivariate FCLT as

$$\lambda^{-1/2}V_0^{-1/2}\sum_{t=1}^{[m\lambda]}X_t\varepsilon_t \Rightarrow \lambda - W(\lambda),\ \ \lambda \in [0,\infty)$$

where $V_0 = lim_{m\to\infty}m^{-1}E(S_mS_m^{'})$ with $S_m = \sum_{t=1}^{m}X_t\varepsilon_t$, and $\underline{W}(\lambda)$ is a k-dimensional Wiener process. Following this, the fluctuation detector can be defined as

$$\hat{Z}_n = nD_m^{-1/2}\left(\hat{\beta}_n - \hat{\beta}_m\right), n \tag{31.44}$$

where $D_m = M_m^{-1}V_0\,M_{mp}^{-1}$ is $O_p(1)$ and uniformly positive definite such that

$$\left(\sum_{t=1}^{m}X_tX_t^{'}/m\right) - M_m \to 0.$$

The essential ingredient of this FL detector is the deviation of the updated parameter estimate $\hat{\beta}_n$ from the historical parameter estimate $\hat{\beta}_m$. Then the conclusion can be obtained as

$$(a) : \lambda^{-1/2}\hat{z}_{[m\lambda]} \Rightarrow \lambda \to \underline{W}^0(\lambda),\ \ \lambda \in [1,\infty),$$

where $\underline{W}^0(\lambda)$ is a k-dimensional Brownian bridge;

$$\begin{aligned}(b) : \lim{}_{m\to\infty}P&\left\{\left|\hat{z}^i_{[m\lambda]}\right| \geq m^{1/2}\frac{n-m}{m}\left[\frac{n}{n-m}\left[a^2 + In\left(\frac{n}{n-m}\right)\right]\right]^{1/2}\right\},\\ &(\text{for some} \geq m \quad \text{and some } i)\\ &= 1 - [1 - 2[1 - \Phi(a) + a\phi(a)]]^k,\end{aligned}$$

where $\hat{z}^i_{[m\lambda]}$ is the ith component of $\hat{Z}^i_{[m\lambda]}$. This conclusion summarizes the monitoring procedure based on the fluctuation of $\hat{\beta}_n(n > m)$ relative to $\hat{\beta}_m$. Similarly, given the number of regressors, k, and arbitrary probability of type I error, the monitoring boundary can be calculated with a derivative a^2. More details about cases of the one-time parameter shift in both FL and CUSUM monitoring are referred to Chu et al. (1996).

31.6 Concluding Remarks

In this chapter, we discuss two classes of structural breaks – the retrospective tests and the monitoring tests. From the investor's points of view, it is crucial to use suitable structural break tests to detect turning points of the financial and macroeconomics data accurately and further adjust the portfolios immediately. With the increase of the data property caused by highly frequent market fluctuations, the currently existing structural break tests could not fully detect locations of breaks without the issue of size distortion. More precisely, more and more sophisticated structural breaks tests with respect to any possible market performances would be expected to be created in the near future.

References

Andrews, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica, 61*, 821–856.

Andrews, D. W. K., & Kim, J. (2003). End-of-sample cointegration breakdown tests. Cowles Foundation Discussion Paper No. 1404.

Bai, J., & Perron, P. (1996). Computation and analysis of multiple structural change models. Manuscript in Preparation, Universite de Montreal.

Bai, J., & Perron, P. (1998). Estimation and testing linear models with multiple structural changes. *Econometrica, 66*, 47–78.

Beran, J. (1991). M-estimators of location for data with slowly decaying serial correlations. *Biometrika, 76*, 261–269.

Beran, J., & Terrin, N. (1996). Testing for a change of the long-memory parameter. *Biometrika, 83*, 627–638.

Campbell, J. Y., & Perron, P. (1991). Pitfalls and opportunities: What macroeconomists should know about unit roots. *National Bureau of Economic Research Macroeconomics Annual, 6*, 201–210.

Cavaliere, G. (2004). Unit root tests under time-varying variances. *Econometric Reviews, 23*, 259–292.

Cavaliere, G., & Taylor, A. M. R. (2006). Testing for unit roots in time series models with non-stationary volatility. *Journal of Econometrics, 140*, 919–947.

Chan, N. H., & Tran, L. T. (1989). On the first order autoregressive process with infinite variance. *Econometric Theory, 5*, 354–362.

Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica, 28*, 591–605.

Chu, C. J., Stinchcombe, M., & White, H. (1996). Monitoring structural change. *Econometrica, 64*, 1045–1065.

Clements, M. P., & Hendry, D. F. (1998). Forecasting economic processes. *International Journal of Forecasting, 14*, 111–131.

Clements, M. P., & Hendry, D. F. (1999). On winning forecasting competitions in economics. *Spanish Economic Review, 1*, 123–160.

Davies, R. B. (1977). Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika, 64*, 247–254.

Diba, B. T., & Grossman, H. I. (1988). Explosive rational bubbles in stock prices? *The American Economic Review, 78*, 520–530.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association, 74*, 427–431.

Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica, 49*, 1057–1072.

Evans, G. W. (1991). Pitfalls in testing for explosive bubbles in asset prices. *American Economic Review, 81*, 922–930.

Granger, C. (1980). Long memory relationships and aggregation of dynamics models. *Journal of Econometrics, 14*, 227–238.

Granger, C. W. J., & Joyeux, R. (1980). An introduction to long-memory time series and fractional differencing. *Journal of Time Series, 1*, 15–29.

Hamori, S., & Tokihisa, A. (1997). Testing for a unit root in the presence of a variance shift. *Economics Letters, 57*, 245–253.

Hansen, B. E. (1992). Tests for parameter instability in regressions with I(1) processes. *Journal of Business and Economic Statistics, 10*, 321–336.

Hidalgo, J., & Robinson, P. M. (1996). Testing for structural change in a long-memory environment. *Journal of Econometrics, 70*, 159–174.

Horvath, L., & Kokoszka, P. S. (2003). A bootstrap approximation to a unit root test statistic for heavy tailed observations. *Statistics and Probability Letters, 62*, 163–173.

Hsiao, C., & Robinson, P. M. (1978). Efficient estimation for a dynamic error-shock model. *International Economic Review, 19*, 467–480.

Jach, A., & Kokoszka, P. S. (2004). Subsampling unit root tests for heavy-tailed observations. *Methodology and Computing in Applied Probability, 6*, 73–97.

Kazakevicius, V., Leipus, R., & Viano, M. (2004). Stability of random coefficient ARCH models and aggregation schemes. *Journal of Econometrics, 120*, 139–158.

Kim, T. H., Leybourne, S., & Newbold, P. (2002). Unit root tests with a break in innovation variance. *Journal of Econometrics, 109*, 365–387.

Kindleberger, C. P. (2005). *Manias, panics, and crashes: A history of financial crises*. New Jersey: Wiley.

Kirman, A., & Teyssiere, G. (2005). Testing for bubbles and change-points. *Journal of Economic Dynamic Control, 29*, 765–799.

Kokoszka, P. S., & Parfionovas, A. (2004). Bootstrap unit root tests for heavy tailed time series. In S. T. Rachev (Ed.), *Handbook of computational and numerical methods in finance* (pp. 175–197). Boston: Birkhäuser.

Kuan, C., & Hsu, C. (1998). Change-point estimation of fractionally integrated processes. *Journal of Time Series Analysis, 19*, 693–708.

Leybourne, S., Taylor, R., & Kim, T. (2006). Cusum of squares-based tests for a change in persistence. *Journal of Time Series Analysis, 28*, 408–433.

Lutkepohl, H. (1988). Prediction tests for structural stability. *Journal of Econometrics, 39*, 267–296.

Lutkepohl, H. (1989). Prediction tests for structural stability of multiple time series. *Journal of Business Economic Statistics, 7*, 129–135.

Mandelbrot, B. B., & Vanness, J. W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Review, 10*, 422–437.

Paparoditis, E., & Politis, D. (2003). Residual-based block bootstrap for unit root testing. *Econometrica, 71*, 813–855.

Perron, P., & Ng, S. (1996). Useful modifications to some unit root tests with dependent errors and their local asymptotic properties. *Review of Economic Studies, 63*, 435–463.

Phillips, P. C. B., & Hansen, B. E. (1990). Statistical inference in instrumental variables regression with I(1) processes. *Review of Economic Studies, 57*, 99–125.

Phillips, P. C. B., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika, 75*, 335–346.

Phillips, P. C. B., Wu, Y., & Yu, J. (2009). Explosive behavior in the 1990s Nasdaq: When did exuberance escalate asset values?. Cowles Foundation Discussion Paper No.1699

Phillips P. C. B., Shi, S. & Yu, J. (2011). Testing for multiple bubbles. Singapore Management University working paper.

Quandt, R. E. (1960). Tests of the hypothesis that a linear regression system obeys two separate regimes. *Journal of the American Statistical Association, 55*, 324–330.

Quintos, C. E., & Phillips, P. C. B. (1993). Parameter constancy in cointegrating regressions. *Empirical Economics, 18*, 675–706.

Robinson, H., & Siegmund, D. (1970). Boundary crossing probabilities for the Wiener process and sample sums. *The Annals of Mathematical Statistics, 41*, 1410–1429.

Robinson, P. M., & Zaffaroni, P. (1998). Nonlinear time series with long memory: A model for stochastic volatility. *Journal of Statistical Planning and Inference, 68*, 359–371.

Segen, J., & Sanderson, A. (1980). Detection of changes in a time series. *IEEE Transaction in Information Theory, 26*, 249–257.

Shiller, R. J. (2000). *Irrational exuberance*. Princeton: Princeton University Press.

Sibbertsen, P., & Kruse, R. (2009). Testing for a break in persistence under long-range dependencies. *Journal of Time Series Analysis, 30*, 268–285.

Stock, J. H. (1994). Unit roots, structural breaks and trends, Chapter 46. In *Handbook of econometrics* (pp. 2740–2841). Amsterdam: Elsevier

Taylor, S. (1986). *Modeling financial time series*. New Jersey: Wiley.

Welch, I., & Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies, 21*, 1455–1508.

Wu, G. & Xiao, Z. (2008). Are there speculative bubbles in stock markets? Evidence from an alternative approach. *Statistics and its Interface, 1*, 307–320.

Xiao, Z., & Phillips, P. C. B. (2002). A CUSUM test for cointegration using regression residuals. *Journal of Econometrics, 108*, 43–61.

Yajima, Y. (1988). On estimation of a regression model with long-memory stationary errors. *Annals of Statistics, 16*, 791–807.

Zaffaroni, P. (2000). Contemporaneous aggregation of GARCH processes. *Journal of Time Series Analysis, 28*, 521–544.

Zaffaroni, P. (2006). Aggregation and memory of models of changing volatility. *Journal of Econometrics, 136*, 237–249.

Consequences for Option Pricing of a Long Memory in Volatility

32

Stephen J. Taylor

Contents

S.J. Taylor
Lancaster University Management School, Lancaster, UK
e-mail: s.taylor@lancaster.ac.uk

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_32,

Abstract

Conditionally heteroscedastic time series models are used to describe the volatility of stock index returns. Volatility has a long memory property in the most general models and then the autocorrelations of volatility decay at a hyperbolic rate; contrasts are made with popular, short memory specifications whose autocorrelations decay more rapidly at a geometric rate.

Options are valued for ARCH volatility models by calculating the discounted expectations of option payoffs for an appropriate risk-neutral measure. Monte Carlo methods provide the expectations. The speed and accuracy of the calculations is enhanced by two variance reduction methods, which use antithetic and control variables. The economic consequences of a long memory assumption about volatility are documented, by comparing implied volatilities for option prices obtained from short and long memory volatility processes.

Results are given for options on the S & P 100-share index, with lives up to 2 years. The long memory assumption is found to have a significant impact upon the term structure of implied volatilities and a relatively minor impact upon smile shapes. These conclusions are important because evidence for long memory in volatility has been found in the prices of many assets.

Keywords

ARCH models • Implied volatility • Index options • Likelihood maximization • Long memory • Monte Carlo • Option prices • Risk-neutral pricing • Smile shapes • Term structure • Variance reduction methods

32.1 Introduction

Long memory effects in a stochastic process are effects that decay too slowly to be explained by stationary processes defined by a finite number of autoregressive and moving-average terms. Long memory is often represented by fractional integration of shocks to the process, which produces autocorrelations which decrease at a hyperbolic rate compared with the faster geometric rate of stationary ARMA processes.

Long memory in volatility occurs when the effects of volatility shocks decay slowly. This phenomenon can be identified from the autocorrelations of measures of realized volatility. Two influential examples are the study of absolute daily returns from stock indices by Ding et al. (1993) and the investigation of daily sums of squared 5-min returns from exchange rates by Andersen et al. (2001b).

Stochastic volatility causes option prices to display both smile and term structure effects. An implied volatility obtained from the Black-Scholes formula then depends on both the exercise price and the time until the option expires. Exact calculation of smile and term effects is only possible for special volatility processes, with the results of Heston (1993) being a notable example. Monte Carlo methods are usually necessary when the volatility process has a long memory and these were first applied by Bollerslev and Mikkelsen (1996, 1999).

This chapter documents the economic consequences of a long memory assumption about volatility. This is achieved by comparing implied volatilities for option prices obtained from short and long memory specifications. It is necessary to use a long history of asset prices when applying a long memory model and this chapter uses levels of the S & P 100 index from 1984 to 1998. For this data it is found that a long memory assumption has a significant economic impact upon the term structure of implied volatilities and a relatively minor impact upon smile effects. Some related empirical results are provided by Ohanissian et al. (2004).

Option traders have to make assumptions about the volatility process. The effects of some assumptions are revealed by the prices of options with long lives. Bollerslev and Mikkelsen (1999) find that the market prices of exchange traded options on the S & P 500 index, with lives between 9 months and 3 years, are described more accurately by a long memory pricing model than by the short memory alternatives. Thus these option prices reflect the long memory phenomenon in volatility, although it is found that significant biases remain unexplained.

Three explanatory sections precede the illustrative option pricing results in Sect. 32.5. Section 32.2 defines and characterizes long memory and then reviews the empirical evidence for these characteristics in volatility. The empirical evidence for the world's major markets appears compelling and explanations for the source of long memory effects in volatility are summarized.

Section 32.3 describes parsimonious volatility models which incorporate long memory, either within an ARCH or a stochastic volatility framework. The former framework is easier to use and we focus on applying the fractionally integrated extension of the exponential GARCH model, which is known by the acronym FIEGARCH. An important feature of applications is the unavoidable truncation of an autoregressive component of infinite order. Empirical results are provided for 10 years of S & P 100 returns.

Section 32.4 provides the option pricing methodology. Contingent claim prices are obtained by simulating terminal payoffs using an appropriate risk-neutral measure. Numerical methods enhance the accuracy of the simulations and these are described in Appendix 2.

Section 32.5 compares implied volatilities for European option prices obtained from short and long memory volatility specifications, for hypothetical S & P 100 options whose lives range from 1 month to 2 years. Options are valued on 10 dates, one per annum from 1989 to 1998. The major impact of the long memory assumption is seen to be the very slow convergence of implied volatilities to a limit as the option life increases. This convergence is so slow that the limit cannot be estimated precisely. Section 32.6 contains conclusions.

32.2 Long Memory

32.2.1 Definitions

Definitions which categorize stochastic processes as having either a short memory or a long memory can be found in Brockwell and Davis (1991), Baillie (1996), Granger and Ding (1996), and Taylor (2005). The fundamental characteristic of a long memory process is that dependence between variables separated by τ time units does not decrease rapidly as τ increases.

Consider a covariance stationary stochastic process $\{x_t\}$ that has variance σ^2 and autocorrelations ρ_τ, spectral density $f(\omega)$, and n-period variance ratios V_n defined by

$$\rho_\tau = \mathrm{cor}(x_t, x_{t+\tau}), \tag{32.1}$$

$$f(\omega) = \frac{\sigma^2}{2\pi} \sum_{\tau=-\infty}^{\infty} \rho_\tau \cos(\tau\omega), \quad \omega > 0, \tag{32.2}$$

$$V_n = \frac{\mathrm{var}(x_{t+1} + \ldots + x_{t+n})}{n\sigma^2} = 1 + 2\sum_{\tau=1}^{n-1} \frac{n-\tau}{n}\rho_\tau \tag{32.3}$$

Then a covariance stationary process is here said to have a short memory if $\sum_{\tau=1}^{n} \rho_\tau$ converges as $n \to \infty$; otherwise it is said to have a long memory. A short memory process then has

$$\sum_{\tau=1}^{n} \rho_\tau \to C_1, \; f(\omega) \to C_2, \; V_n \to C_3, \;\; \text{as} \;\; n \to \infty, \;\; \omega \to 0, \tag{32.4}$$

for constants C_1, C_2, C_3. Examples are provided by stationary ARMA processes. These processes have geometrically bounded autocorrelations, so that $|\rho_\tau| \le C\phi^\tau$ for some $C > 0$ and $1 > \phi > 0$, and hence Eq. 32.4 is applicable.

In contrast to the above results, all the limits given by Eq. 32.4 do not exist for a typical covariance stationary long memory process. Instead, it is typical that the autocorrelations have a hyperbolic decay, the spectral density is unbounded for low frequencies, and the variance ratio increases without limit. Appropriate limits are then provided for some positive $d < \frac{1}{2}$ by

$$\frac{\rho_\tau}{\tau^{2d-1}} \to D_1, \; \frac{f(\omega)}{\omega^{-2d}} \to D_2, \; \frac{V_n}{n^{2d}} \to D_3, \;\; \text{as} \;\; n \to \infty, \;\; \omega \to 0, \tag{32.5}$$

for positive constants D_1, D_2, D_3. The limits given by Eq. 32.5 characterize the stationary long memory processes that are commonly used to represent long memory in volatility. The fundamental parameter d can be estimated from data using a regression, either of $\ln\left(\hat{f}(\omega)\right)$ on ω or of $\ln\left(\hat{V}_n\right)$ on n as in, for example, Andersen et al. (2001a)

32.2.2 Fractionally Integrated White Noise

An important example of a long memory process is a stochastic process $\{y_t\}$ which requires fractional differencing to obtain a set of independent and identically distributed residuals $\{\varepsilon_t\}$. Following Granger and Joyeux (1980) and Hosking (1981), such a process is defined using the filter

$$(1-L)^d = 1 - dL + \frac{d(d-1)}{2!}L^2 - \frac{d(d-1)(d-2)}{3!}L^3 + \ldots \tag{32.6}$$

where L is the usual lag operator, so that $Ly_t = y_{t-1}$. Then a fractionally integrated white noise (FIWN) process $\{y_t\}$ is defined by

$$(1-L)^d y_t = \varepsilon_t \tag{32.7}$$

with the ε_t assumed to have zero mean and variance σ_ε^2. Throughout this chapter it is assumed that the differencing parameter d is constrained by $0 \le d < 1$.

The mathematical properties of FIWN are summarized in Baillie (1996). The process is covariance stationary if $d < \frac{1}{2}$ and then the following results apply. First, the autocorrelations are given by

$$\rho_1 = \frac{d}{1-d}, \quad \rho_2 = \frac{d(d+1)}{(1-d)(2-d)}, \quad \rho_3 = \frac{d(d+1)(d+2)}{(1-d)(2-d)(3-d)}, \ldots\ldots \tag{32.8}$$

or, in terms of the gamma function,

$$\rho_\tau = \frac{\Gamma(1-d)\Gamma(\tau+d)}{\Gamma(d)\Gamma(\tau+1-d)}, \tag{32.9}$$

with

$$\frac{\rho_\tau}{\tau^{2d-1}} \to \frac{\Gamma(1-d)}{\Gamma(d)} \quad \text{as} \quad \tau \to \infty. \tag{32.10}$$

Second, the spectral density is

$$f(\omega) = \frac{\sigma_\varepsilon^2}{2\pi}\left|1 - e^{-i\omega}\right|^{-2d} = \frac{\sigma_\varepsilon^2}{2\pi}\left[2\sin\left(\frac{\omega}{2}\right)\right]^{-2d}, \quad \omega > 0, \tag{32.11}$$

so that

$$f(\omega) \cong \frac{\sigma_\varepsilon^2}{2\pi}\omega^{-2d} \text{ for } \omega \text{ near } 0. \tag{32.12}$$

Also,

$$\frac{V_n}{n^{2d}} \to \frac{\Gamma(1-d)}{(1+2d)\Gamma(1+d)} \text{ as } n \to \infty. \tag{32.13}$$

When $d \geq \frac{1}{2}$, the FIWN process has infinite variance and thus the autocorrelations are not defined, although the process has some stationarity properties for $\frac{1}{2} \leq d < 1$.

32.2.3 Evidence for Long Memory in Volatility

When returns r_t can be represented as $r_t = \mu + \sigma_t u_t$, with σ_t representing volatility and independent of an i.i.d. standardized return u_t, it is often possible to make inferences about the autocorrelations of volatility from the autocorrelations of either $|r_t - \mu|$ or $(r_t - \mu)^2$; see Taylor (2005, 2008). In particular, evidence for long memory in powers of daily absolute returns is also evidence for long memory in volatility. Ding et al. (1993) observe hyperbolic decay in the autocorrelations of powers of daily absolute returns obtained from US stock indices. Dacorogna et al. (1993) observe a similar hyperbolic decay in 20-min absolute exchange rate returns. Breidt et al. (1998) find that spectral densities estimated from the logarithms of squared index returns have the shape expected from a long memory process at low frequencies. Ohanissian et al. (2008) accept the null hypothesis of long memory for exchange rate volatility, by assessing the long memory implication that d is invariant under temporal aggregation. Further evidence for long memory in volatility has been obtained by fitting appropriate fractionally integrated ARCH models and then testing the null hypothesis $d = 0$ against the alternative $d > 0$. Bollerslev and Mikkelsen (1996) use this test to support long memory models for US stock index volatility.

Direct evidence for long memory in volatility uses high-frequency data to construct accurate estimates of the volatility process. The estimated quadratic variation of the logarithm of the price process during a 24-h period denoted by t can be estimated from intraday returns $r_{t,j}$ by calculating

$$\hat{\sigma}_t^2 = \sum_{j=1}^{N} r_{t,j}^2. \tag{32.14}$$

The estimate $\hat{\sigma}_t^2$ will be very close to the integral of the latent volatility during the same 24-h period providing N is large but not so large that the bid-ask spread and other microstructure effects introduce bias into the estimate. Using 5-min returns provides conclusive evidence for long memory effects in the estimates $\hat{\sigma}_t^2$ in four studies: Andersen et al. (2001b) for 10 years of DM/$ and Yen/$ rates; Andersen et al. (2001b) for 5 years of stock prices for the 30 components of the Dow-Jones index; Ebens (1999) for 15 years of the same index; and Areal and Taylor (2002) for 8 years of FTSE-100 stock index futures prices. These papers provide striking evidence that time series of estimates $\hat{\sigma}_t^2$ display all three properties of a long memory process: hyperbolic decay in the autocorrelations, spectral densities at low frequencies that are proportional to ω^{-2d}, and variance ratios whose logarithms are very close to linear functions of the aggregation period n. It is also found that estimates of d are between 0.3 and 0.5, with most estimates close to 0.4.

32.2.4 Explanations of Long Memory in Volatility

Granger (1980) shows that long memory can be a consequence of aggregating short memory processes; specifically if AR(1) components are aggregated and if the AR(1) parameters are drawn from a beta distribution, then the aggregated process converges to a long memory process as the number of components increases. Andersen and Bollerslev (1997) develop Granger's theoretical results in more detail for the context of aggregating volatility components and also provide supporting empirical evidence obtained from only 1 year of 5-min returns. It is plausible to assert that volatility reflects several sources of news, that the persistence of shocks from these sources depends on the source, and hence that total volatility may follow a long memory process. Scheduled macroeconomic news announcements are known to create additional volatility that is very short-lived (Ederington and Lee 1993), while other sources of news that have a longer impact on volatility are required to explain volatility clustering effects that last several weeks.

Gallant et al. (1999) estimate a volatility process for daily IBM returns that is the sum of only two short memory components, yet the sum is able to mimic long memory. They also show that the sum of a particular pair of AR(1) processes has a spectral density function very close to that of fractionally integrated white noise with $d = 0.4$ for frequencies $\omega \geq 0.01\pi$. Consequently, evidence for long memory may be consistent with a short memory process that is the sum of a small number of components whose spectral density happens to resemble that of a long memory process except at extremely low frequencies. Attempts to distinguish between true long memory and short memory models which mimic long memory behavior include Ohanissian et al. (2008) and Pong et al. (2008).

Barndorff-Nielsen and Shephard (2001) model volatility in continuous time as the sum of a few short memory components. Their analysis of 10 years of 5-min DM/$ returns shows that the sum of four volatility processes is able to provide an excellent match to the autocorrelations of squared 5-min returns, which exhibit the long memory property of hyperbolic decay.

32.3 Long Memory Volatility Models

A general set of long memory stochastic processes can be defined by first applying the filter $(1 - L)^d$ and then assuming that the filtered process is a stationary ARMA (p, q) process. This defines the ARFIMA (p, d, q) models of Granger (1980), Granger and Joyeux (1980), and Hosking (1981). This approach can be used to obtain long memory models for volatility, by extending various specifications of short memory volatility processes. We consider both ARCH and stochastic volatility specifications.

32.3.1 ARCH Specifications

The conditional distributions of returns r_t are defined for ARCH models using information sets I_{t-1} which are here assumed to be previous returns $\{r_{t-i}, i \geq 1\}$, conditional mean functions $\mu_t(I_{t-1})$, conditional variance functions $h_t(I_{t-1})$, and a probability distribution D for standardized returns z_t. Then the terms

$$z_t = \frac{r_t - \mu_t}{\sqrt{h_t}} \tag{32.15}$$

are independently and identically distributed with distribution D and have zero mean and unit variance.

Baillie (1996) and Bollerslev and Mikkelsen (1996) both show how to define a long memory process for h_t by extending either the GARCH models of Bollerslev (1986) or the exponential ARCH models of Nelson (1991). The GARCH extension cannot be recommended because the returns process then has infinite variance for all positive values of d, which is incompatible with the stylized facts for asset returns. For the exponential extension, however, $\ln(h_t)$ is covariance stationary for $d < \frac{1}{2}$; it may then be conjectured that the returns process has finite variance for particular specifications of h_t.

Like Bollerslev and Mikkelsen (1996, 1999), this chapter applies the FIEGARCH(1, d, 1) specification:

$$\ln(h_t) = \alpha + (1 - \phi L)^{-1}(1 - L)^{-d}(1 + \psi L)g(z_{t-1}), \tag{32.16}$$

$$g(z_t) = \theta z_t + \gamma(|z_t| - C), \tag{32.17}$$

with α, ϕ, d, ψ respectively denoting the location, autoregressive, differencing, and moving-average parameters of $\ln(h_t)$. The i.i.d. residuals $g(z_t)$ depend on a symmetric response parameter γ and an asymmetric response parameter θ which allows the conditional variances to depend on the signs of the terms z_t; these residuals have zero mean because C is defined to be the expectation of $|z_t|$. The EGARCH(1,1) model of Nelson (1991) is given by $d = 0$. If $\phi = \psi = 0$ and $d > 0$, then $\ln(h_t) - \alpha$ is a fractionally integrated white noise process. In general, $\ln(h_t)$ is an ARFIMA(1, d, 1) process.

Calculations using Eq. 32.16 require truncation of a series expansion in the lag operator L. The relevant formulae are listed in Appendix 1. We apply the following ARMA(N, 1) approximation:

$$\ln(h_t) = \alpha + \sum_{j=1}^{N} b_j \left[\ln\left(h_{t-j}\right) - \alpha\right] + g(z_{t-1}) + \psi g(z_{t-2}), \tag{32.18}$$

with the coefficients b_j defined by Eqs. 32.31 and 32.32.

32.3.2 Estimates for the S & P 100 Index

Representative parameters are used in Sect. 32.5 to illustrate the option pricing consequences of long memory in volatility. These parameters are estimated from daily returns r_t for the S & P 100 index, excluding dividends, calculated from index levels p_t as $r_t = \ln\left(p_t/p_{t-1}\right)$.

The conditional variances are evaluated for $t \geq 1$ by setting $N = 1{,}000$ in Eq. 32.18, with $\ln\left(h_{t-j}\right)$ replaced by α and $g(z_{t-j})$ replaced by zero whenever $t - j \leq 0$. The log-likelihood function is calculated for the 2,528 trading days during the 10-year estimation period from 3 January 1989 to 31 December 1998, which corresponds to the times $1,221 \leq t \leq 3,748$ for our dataset; thus the first 1,220 returns are reserved for the calculation of conditional variances before 1989 which are needed to evaluate the subsequent conditional variances.

Results are first discussed when returns have a constant conditional mean which is estimated by the sample mean. The conditional variances are obtained recursively from Eqs. 32.15, 32.17, 32.18, and 32.32. The conditional distributions are assumed to be normal when defining the likelihood function. This assumption is known to be false but it is made to obtain consistent parameter estimates (Bollerslev and Wooldridge 1992). Preliminary maximizations of the likelihood showed that a suitable value for $C = E[|z_t|]$ is 0.737, compared with $\sqrt{2/\pi} \cong 0.798$ for the standard normal distribution. They also showed that an appropriate value of the location parameter α of $\ln\left(h_t\right)$ is -9.56; the log-likelihood is not sensitive to minor deviations from this level because α is multiplied by a term $1 - \sum_{j=1}^{N} b_j$ in Eq. 32.18 which is small for large N. Consequently, the results summarized in Table 32.1 are given by maximizing the log-likelihood function over some or all of the parameters $\theta, \gamma, \phi, \psi, d$.

The estimates of θ and γ provide the usual result for a series of US stock index returns that changes in volatility are far more sensitive to the values of negative returns than those of positive returns, as first reported by Nelson (1991). When z_t is negative, $g(z_t) = (\gamma - \theta)(-z_t) - \gamma C$, otherwise $g(z_t) = (\gamma + \theta)z_t - \gamma C$. The ratio $(\gamma - \theta)/(\gamma + \theta)$ is at least 4 and hence is substantial for the estimates presented in Table 32.1.

Table 32.1 Parameter estimates for short and long memory ARCH models

Model	Constraints	θ	γ	ϕ	ψ	d	RLL
AR(1)	$\psi = d = 0$	−0.06	0.10	0.982			0.0
ARMA(1,1)	$d = 0$	−0.09	0.15	0.988	−0.44		3.0
FI(d)	$\phi = \psi = 0$	−0.12	0.19			0.66	13.9
ARFI(1, d)	$\psi = 0$	−0.11	0.17	0.30		0.59	17.2
ARFIMA(1, d, 1)	None	−0.10	0.15	−0.16	0.98	0.57	21.7
ARFIMA(1, d, 1)	$\phi + \psi + d \leq 1$	−0.12	0.18	−0.27	0.68	0.59	19.7
ARFIMA(1, d, 1)	$d = 0.4$, $\phi + \psi \leq 0.6$	−0.11	0.18	0.64	−0.04	0.4	11.4

Parameters are estimated by maximizing the log-likelihood of daily returns from the S & P 100 index, from 3 January 1989 to 31 December 1998. Returns are modelled as
$r_t = \mu + \sqrt{h_t} z_t$,
$\ln(h_t) = \alpha + (1 - \phi L)^{-1}(1 - L)^{-d}(1 + \psi L) g(z_{t-1})$,
$g(z_t) = \theta z_t + \gamma(|z_t| - C)$
The z_t are assumed to be i.i.d., standard normal variables when defining the likelihood function. The values $C = 0.737$ and $\alpha = -9.56$ are used for all likelihood calculations. The relative log-likelihood (RLL) for a model equals the maximum log-likelihood (MLL) for that model minus the MLL for the AR(1) model. The MLL for the AR(1) model is 8561.6

The first two rows of Table 32.1 report estimates for short memory specifications of the conditional variance. The AR(1) specification has a persistence of 0.982 which is typical for this volatility model. The ARMA (1,1) specification has an additional parameter and increases the log-likelihood by 3.0. The third row shows that the fractional differencing filter alone ($d > 0$, $\phi = \psi = 0$) provides a better description of the volatility process than the ARMA(1,1) specification; with $d = 0.66$ the log-likelihood increases by 10.9. A further increase of 7.8 is then possible by optimizing over all three volatility parameters, d, ϕ, and ψ, to give the parameter estimates[1] in the fifth row of Table 32.1.

The estimates for the most general specification identify two issues of concern. First, d equals 0.57 for our daily data which is more than the typical estimate of 0.4 produced by the studies of higher frequency data mentioned in Sect. 32.2.3. The same issue arises in Bollerslev and Mikkelsen (1996) with d estimated as 0.63 (standard error 0.06) from 9,559 daily returns of the S & P 500 index, from 1953 to 1990. Second, the sum $d + \phi + \psi$ equals 1.39. As this sum equals ψ_1 in Eqs. 32.35 and 32.36, more weight is then given to the volatility shock at time $t - 2$ than to the shock at time $t - 1$ when calculating $\ln(h_t)$. This is counterintuitive. To avoid this outcome, the constraint $d + \phi + \psi \leq 1$ is applied and the results given in the penultimate row of Table 32.1 are obtained. The log-likelihood is then reduced by 2.0. Finally, if d is constrained to be 0.4, then the log-likelihood is reduced by an additional 8.3.

[1] A conservative robust standard error for our estimate of d is 0.12, using information provided by Bollerslev and Mikkelsen (1996).

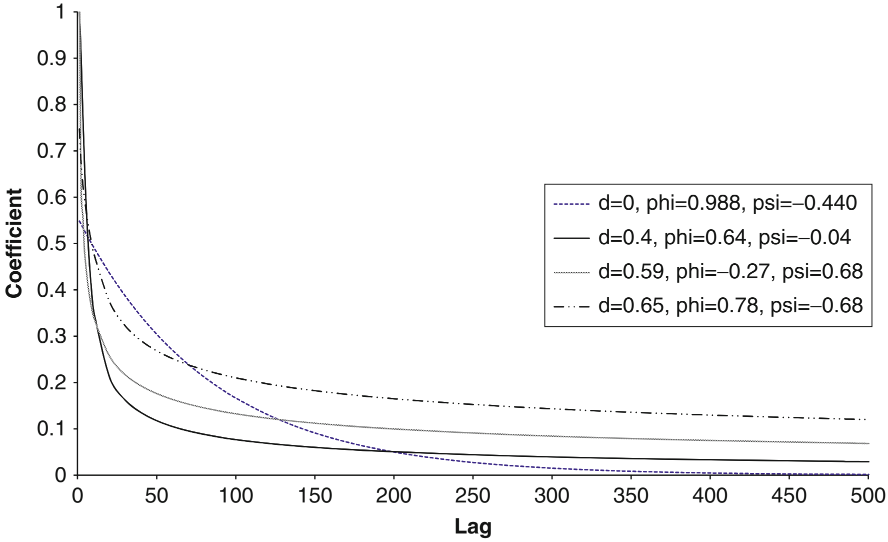


Fig. 32.1 Moving-average coefficients for four ARFIMA(1, d, 1) processes

The estimates obtained here for ϕ and ψ, namely, -0.27 and 0.68 for the most general specification, are rather different to the 0.78 and -0.68 given by Bollerslev and Mikkelsen (1999, Table 1), although the estimates of d are similar, namely, 0.59 and 0.65. However, the moving-average representations obtained from these sets of parameter estimates are qualitatively similar. This is shown on Fig. 32.1 which compares the moving-average coefficients ψ_j defined by (32.36). The coefficients are positive and monotonic decreasing for the four sets of parameter values used to produce Fig. 32.1. They show the expected hyperbolic decay when $d > 0$ and a geometric decay when $d = 0$. The values of b_j in Eqs. 32.32 and 32.38 that are used to calculate the conditional variances decay much faster. For each curve on Fig. 32.1, $\psi_{10} > 0.33$ and $\psi_{100} > 0.07$ while $0 < b_{10} < 0.02$ and $0 < b_{100} < 0.0003$.

The results reported in Table 32.1 are for a constant conditional mean, $\mu_t = \mu$. Alternative specifications such as $\mu_t = \mu + \beta r_{t-1}$, $\mu_t = \mu - \frac{1}{2}h_t$, and $\mu_t = \mu + \lambda\sqrt{h_t}$ give similar values of the log-likelihood when the volatility parameters are set to the values in the final row of Table 32.1. First, including the lagged return r_{t-1} is not necessary because the first-lag autocorrelation of the S & P 100 returns equals -0.022 and is statistically insignificant. Second, including the adjustment $-\frac{1}{2}h_t$ makes the conditional expectation of $(p_t - p_{t-1})/p_{t-1}$ constant when the conditional distribution is normal. The adjustment reduces the log-likelihood by an unimportant 0.3. Third, incorporating the ARCH-M parameter λ gives an optimal value of 0.10 and an increase in the log-likelihood of 1.5. This increase is not significant using a non-robust likelihood-ratio test at the 5 % level.

32.3.3 Stochastic Volatility Specifications

Two shocks per unit time characterize stochastic volatility (SV) models, in contrast to the single shock z_t that appears in ARCH models. A general framework for long memory stochastic volatility models is given for returns r_t by

$$r_t = \mu + \sigma_t u_t \tag{32.19}$$

with $\ln(\sigma_t)$ following an ARFIMA(p, d, q) process. For example, with $p = q = 1$,

$$\ln(\sigma_t) = \alpha + (1 - \phi L)^{-1}(1 - L)^{-d}(1 + \psi L)v_t. \tag{32.20}$$

This framework has been investigated by Breidt et al. (1998), Harvey (1998), and Bollerslev and Wright (2000), all of whom provide results for the simplifying assumption that the two i.i.d. processes $\{u_t\}$ and $\{v_t\}$ are independent. This assumption can be relaxed and has been for short memory applications (Taylor 1994; Shephard 1996).

Parameter estimation is difficult for SV models, compared with ARCH models, because SV models have twice as many random innovations as observable variables. Breidt et al. (1998) describe a spectral-likelihood estimator and provide results for a CRSP index from 1962 to 1989. For the ARFIMA(1, d, 0) specification, they estimate $d = 0.44$ and $\phi = 0.93$. Bollerslev and Wright (2000) provide detailed simulation evidence about semiparametric estimates of d, related to the frequency of the observations.

It is apparent that the ARCH specification (Eqs. 32.15–32.17) has a similar structure to the SV specification (Eqs. 32.19–32.20). Short memory special cases of these specifications, given by $d = q = 0$, have similar moments (Taylor 1994). This is a consequence of the special cases having the same bivariate diffusion limit when appropriate parameter values are defined for increasingly frequent observations (Nelson 1990; Duan 1997). It seems reasonable to conjecture that the multivariate distributions for returns defined by (32.15–17) and (32.19–20) are similar, with the special case of independent shocks $\{u_t\}$ and $\{v_t\}$ corresponding to the symmetric ARCH model that has $\theta = 0$ in Eq. 32.17.

32.4 Option Pricing Methodology

32.4.1 A Review of SV and ARCH Methods

The pricing of options when volatility is stochastic and has a short memory has been studied using a variety of methods. The most popular methods commence with separate diffusion specifications for the asset price and its volatility. These are called stochastic volatility (SV) methods. Option prices then depend on several parameters including a volatility risk premium and the correlation between the differentials of the Wiener processes in the separate diffusions.

The closed-form solution of Heston (1993) assumes that volatility follows a square-root process and permits a general correlation and a nonzero volatility risk premium; for applications see, for example, Bakshi et al. (1997) and for extensions see Duffie et al. (2000).

There has been less research into option pricing for short memory ARCH models. Duan (1995) provides a valuation framework and explicit results for the GARCH(1,1) process that can be extended to other ARCH specifications. Ritchken and Trevor (1999) provide an efficient lattice algorithm for GARCH (1,1) processes and extensions for which the conditional variance depends on the previous value and the latest return innovation. Recent innovations are provided by Christoffersen et al. (2008, 2010).

Methods for pricing options when volatility has a long memory have been described by Comte and Renault (1998) and Bollerslev and Mikkelsen (1996, 1999). The former authors provide analysis within a bivariate diffusion framework. They replace the usual Wiener process in the volatility equation by fractional Brownian motion. However, their option pricing formula appears to require independence between the Wiener process in the price equation and the volatility process which is not consistent with the empirical evidence for stock returns.

The most practical way to price options with long memory in volatility is probably based upon ARCH models, as demonstrated by Bollerslev and Mikkelsen (1999). We follow the same strategy. From the asymptotic results in Duan (1997), also discussed in Ritchken and Trevor (1999), it is anticipated that insights about options priced from a long memory ARCH model will be similar to the insights that can be obtained from a related long memory SV model.

32.4.2 The ARCH Pricing Framework

When pricing options it will be assumed that returns are calculated from prices (or index levels) as $r_t = \ln(p_t/p_{t-1})$ and hence exclude dividends. A constant risk-free interest rate and a constant dividend yield will also be assumed and, to simplify the notation and calculations, it will be assumed that interest and dividends are paid once per trading period. Conditional expectations are defined with respect to current and prior price information represented by $I_t = \{p_{t-i}, i \geq 0\}$.

To obtain fair option prices in an ARCH framework, it is necessary to make additional assumptions in order to obtain a risk-neutral measure Q. Duan (1995) and Bollerslev and Mikkelsen (1999) provide sufficient conditions to apply a risk-neutral valuation methodology. For example, it is sufficient that a representative agent has constant relative risk aversion and that returns and aggregate growth rates in consumption have conditional normal distributions. Kallsen and Taqqu (1998) derive the same solution as Duan (1995) without making assumptions about utility functions and consumption. Instead, they assume that intraday prices are determined by geometric Brownian motion with volatility determined once a day from a discrete-time ARCH model.

32.4.3 Implementation

At time t', measured in trading periods, the fair price of a European contingent claim which has value $y_{t'+n}(p_{t'+n})$ at the terminal time $t' + n$ is given by

$$y_{t'} = E^Q\left[e^{-\rho n} y_{t'+n}\left(p_{t'+n}\right) \ |I_{t'}\right] \tag{32.21}$$

with ρ the risk-free interest rate for one trading period. We now specify an appropriate way to simulate $p_{t'+n}$ under a risk-neutral measure Q. Monte Carlo methods are then used to estimate the conditional expectation in Eq. 32.21.

Following Duan (1995), it is assumed that the distribution of observed returns is defined by some probability measure P, for which

$$r_t|I_{t-1} \sim^P N(\mu_t, h_t), \tag{32.22}$$

with

$$z_t = \frac{r_t - \mu_t}{\sqrt{h_t}} \sim^P i.i.d.N(0,1) \tag{32.23}$$

It is also assumed that the distributions in a risk-neutral framework are defined by a measure Q, with

$$r_t|I_{t-1} \sim^Q N\left(\rho - \delta - \frac{1}{2}h_t, h_t\right), \tag{32.24}$$

and

$$z_t^* = \frac{r_t - \left(\rho - \delta - \frac{1}{2}h_t\right)}{\sqrt{h_t}} \sim^Q i.i.d.N(0,1). \tag{32.25}$$

Here δ is the dividend yield, which corresponds to a dividend payment of $d_t = (e^{\delta} - 1)p_t$ per share at time t. Then $E^Q[p_t|I_{t-1}] = e^{\rho-\delta}p_{t-1}$ and the expected value at time t of one share and the dividend payment is $E^Q[p_t + d_t|I_{t-1}] = e^{\rho}p_{t-1}$, as required in a risk-neutral framework. Note that the conditional means are different for measures P and Q, but the functions $h_t(p_{t-1}, p_{t-2}, \ldots.)$ that define the conditional variances for the two measures are identical.

Option prices depend on the specifications for μ_t and h_t. We again follow Duan (1995) and assume that

$$\mu_t = \rho - \delta - \frac{1}{2}h_t + \lambda\sqrt{h_t} \tag{32.26}$$

with λ representing a risk premium parameter. Then the conditional expectations of r_t for measures P and Q differ by $\lambda\sqrt{h_t}$ and

$$z_t - z_t^* = -\lambda. \tag{32.27}$$

Option prices are evaluated when the conditional variances are given by an ARMA(N, 1) approximation to the FIEGARCH(1, d, 1) specification. From Eqs. 32.18 and 32.27,

$$\left(1-\sum_{j=1}^{N} b_j L^j\right)(\ln(h_t)-\alpha) = (1+\psi L)g(z_{t-1}) = (1+\psi L)g\left(z_{t-1}^{*}-\lambda\right), \quad (32.28)$$

and

$$g(z_t) = \theta z_t + \gamma(|z_t| - C) \quad (32.29)$$

The autoregressive coefficients b_j are functions of ϕ and d, which are defined in Appendix 1. Efficient numerical methods for simulating $p_{t'+n}$ are described in Appendix 2.

32.5 Illustrative Long Memory Option Prices

32.5.1 Inputs

Many parameters and additional inputs are required to implement the FIEGARCH option pricing methodology. To apply that methodology to value European options, we specify 18 numbers, a price history, and a random number generator, as follows:

- Contractual parameters – time until exercise T measured in years, the exercise price X, and whether a call or a put option.
- The current asset price $S = p_{t'}$ and a set of previous prices $\{p_t, 1 \le t < t'\}$.
- Trading periods per annum M, such that consecutive observed prices are separated by $1/M$ years and likewise for simulated prices $\{p_t, t' < t \le t' + n\}$ with $n = MT$.
- Risk-free annual interest rate R, from which the trading period rate $\rho = R/M$ is obtained.
- Annual dividend yield D giving a constant trading period payout rate of $\delta = D/M$; both R and D are continuously compounded and applicable for the life of the option contract.
- The risk premium λ for investment in the asset during the life of the option, such that one-period conditional expected returns are $\mu_t = \rho - \delta - \frac{1}{2}h_t + \lambda\sqrt{h_t}$, for the real-world measure P.
- Parameters m and λ' that define conditional expected returns during the time period of the observed prices by $\mu_t = m - \frac{1}{2}h_t + \lambda'\sqrt{h_t}$, again for measure P.
- Eight parameters that define the one-period conditional variances h_t. The integration level d, the autoregressive parameter ϕ, and the truncation level N determine the terms b_j in the AR(N) filter in Eq. 32.28. The mean α and the moving-average parameter ψ complete the ARMA(N, 1) specification for $\ln(h_t)$ In Eq. 32.28. The values of the shocks to the ARMA(N, 1) process depend on γ

and θ, which, respectively, appear in the symmetric function $\gamma(|z_t| - C)$ and the asymmetric function θz_t whose total determines the shock term $g(z_t)$; the constant C is a parameter C' for observed prices but is $\sqrt{2/\pi}$ when returns are simulated.

- K, the number of independent simulations of the terminal asset price $S_T = p_{t'+n}$.
- A set of Kn pseudorandom numbers distributed uniformly between 0 and 1, from which pseudorandom standard normal variates can be obtained. These numbers typically depend on a seed value and a deterministic algorithm.

32.5.2 Parameter Selections

Option values are tabulated for hypothetical European options on the S & P 100 index. Options are valued for 10 dates defined by the last trading days of the 10 years from 1989 to 1998 inclusive. For valuation dates from 1992 onwards, the size of the price history is set at $t' = 2{,}000$; for previous years the price history commences on 6 March 1984 and $t' < 2{,}000$. It is assumed that there are $M = 252$ trading days in 1 year and hence exactly 21 trading days in one simulated month. Option values are tabulated when T is 1, 2, 3, 6, 12, 18, and 24 months.

Table 32.2 lists the parameter values used to obtain the main results. The annualized risk-free rate and dividend yield are set at 5 % and 2 %, respectively. The risk parameter λ is set at 0.028 to give[2] an annual equity risk premium of 6 %. The mean return parameter m is set to the historic mean of the complete set of S & P 100 returns from March 1984 to December 1998 and λ' is set to zero.

There are two sets of values for the conditional variance process because the primary objective here is to compare option values when volatility is assumed to have either a short or a long memory. The long memory parameter set takes the integration level to be $d = 0.4$, because this is an appropriate level based upon the high frequency reviewed in Sect. 32.2.3. The remaining variance parameters are then based on Table 32.1; as the moving-average parameter is small, it is set to zero and the autoregressive parameter is adjusted to retain the unit total, $d + \phi + \psi = 1$. The AR filter[3] is truncated at lag 1,000, although the results obtained will nevertheless be referred to as long memory results. The short memory parameters are similar to those for the AR(1) estimates provided in Table 32.1. The parameters γ and θ are both 6 % less in Table 32.2 than in Table 32.1 to ensure that selected moments are matched for the short and long memory specifications; the unconditional mean and variance of $\ln(h_t)$ are then matched for the historic measure P, although the unconditional means differ by approximately 0.10 for the risk-neutral measure Q as noted in Appendix 3.

[2]The conditional expectations of r_t for measures P and Q differ by $\lambda\sqrt{h_t}$ and a typical average value of $\sqrt{h_t}$ is 0.00858. Assuming 253 trading days in 1 year gives the stated value of λ.

[3]The filter coefficients sum to $b_1 + \ldots + b_{1,000} = 0.983$. After $b_1 = 1$ and $b_2 = -0.12$, all of the coefficients are near zero, with $b_{100} = 0.00017$ and $b_{1,000} = 7 \times 10^{-6}$.

Table 32.2 Parameter values for option price calculations

Trading periods per annum M	252	
Risk-free interest rate ρ	$0.05/M$	
Conditional mean		
Historic intercept m	$0.161/M$	
Historic equity risk premium term λ'	0	
Dividend yield δ	$0.02/M$	
Future equity risk premium term λ	0.028	
Conditional variance		
	Short memory	*Long memory*
Integration level d	0	0.4
Truncation limit N		1,000
Mean α of $E^P[\ln(h_t)]$	−9.56	−9.56
Autoregressive parameter ϕ	0.982	0.6
Moving-average parameter ψ	0	0
Asymmetric shock parameter θ	−0.056	−0.11
Symmetric shock parameter γ	0.094	0.18
Historical value of $E^P[\|z\|]$	0.737	0.737

Options are valued on the final trading day of ten consecutive years, from 1989 to 1998. The returns history is drawn from the set of daily returns from the S & P 100 index from 6 March 1984 to 31 December 1998. A set of $t' = 2{,}000$ historical returns is used from 1992 onwards, and as many as possible before then. The current level of the index is reset to $S = 100$ when option values are determined

Option prices are estimated from $K = 10,000$ independent simulations of prices $\{p_t, t' < t \leq t' + n\}$ with $n = 504$. Applying the antithetic and control variate methods described in Appendix 2 then produces results for a long memory process in about 12 min, when the processor speed is 2 GHz. Most of the time is spent evaluating the high-order AR filter; the computation time is about 1 min for the short memory process.

32.5.3 Comparisons of Implied Volatility Term Structures

The values of all options are reported using annualized implied volatilities rather than prices. Each implied volatility (IV) is calculated from the Black-Scholes formula, adjusted for continuous dividends. The complete set of IV outputs for one set of inputs forms a matrix with rows labelled by the exercise prices X and columns labelled by the times to expiry T; examples are given in Tables 32.5 and 32.6 and are discussed later.

Initially we only consider at-the-money options, for which the exercise price equals the forward price $F = Se^{(R-D)T}$, with IV values obtained by interpolation across two adjacent values of X. As T varies, the IV values represent the term structure of implied volatility. Tables 32.3 and 32.4, respectively, summarize these term structures for the short and long memory specifications. The same information is plotted on Figs. 32.2

Table 32.3 At-the-money implied volatilities for a short memory volatility process

Implied volatilities for options that expire after 1, 2, 3, 6, 12, 18, and 24 months							
	0	1	2	3	6	12	18	24
Year								
1989	0.1211	0.1238	0.1267	0.1291	0.1340	0.1384	0.1404	0.1414
1990	0.1378	0.1380	0.1388	0.1395	0.1409	0.1421	0.1427	0.1431
1991	0.1213	0.1240	0.1271	0.1296	0.1339	0.1383	0.1403	0.1413
1992	0.1085	0.1129	0.1175	0.1210	0.1288	0.1356	0.1385	0.1401
1993	0.0945	0.1008	0.1069	0.1119	0.1226	0.1321	0.1363	0.1382
1994	0.1165	0.1201	0.1236	0.1263	0.1321	0.1375	0.1398	0.1411
1995	0.1118	0.1158	0.1200	0.1234	0.1304	0.1367	0.1391	0.1405
1996	0.1462	0.1449	0.1446	0.1445	0.1442	0.1438	0.1440	0.1441
1997	0.1883	0.1791	0.1730	0.1683	0.1599	0.1526	0.1500	0.1486
1998	0.1694	0.1640	0.1607	0.1581	0.1531	0.1488	0.1476	0.1469
Mean	0.1315	0.1323	0.1339	0.1352	0.1380	0.1406	0.1419	0.1425
St. dev.	0.0291	0.0243	0.0205	0.0175	0.0116	0.0063	0.0043	0.0032

The parameters of the EGARCH price process are listed in Table 32.2. The half-life of a volatility shock is 1.8 months

European option prices are estimated from 10,000 simulations of the asset prices on the dates that the options expire. The implied volatilities are for at-the-money options whose exercise prices equal the forward rates for the expiry dates. The standard errors of the implied volatilities are between 0.0001 and 0.0003

Options are valued on the final trading day of ten consecutive years. The returns history is drawn from the set of daily returns from the S & P 100 index from 6 March 1984–31 December 1998. A set of $t' = 2{,}000$ historical returns is used from 1992 onwards and as many as possible before then. The column for $T = 0$ provides the annualized volatilities on the valuation dates

and 32.3, respectively. The IV values for $T = 0$ are obtained from the conditional variances on the valuation dates. The standard errors of the tabulated implied volatilities increase with T. The maximum standard errors for at-the-money options are, respectively, 0.0003 and 0.0004 for the short and long memory specifications.

The ten IV term structures for the short memory specification commence between 9.5 % (1993) and 18.8 % (1997) and converge towards the limiting value of 14.3 %. The initial IV values are near the median level from 1989 to 1991, are low from 1992 to 1995, and are high from 1996 to 1998. Six of the term structures slope upwards, two are almost flat, and two slope downwards. The shapes of these term structures are completely determined by the initial IV values because the volatility process is Markovian.

There are three clear differences between the term structures for the short and long memory specifications that can be seen by comparing Figs. 32.2 and 32.3. First, the long memory term structures can and do intersect because the volatility process is not Markovian. Second, some of the term structures have sharp kinks for the first month. This is particularly noteworthy for 1990 and 1996 when the term structures are not monotonic. For 1990, the initial value of 14.1 % is followed by 15.6 % at 1 month and a gradual rise to 16.2 % at 6 months and a subsequent slow decline. For 1996, the term structure commences at 15.6 %,

Table 32.4 At-the-money implied volatilities for a long memory volatility process

Implied volatilities for options that expire after 1, 2, 3, 6, 12, 18, and 24 months								
	0	1	2	3	6	12	18	24
Year								
1989	0.1194	0.1356	0.1403	0.1429	0.1467	0.1496	0.1507	0.1515
1990	0.1413	0.1556	0.1592	0.1609	0.1624	0.1624	0.1614	0.1607
1991	0.1215	0.1368	0.1416	0.1441	0.1478	0.1502	0.1510	0.1516
1992	0.1239	0.1261	0.1283	0.1301	0.1338	0.1375	0.1395	0.1409
1993	0.1080	0.1101	0.1128	0.1150	0.1195	0.1245	0.1272	0.1292
1994	0.1114	0.1189	0.1213	0.1228	0.1256	0.1284	0.1300	0.1314
1995	0.0965	0.1041	0.1067	0.1085	0.1119	0.1160	0.1187	0.1210
1996	0.1564	0.1357	0.1311	0.1295	0.1283	0.1288	0.1297	0.1308
1997	0.1697	0.1650	0.1626	0.1609	0.1574	0.1540	0.1523	0.1515
1998	0.1734	0.1693	0.1682	0.1672	0.1650	0.1623	0.1610	0.1602
Mean	0.1322	0.1357	0.1372	0.1382	0.1399	0.1414	0.1422	0.1429
St. dev.	0.0267	0.0221	0.0211	0.0204	0.0186	0.0165	0.0151	0.0141

The parameters of the FIEGARCH price process are listed in Table 32.2 and include an integration level of $d = 0.4$.

European option prices are estimated from 10,000 simulations of the asset prices on the dates that the options expire. The implied volatilities are for at-the-money options whose exercise prices equal the forward rates for the expiry dates. The standard errors of the implied volatilities are all less than 0.0004

Options are valued on the final trading day of ten consecutive years. The returns history is drawn from the set of daily returns from the S & P 100 index from 6 March 1984–31 December 1998. A set of $t' = 2{,}000$ historical returns is used from 1992 onwards, and as many as possible before then. The column for $T = 0$ provides the annualized volatilities on the valuation dates

falls to 13.6 % after 1 month, and reaches a minimum of 12.8 % after 6 months followed by a slow incline. The eight other term structures are monotonic and only those for 1997 and 1998 slope downwards. Third, the term structures approach their limiting value very slowly.[4] The 2-year IVs range from 12.1 % to 16.1 %, and it is not possible to deduce the limiting value, although 15.0–16.0 % is a plausible range.[5] It is notable that the dispersion between the ten IV values for each T decreases slowly as T increases, from 2.2 % for 1-month options to 1.4 % for 2-year options.

There are substantial differences between the two IV values that are calculated for each valuation date and each option lifetime. Figure 32.4 shows the differences between the at-the-money IVs for the long memory

[4]The results support the conjecture that $IV(T) \cong a_1 + a_2 T^{2d-1}$ for large T with a_2 determined by the history of observed returns.

[5]An estimate of the constant a_1 (defined in the previous footnote) is 16.0 %. An estimate of 15.0 % follows by supposing the long memory limit is 105 % of the short memory limit, based on the limit of $\ln(h_t)$ being higher by 0.1 for the long memory process as noted in Appendix 3. The difference in the limits is a consequence of the risk premium obtained by owning the asset; its magnitude is mainly determined by the pronounced asymmetry in the volatility shock function $g(z_t)$.

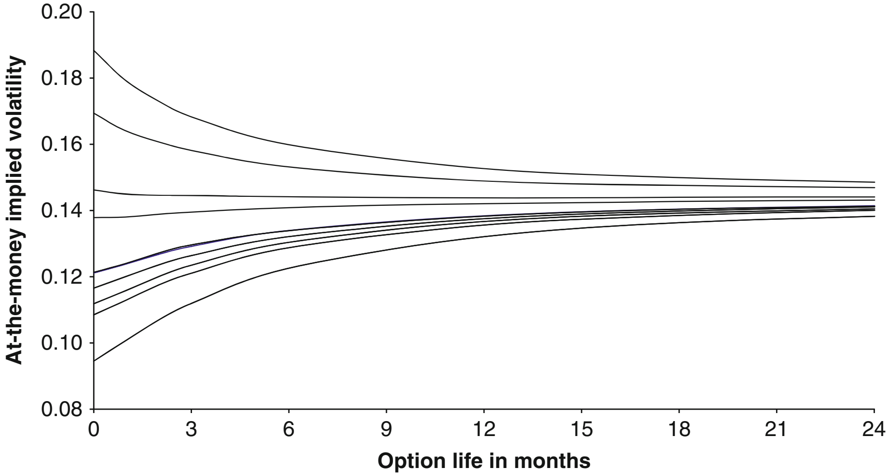


Fig. 32.2 Ten volatility term structures for a short memory process

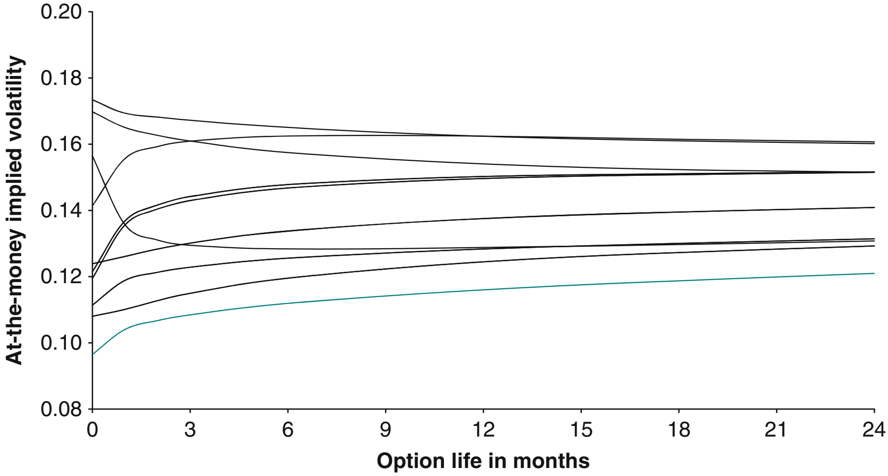


Fig. 32.3 Ten volatility term structures for a long memory process with d = 0.4

specification minus the number for the short memory specification. When $T = 0$, these differences range from -1.9 % (1997) to 1.5 % (1992), for 3 month options from -1.5 % (1995, 1996) to 2.1 % (1990), and for 2-year options from -1.9 % (1995) to 1.7 % (1990). The standard deviation of the ten differences is between 1.1 % and 1.4 % for all values of T considered so it is common for the short and long memory option prices to have IVs that differ by more than 1 %.

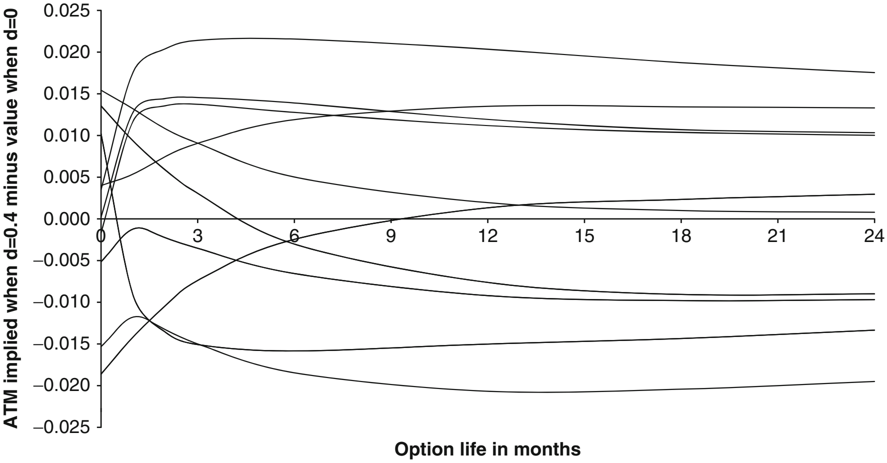


Fig. 32.4 Differences between ten pairs of term structures

32.5.4 Comparisons of Smile Effects

The columns of the IV matrix provide information about the strength of the so-called smile effect for options prices. These effects seem to be remarkably robust to the choice of valuation date and they are not very sensitive to the choice between the short and long memory specifications. This can be seen by considering the ten values of $\Delta IV = IV(T,X_1) - IV(T,X_2)$ obtained for the ten valuation dates, for various values of T, various pairs of exercise prices X_1, X_2, and a choice of volatility process. First, for 1-month options with $S = 100$, $X_1 = 92$, and $X_2 = 108$, the values of ΔIV range from 3.0 % to 3.3 % for the short memory specification and from 3.7 % to 4.0 % for the long memory specification. Second, for 2-year options with $X_1 = 80$ and $X_2 = 120$, the values of ΔIV range from 1.8 % to 2.0 % and from 1.8 % to 1.9 %, respectively, for the short and long memory specifications.

Figure 32.5 shows the smiles for 3-month options valued using the short memory model, separately for the ten valuation dates. As may be expected from the above remarks, the ten curves are approximately parallel to each other. They are almost all monotonic decreasing for the range of exercise prices considered, so that a U-shaped function (from which the idea of a smile is derived) cannot be seen. The near monotonic decline is a standard theoretical result when volatility shocks are negatively correlated with price shocks (Hull 2000). It is also a stylized empirical fact for US equity index options; see, for example, Rubinstein (1994) and Dumas et al. (1998).

Figure 32.6 shows the 3-month smiles for the long memory specification. The shapes on Figs. 32.5 and 32.6 are similar, as all the curves are for the same expiry time, but they are more dispersed on Fig. 32.6 because the long memory effect induces more dispersion in at-the-money IVs. The minima of the smiles are generally near an exercise price of 116. Figure 32.7 shows further long memory smiles, for 2-year options when the forward price is 106.2. The parallel shapes are

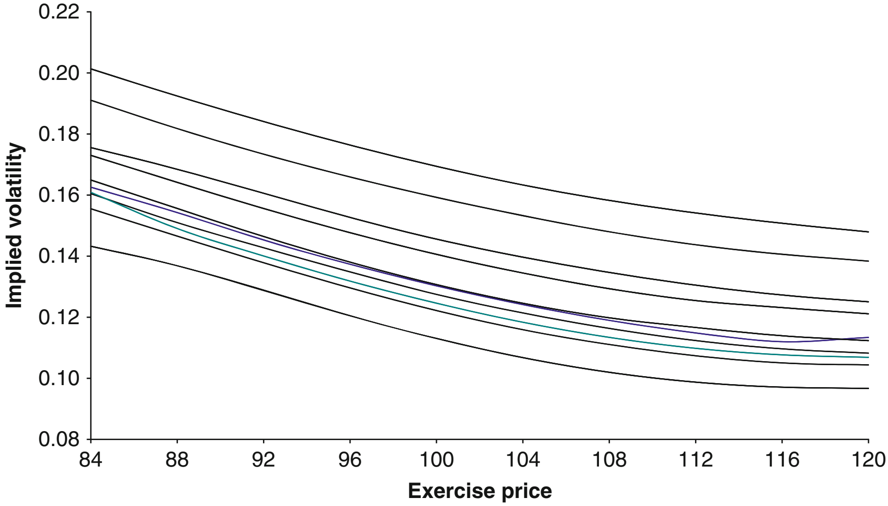


Fig. 32.5 Ten smile shapes for three-month options and a short memory process

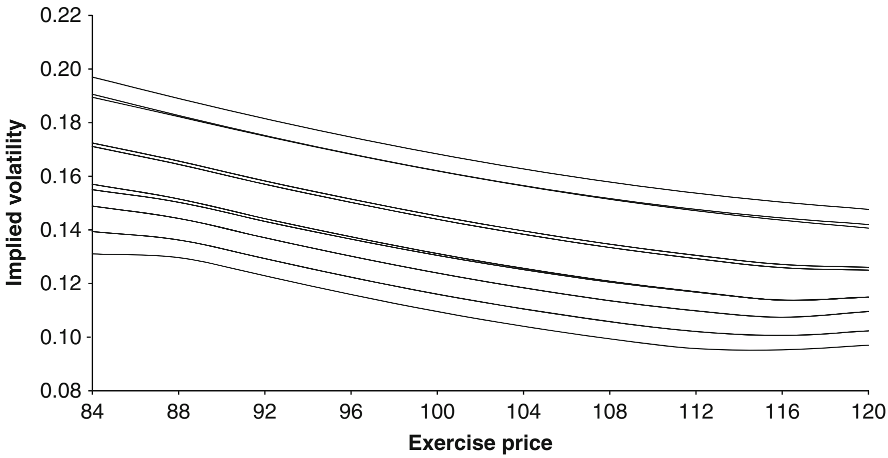


Fig. 32.6 Ten smile shapes for three-month options and a long memory process with d = 0.4

clear; the two highest curves are almost identical, and the third, fourth, and fifth highest curves are almost the same.

Tables 32.5 and 32.6 provide matrices of implied volatilities for options valued on 31 December 1998. When either the call or the put option is deep out-of-the-money, it is difficult to estimate the option price accurately because the risk-neutral probability $q(X)$ of the out-of-the-money option expiring in-the-money is small. Consequently, the IV information has not been presented when the corresponding standard errors exceed 0.002; estimates of $q(X)$ are less than 3 %.

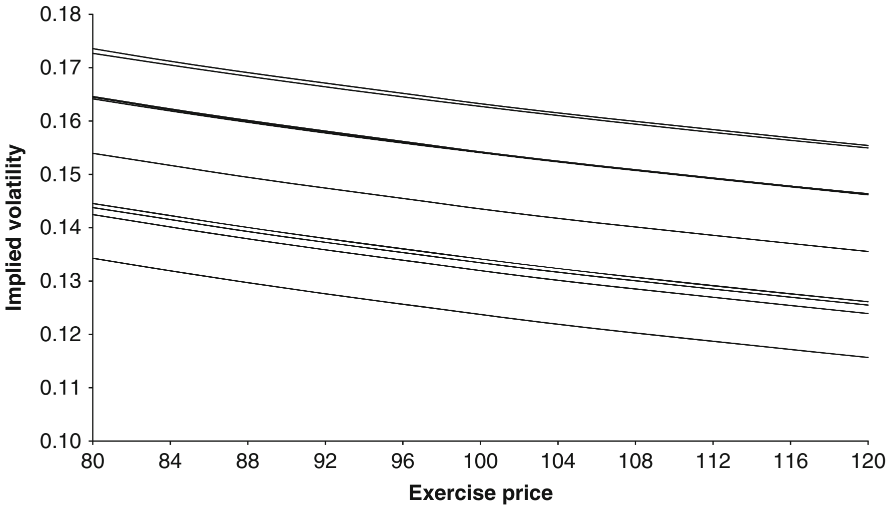


Fig. 32.7 Ten smile shapes for two-year options and a long memory process with d = 0.4

The standard errors of the IVs are least for options that are near to at-the-money and most of them are less than 0.0005 for the IVs listed in Tables 32.5 and 32.6. All the sections of the smiles summarized by Tables 32.5 and 32.6 are monotonic decreasing functions of the exercise price. The IV decreases by approximately 4–5 % for each tabulated section.

32.5.5 Sensitivity Analysis

The sensitivity of the IV matrices to three of the inputs has been assessed for options valued on 31 December 1998. First, consider a change to the risk parameter λ which corresponds to an annual risk premium of 6 % for the tabulated results. From Sect. 32.4.3, option prices should be lower for large T when λ is reduced to zero. Changing λ to zero reduces the at-the-money IV for 2-year options from 16.0 % to 15.4 % for the long memory inputs, with a similar reduction for the short memory inputs. Second, consider reducing the truncation level N in the AR(N) filter from 1,000 to 100. Although this has the advantage of a substantial reduction in the computational time, it changes the IV numbers by appreciable amounts and cannot be recommended; for example, the 2-year at-the-money IV then changes from 16.0 % to 14.7 %.

The smile shapes on Figs. 32.5, 32.6 and 32.7 are heavily influenced by the negative asymmetric shock parameter θ, which is substantial relative to the symmetric shock parameter γ. The asymmetry in the smile shapes can be expected to disappear when θ is zero, which is realistic for some assets including exchange rates. Figures 32.8 and 32.9 compare smile shapes when θ is changed from the

Table 32.5 A matrix of implied volatilities for a short memory volatility process

Implied volatilities for options that expire after 1, 2, 3, 6, 12, 18, and 24 months							
	1	2	3	6	12	18	24
X							
72						*0.1705*	*0.1651*
76					*0.1744*	*0.1670*	*0.1624*
80				*0.1835*	*0.1696*	*0.1637*	*0.1600*
84		*0.1994*	*0.1911*	*0.1771*	*0.1655*	*0.1607*	*0.1577*
88	*0.1959*	*0.1876*	0.1818	0.1711	0.1617	0.1579	0.1556
92	0.1845	0.1780	0.1734	0.1653	0.1579	0.1552	0.1535
96	0.1735	0.1693	0.1660	0.1599	0.1546	0.1526	0.1515
100	0.1644	0.1615	0.1593	0.1549	0.1514	0.1502	0.1496
104	0.1577	0.1550	0.1533	0.1502	0.1481	0.1479	0.1478
108	0.1537	0.1498	0.1480	0.1461	0.1452	0.1457	0.1461
112	*0.1510*	0.1461	0.1437	0.1422	0.1425	0.1436	0.1445
116	*0.1525*	*0.1435*	0.1407	0.1388	0.1400	0.1417	0.1429
120		*0.1421*	*0.1384*	0.1359	0.1376	0.1399	0.1415
126				*0.1325*	0.1345	0.1373	0.1394
132				*0.1304*	0.1317	0.1348	0.1374
138					*0.1293*	0.1328	0.1353
144					*0.1272*	0.1305	0.1336
150					*0.1255*	*0.1286*	0.1322
156						*0.1269*	0.1306
162						*0.1246*	*0.1294*
168						*0.1220*	*0.1282*
174							*0.1277*
180							*0.1274*

The parameters of the EGARCH price process are listed in Table 32.2. The half-life of a volatility shock is 1.8 months

Options are valued on 31 December 1998. The returns history is the set of 2,000 daily returns from the S & P 100 index from 4 February 1991 to 31 December 1998. European option prices are estimated from 10,000 simulations of the asset prices on the dates that the options expire

Implied volatilities shown in Roman font have standard errors (s.e.) that are at most 0.0005 and those shown in italic font have s.e. between 0.0005 and 0.0020; results are not shown when the s.e. exceeds 0.0020, when options are either deep in- or out-of-the-money

values used previously to zero, with γ scaled to ensure the variance of $\ln(h_t)$ is unchanged for measure P. Figure 32.8 shows that the 1-month smile shapes become U-shaped when θ is zero, while Fig. 32.9 shows that the IV are then almost constant for 1-year options.

32.6 Conclusions

The empirical evidence for long memory in volatility is strong, for both equity and foreign exchange markets. This evidence may more precisely be interpreted as

Table 32.6 A matrix of implied volatilities for a long memory volatility process

Implied volatilities for options that expire after 1, 2, 3, 6, 12, 18, and 24 months							
	1	2	3	6	12	18	24
X							
72						*0.1811*	*0.1765*
76					*0.1829*	*0.1783*	*0.1743*
80				*0.1895*	*0.1792*	*0.1754*	*0.1723*
84			*0.1970*	*0.1847*	*0.1760*	*0.1728*	*0.1703*
88		*0.1954*	*0.1889*	*0.1798*	0.1729	0.1703	0.1683
92	*0.1953*	0.1860	0.1815	0.1750	0.1699	0.1680	0.1664
96	0.1814	0.1771	0.1746	0.1707	0.1672	0.1657	0.1645
100	0.1699	0.1691	0.1685	0.1666	0.1645	0.1635	0.1627
104	0.1611	0.1623	0.1631	0.1627	0.1617	0.1615	0.1610
108	0.1557	0.1569	0.1581	0.1593	0.1592	0.1596	0.1595
112	*0.1525*	0.1526	0.1538	0.1559	0.1570	0.1577	0.1579
116		*0.1499*	0.1505	0.1529	0.1549	0.1559	0.1564
120		*0.1480*	*0.1480*	0.1504	0.1528	0.1543	0.1550
126			*0.1439*	0.1469	0.1502	0.1520	0.1530
132				*0.1440*	0.1478	0.1497	0.1512
138				*0.1421*	0.1456	0.1477	0.1492
144					*0.1436*	0.1456	0.1473
150					*0.1417*	0.1438	0.1457
156					*0.1402*	*0.1421*	0.1444
162						*0.1409*	0.1428
168						*0.1388*	*0.1416*
174						*0.1370*	*0.1405*
180							*0.1399*

The parameters of the FIEGARCH price process are listed in Table 32.2 and include an integration level of $d = 0.4$

Options are valued on 31 December 1998. The returns history is the set of 2,000 daily returns from the S & P 100 index from 4 February 1991 to 31 December 1998. European option prices are estimated from 10,000 simulations of the asset prices on the dates that the options expire

Implied volatilities shown in Roman font have standard errors (s.e.) that are at most 0.0005, and those shown in italic font have s.e. between 0.0005 and 0.0020; results are not shown when the s.e. exceeds 0.0020, when options are either deep in- or out-of-the-money

evidence for long memory effects, because there are short memory processes that have similar autocorrelations and spectral densities, except at very low frequencies.

The theory of option pricing when volatility follows a discrete-time ARCH process relies on weak assumptions about the continuous-time process followed by prices and the numerical implementation of the theory is straightforward. Application of the theory when the volatility process is fractionally integrated does, however, require pragmatic approximations because the fundamental filter $(1 - L)^d$ is an infinite order polynomial that must be truncated at some power N. Option prices are sensitive to the truncation point N, so that large values and long price histories from an assumed stationary process are required.

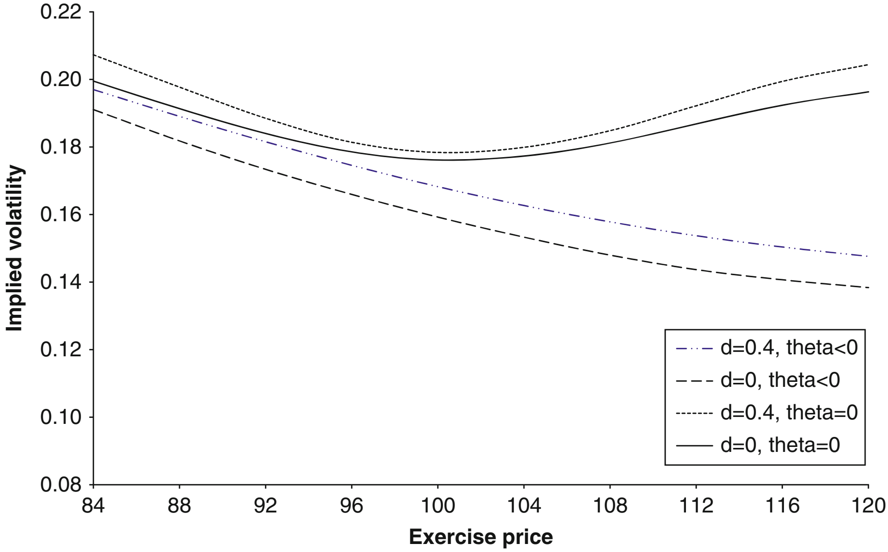


Fig. 32.8 Impact of asymmetric volatility shocks on one-month options

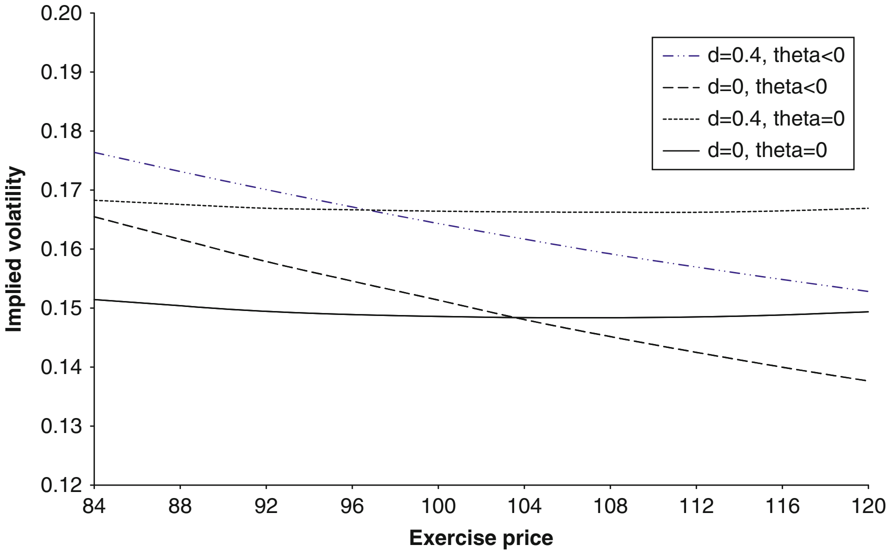


Fig. 32.9 Impact of asymmetric volatility shocks on one-year options

The term structure of implied volatility for at-the-money options can be notably different for short and long memory ARCH specifications applied to the same price history. Long memory term structures have more variety in their shapes. They may have kinks for short maturity options and they may not have a monotonic shape. Also, term structures calculated on different valuation dates sometimes intersect each other. None of these possibilities occurs for a Markovian short memory specification. Long memory term structures do not converge rapidly to a limit as the lifetime of options increases. It is difficult to estimate the limit for the typical value $d = 0.4$.

Implied volatilities as functions of exercise prices have similar shapes for short and long memory specifications. The differences in these shapes are minor in comparison to the differences in the term structure shapes.

It is common for the short and long memory implied volatilities to differ by more than 1 % for options on the S & P 100 index, regardless of the option lifetime and the exercise price; if the short memory implied is at a typical level of 14 %, then the long memory implied is often below 13 % or above 15 %. Consequently, the economic consequences of a long memory assumption are important.

Acknowledgments The support of Inquire Europe and the helpful advice of Bernard Dumas, Ton Vorst, and Granville Tunnicliffe-Wilson are gratefully acknowledged.

Appendix 1: Series Expansions and Truncation Approximations

Series expansions in the lag operator L are required to evaluate the following conditional variance:

$$\ln(h_t) = \alpha + (1 - \phi L)^{-1}(1 - L)^{-d}(1 + \psi L)g(z_{t-1}). \tag{32.30}$$

We note the results:

$$(1 - L)^d = 1 - \sum_{j=1}^{\infty} a_j L^j, a_1 = d, a_j = \frac{j - d - 1}{j} a_{j-1}, \ j \geq 2, \tag{32.31}$$

$$(1 - \phi L)(1 - L)^d = 1 - \sum_{j=1}^{\infty} b_j L^j, b_1 = d + \phi, b_j = a_j - \phi a_{j-1}, \ \ j \geq 2, \tag{32.32}$$

$$(1 - \phi L)(1 - L)^d (1 + \psi L)^{-1} = 1 - \sum_{j=1}^{\infty} \phi_j L^j,$$

$$\phi_1 = d + \phi + \psi, \phi_j = b_j - (-\psi)^j + \sum_{k=1}^{j-1} (-\psi)^{j-k} b_k, \ \ j \geq 2. \tag{32.33}$$

The autoregressive weights in Eqs. 32.33 can be denoted as $\phi_j\ (d,\phi,\psi)$. Also,

$$(1-\phi L)^{-1}(1-L)^{-d}(1+\psi L)=1+\sum_{j=1}^{\infty}\psi_j L^j, \tag{32.34}$$

$$\psi_1=d+\phi+\psi,\ \psi_j=-\phi_j(-d,-\psi,-\phi) \tag{32.35}$$

It is necessary to truncate the infinite summations when evaluating empirical conditional variances. Truncation after N terms of the summations in Eqs. 32.35, 32.33, and 32.32, respectively, gives the MA(N), AR(N), and ARMA(N, 1) approximations:

$$\ln(h_t)=\alpha+g(z_{t-1})+\sum_{j=1}^{N}\psi_j g(z_{t-j-1}), \tag{32.36}$$

$$\ln(h_t)=\alpha+\sum_{j=1}^{N}\phi_j\left[\ln(h_{t-j})-\alpha\right]+g(z_{t-1}), \tag{32.37}$$

$$\ln(h_t)=\alpha+\sum_{j=1}^{N}b_j\left[\ln(h_{t-j})-\alpha\right]+g(z_{t-1})+\psi g(z_{t-2}). \tag{32.38}$$

As $j\to\infty$, the coefficients b_j and ϕ_j converge much more rapidly to zero than the coefficients ψ_j. Consequently it is best to use either the AR or the ARMA approximation.

Appendix 2: Simulation Methods

Suppose there are returns observed at times $1\le t\le t'$, whose distributions are given by measure P, and that we then want to simulate returns for times $t>t'$ using measure Q. Then ln (h_t) is calculated for $1\le t\le t'+1$ using the observed returns, with ln $(h_t)=\alpha$ and $g(z_t)=0$ for $t<1$, followed by simulating $z_t^*\sim^Q N(0,1)$ and hence obtaining r_t and ln (h_{t+1}) for $t>t'$.

Care is required when calculating the constant C in Eq. 32.29 because the observed conditional distributions of the terms z_t are not normal while the simulations assume that they are. Consequently, we define

$$\begin{aligned} C &= C', & t&\le t',\\ &=\sqrt{2/\pi}, & t&>t', \end{aligned} \tag{32.39}$$

for a constant C' estimated from observed returns. An alternative method, described by Bollerslev and Mikkelsen (1999), is to simulate from the sample distribution of standardized observed returns.

Standard variance reduction techniques substantially enhance the accuracy of Monte Carlo estimates of contingent claim prices. Our antithetic method uses one i.i.d. $N(0, 1)$ sequence $\{z_t^*\}$ to define the further i.i.d. $N(0, 1)$ sequences, $\{-z_t^*\}$, $\{z_t^\bullet\}$, and $\{-z_t^\bullet\}$, with the terms $z_t^\bullet$ chosen so that there is negative correlation between $|z_t^*|$ and $|z_t^\bullet|$; this is achieved by defining $\Phi(z_t^\bullet) + \Phi(z_t^*) = 1 + \frac{1}{2}\text{sign}(z_t^*)$. The four sequences provide claim prices whose average, $\bar{y}$ say, is much less variable than the claim price from a single sequence. An overall average $\hat{y}$ is then obtained from a set of K values $\{\bar{y}_k, 1 \le k \le K\}$.

The control variate method makes use of an unbiased estimate $\hat{y}_{CV}$ of a known parameter y_{CV}, such that $\hat{y}$ is positively correlated with $\hat{y}_{CV}$. A suitable parameter, when pricing a call option in an ARCH framework, is the price of a call option when volatility is deterministic. The deterministic volatility process is defined by replacing all terms $\ln(h_t)$, $t > t' + 1$, by their expectations under P conditional on the history $I_{t'}$. Then y_{CV} is given by the obvious modification of the Black-Scholes formula, while $\hat{y}_{CV}$ is obtained by using the same $4K$ sequences of i.i.d. variables that define $\hat{y}$. Finally, a more accurate estimate of the option price is then given by $\tilde{y} = \hat{y} - \beta(\hat{y}_{CV} - y_{CV})$ with β chosen to minimize the variance of $\tilde{y}$.

Appendix 3: Impact of a Volatility Risk Premium

On average the term structure of implied volatilities will slope upwards for the FIEGARCH option pricing model. This occurs because the expectation of $\ln(h_t)$ depends on the measure[6] when $\lambda \neq 0$. The unconditional expectation equals α for measure P. It is different for measure Q because

$$E^Q\left[g\left(z_t^* - \lambda\right)\right] = -\lambda\theta + \gamma\left(E^Q\left[\left|z_t^* - \lambda\right|\right] - \sqrt{2/\pi}\right) \cong -\lambda\theta + \frac{\lambda^2\gamma}{\sqrt{2\pi}} \tag{32.40}$$

when λ is small, and this expectation is in general not zero.[7] For a fixed t', as $t \to \infty$,

$$E^Q[\ln(h_t)\;|I_{t'}] \to \alpha + \left(1 - \sum_{j=1}^{N} b_j\right)^{-1}(1+\psi)E^Q\left[g\left(z_t^* - \lambda\right)\right] \tag{32.41}$$

The difference between the P and Q expectations of $\ln(h_t)$ could be interpreted as a volatility risk premium. This premium is typically negative, because typically $\lambda > 0$, $\theta \le 0$ and $\gamma > 0$. Furthermore, when θ is negative, the dominant term in Eq. 32.40 is $-\lambda\theta$, because λ is always small, and then the premium reflects the degree of asymmetry in the volatility shocks $g(z_t)$.

[6]The dependence of moments of h_t on the measure is shown by Duan (1995, p. 19) for the GARCH (1,1) model.

[7]When $z \sim N(0,1)$, $E[|z - \lambda|] = \sqrt{2/\pi}\exp\left(-\frac{1}{2}\lambda^2\right) + \lambda(2\Phi(\lambda) - 1)$ with Φ the cumulative distribution function of z.

The magnitude of the volatility risk premium can be important and, indeed, the quantity defined by the limit in Eq. 32.41 becomes infinite[8] as $N \to \infty$ when d is positive. A plausible value of λ for the S & P 100 index is 0.028, obtained by assuming that the equity risk premium is 6 % per annum. For the short memory parameter values in the first row of Table 32.1, when $d = 0$ and $N = 1{,}000$, the limit of $E^Q[\ln(h_t) \,|I_t'] - \alpha$ equals 0.10. This limit increases to 0.20 for the parameter values in the final row of Table 32.1, when $d = 0.4$ and $N = 1{,}000$. The typical effect of adding 0.2 to $\ln(h_t)$ is to multiply standard deviations $\sqrt{h_t}$ by 1.1 so that far-horizon expected volatilities, under Q, are slightly higher than might be expected from historical standard deviations.

References

Andersen, T. G., & Bollerslev, T. (1997). Heterogeneous information arrivals and return volatility dynamics: Uncovering the long-run in high frequency returns. *Journal of Finance, 52*, 975–1005.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Ebens, H. (2001a). The distribution of stock return volatility. *Journal of Financial Economics, 61*, 43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2001b). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association, 96*, 42–55.

Areal, N. M. P. C., & Taylor, S. J. (2002). The realized volatility of FTSE-100 futures prices. *Journal of Futures Markets, 22*, 627–648.

Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics, 73*, 5–59.

Bakshi, G., Cao, C., & Chen, Z. (1997). Empirical performance of alternative option pricing models. *Journal of Finance, 52*, 2003–2049.

Barndorff-Nielsen, O. E., & Shephard, N. (2001). Non-Gaussian Ornstein-Uhlenbeck based models and some of their uses in financial economics. *Journal of the Royal Statistical Society B, 63*, 167–241.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics, 31*, 307–327.

Bollerslev, T., & Mikkelsen, H. O. (1996). Modeling and pricing long memory in stock market volatility. *Journal of Econometrics, 73*, 151–184.

Bollerslev, T., & Mikkelsen, H. O. (1999). Long-term equity anticipation securities and stock market volatility dynamics. *Journal of Econometrics, 92*, 75–99.

Bollerslev, T., & Wooldridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time varying covariances. *Econometric Reviews, 11*, 143–172.

Bollerslev, T., & Wright, J. H. (2000). Semiparametric estimation of long-memory volatility dependencies: The role of high-frequency data. *Journal of Econometrics, 98*, 81–106.

Breidt, F. J., Crato, N., & de Lima, P. (1998). The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics, 83*, 325–348.

Brockwell, P. J., & Davis, R. A. (1991). *Time series: Theory and methods* (2nd ed.). New York: Springer.

Christoffersen, P., Jacobs, K., Ornthanalai, C., & Wang, Y. (2008). Option valuation with long-run and short-run volatility components. *Journal of Financial Economics, 90*, 272–297.

Christoffersen, P., Dorion, C., Jacobs, K., & Wang, Y. (2010). Volatility components: Affine restrictions and non-normal innovations. *Journal of Business and Economic Statistics, 28*, 483–502.

Comte, F., & Renault, E. (1998). Long memory in continuous-time stochastic volatility models. *Mathematical Finance, 8*, 291–323.

[8]As $(1 - L)^d 1 = 0$ for $d > 0$, it follows from Eqs. 32.31 and 32.32 that $\sum_{j=1}^{\infty} a_j = \sum_{j=1}^{\infty} b_j = 1$.

Dacorogna, M. M., Muller, U. A., Nagler, R. J., Olsen, R. B., & Pictet, O. V. (1993). A geographical model for the daily and weekly seasonal volatility in the foreign exchange market. *Journal of International Money and Finance, 12*, 413–438.

Ding, Z., Granger, C. W. J., & Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance, 1*, 83–106.

Duan, J.-C. (1995). The GARCH option pricing model. *Mathematical Finance, 5*, 13–32.

Duan, J.-C. (1997). Augmented GARCH(p, q) process and its diffusion limit. *Journal of Econometrics, 79*, 97–127.

Duffie, D., Pan, J., & Singleton, K. J. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica, 68*, 1343–1376.

Dumas, B., Fleming, J., & Whaley, R. E. (1998). Implied volatility functions: Empirical tests. *Journal of Finance, 53*, 2059–2106.

Ebens, H. (1999). Realized stock volatility. Working paper, John Hopkins University.

Ederington, L. H., & Lee, J. H. (1993). How markets process information: News releases and volatility. *Journal of Finance, 49*, 1161–1191.

Gallant, A. R., Hsu, C. T., & Tauchen, G. E. (1999). Using daily range data to calibrate volatility diffusions and extract the forward integrated variance. *Review of Economics and Statistics, 81*, 617–631.

Granger, C. W. J. (1980). Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics, 14*, 227–238.

Granger, C. W. J., & Ding, Z. (1996). Varieties of long memory models. *Journal of Econometrics, 73*, 61–77.

Granger, C. W. J., & Joyeux, R. (1980). An introduction to long-memory time series models and fractional differencing. *Journal of Time Series Analysis, 1*, 15–29.

Harvey, A. C. (1998). Long-memory in stochastic volatility. In J. Knight & S. Satchell (Eds.), *Forecasting volatility in financial markets*. London: Butterworth-Heinemann.

Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies, 6*, 327–343.

Hosking, J. (1981). Fractional differencing. *Biometrika, 68*, 165–176.

Hull, J. (2000). *Options, futures and other derivative securities* (4th ed.). New Jersey: Prentice-Hall.

Kallsen, J., & Taqqu, M. (1998). Option pricing in ARCH-type models. *Mathematical Finance, 8*, 13–26.

Nelson, D. B. (1990). ARCH models as diffusion approximations. *Journal of Econometrics, 45*, 7–38.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica, 59*, 347–370.

Ohanissian, A., Russell, J. R., Tsay, R. S. (2004). True or spurious long memory in volatility: Does it matter for pricing options? Working paper, University of Chicago.

Ohanissian, A., Russell, J. R., & Tsay, R. S. (2008). True or spurious long memory? A new test. *Journal of Business and Economic Statistics, 26*, 161–175.

Pong, S., Shackleton, M. B., & Taylor, S. J. (2008). Distinguishing short and long memory volatility specifications. *Econometrics Journal, 11*, 617–637.

Ritchken, P., & Trevor, R. (1999). Pricing options under generalized GARCH and stochastic volatility processes. *Journal of Finance, 54*, 377–402.

Rubinstein, M. (1994). Implied binomial trees. *Journal of Finance, 49*, 771–818.

Shephard, N. (1996). Statistical aspects of ARCH and stochastic volatility. In D. R. Cox, D. V. Hinkley, & O. E. Barndorff-Nielsen (Eds.), *Likelihood, time series with econometric and other applications* (pp. 1–67). London: Chapman and Hall.

Taylor, S. J. (1994). Modeling stochastic volatility: A review and comparative study. *Mathematical Finance, 4*, 183–204.

Taylor, S. J. (2005). *Asset price dynamics, volatility, and prediction*. Princeton: Princeton University Press.

Taylor, S. J. (2008). *Modelling financial time series* (2nd ed.). Singapore: World Scientific Publishing.

Seasonal Aspects of Australian Electricity Market

33

Vikash Ramiah, Stuart Thomas, Richard Heaney, and Heather Mitchell

Contents

Abstract

Australian electricity spot prices differ considerably from equity spot prices in that they contain an extremely rapid mean-reversion process. The electricity spot price could increase to a market cap price of AU$12,500 per megawatt hour (MWh) and revert back to a mean level (AUD$30) within a half-hour interval. This has implications for derivative pricing and risk management. For example, while the Black and Scholes option pricing model works reasonably well for equity market-based securities, it performs poorly for commodities like electricity. Understanding the dynamics of electricity spot prices and demand is also

V. Ramiah (✉)
School of Economics, Finance and Marketing, RMIT University, Melbourne, Australia
e-mail: vikash.ramiah@rmit.edu.au

S. Thomas • H. Mitchell
RMIT University, Melbourne, VIC, Australia
e-mail: stuart.thomas@rmit.edu.au; heather.mitchell@rmit.edu.au

R. Heaney
Accounting and Finance, The University of Western Australia, Perth, Australia
e-mail: richard.heaney@uwa.edu.au

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_33,

important in order to correctly forecast electricity prices. We develop econometric models for seasonal patterns in both price returns and proportional changes in demand for electricity. We also model extreme spikes in the data. Our study identifies both seasonality effects and dramatic price reversals in the Australian electricity market. The pricing seasonality effects include time-of-day, day-of-week, monthly, and yearly effects. There is also evidence of seasonality in demand for electricity.

Keywords

Electricity • Spot price • Seasonality • Outlier • Demand • Econometric modelling

33.1 Introduction

On Monday 11th of January 2010, the temperature in the state of Victoria in Australia peaked to 42.8 °C. This led to an increase in demand for electricity as households turned on their air conditioning systems. During that half-hour interval, the wholesale price of electricity surged to over $9,000 and then reverted to the average price of around $30 in the subsequent half-hour period. Exercising a call option around that time would have been extremely profitable. What causes the spot price of electricity to exhibit such erratic behavior? We provide a partial explanation of this behavior in this study.

Unpredictable jumps in electricity spot price are usually referred to as “spikes” because jumps tend to happen very quickly, with equally quick reversion to mean levels. These events create a challenge for quantitative analysts when it comes to pricing electricity derivatives. Before one can price these derivatives, one has to be able to effectively forecast the spot price distribution. To date, this problem remains an unresolved matter. For instance, the Black and Scholes (1973) option pricing model fails to adequately price an option written on the electricity spot price because of the incidence of these pricing spikes. The Black and Scholes model works reasonably well in the equity markets whereby the equity time series takes a longer time to trend upwards or downwards. Given the rapid price reversals evident in the electricity spot prices, the Black and Scholes model does not perform particularly well. The key characteristic of non-storability of electricity partially explains why electricity spot price increases significantly during a short-term interval. As the possibility of storing large amounts of electricity does not exist, this product has to be consumed and produced simultaneously. Large price swings occur in this market mainly because supply is inelastic.

The electricity pricing literature provides clear evidence of the spikes. Thomas et al. (2011), Huisman and Huurman (2003), and Goto and Karolyi (2004) show non-normality manifested as positive skewness and extreme leptokurtosis; Johnson and Barz (1999) observe mean reversion in the long run; Kaminski (1997), Clewlow and Strickland (2000), and Eichler et al. (2012) find evidence of extreme behavior with fast-reverting spikes; and Bunn and Karakatsani (2003) demonstrate excessive volatility. Further, Escribano et al. (2002) show that electricity volatility can be time varying in a number of countries like in Argentina, New Zealand, Norway, Sweden, and Spain.[1] Other researchers devote their time to modelling the electricity time series itself. For instance, Knittel and Roberts (2001) apply mean-reversion, time-varying mean, jump-diffusion, time-dependent jump intensity, ARMAX, and exponential generalized autoregressive conditional heteroskedasticity (EGARCH) models to hourly prices in the Californian electricity market and conclude that forecasting performance is relatively poor. Kaminski (1997) addresses the spiky characteristic with a random walk jump-diffusion model, but this model ignores the persistent mean reversion, which is a feature of electricity prices, and this avenue is explored further in Clewlow and Strickland (2000). One of the limitations of the jump-diffusion approaches is the assumption that all shocks affecting the price series die out at the same rate. Escribano et al. (2002) identify two additional price components, namely, volatility clustering in the form of GARCH effects and seasonality (emphasized by Lucia and Schwartz 2002), both in the deterministic component of prices and the jump intensity.

The seasonal aspects recognized within the electricity market are also important. The results observed in one geographic region cannot be applied to other areas as the climatic conditions are different. For example, summer time occurs in different months around the globe, and for that reason, it is important to investigate the seasonal aspects within the Australian electricity market.

Seasonality in the electricity market is not limited to price series as seasonal patterns in demand or system load are well documented in the literature. Harvey and Koopman (1993) document intra-daily and intra-week effects and incorporate them into their demand model using splines. Other early studies considered longer-term load forecasting horizons several months into the future, using daily, weekly, or monthly demand data (e.g., Engle et al. 1989). Pardo et al. (2002) employ daily data in a study of Spanish electricity demand and emphasize the importance of daily and monthly seasonal structures. More recent studies consider modelling and forecasting demand over shorter periods using intraday data. In the Australian context, Smith (2000) documents intraday patterns in demand in the New South Wales electricity market and incorporates diurnal variation into a Bayesian semiparametric regression framework to model intraday electricity load data and obtain short-term load forecasts. We extend his analysis by testing whether seasonality exists in other Australian regions.

In the early 1990s, following the release of the Hilmer Report, the Australian electricity industry embarked on a progressive program of deregulation. The Hilmer reforms led to the disaggregation of vertically integrated, government-owned electricity authorities into separate generation transmission and distribution and retail sales sectors in each state. This gave rise to a wholesale market for electricity which is currently managed by the Australian Energy Market Operator (AEMO). AEMO is responsible for five regions, namely, VIC1 (Victoria), NSW1 (New South Wales), QLD1 (Queensland), SA1 (South Australia), and TAS1 (Tasmania).

[1]See Montero et al. (2011) for further evidence in the Spanish market.

Western Australia and the Northern Territory are geographically remote and to date have not been integrated into the "national" market. Physical transmission of power between regions is achieved via interconnectors that physically link the five different states. The spot electricity market in the AEMO is where all market generators and market customers settle their electricity sales and purchases based on a spot price. The market participants are generally large consumers of electricity, generators, and speculators. Large consumers of electricity like car manufacturers and supermarkets enter this market to enjoy a lower cost of production and to hedge their positions. Speculators, on the other hand, trade future contracts, forwards, options, caps, floors, and swaps for profit motives.

33.2 Australian Electricity Spot Price

The half-hourly pool price data is sourced directly from AEMO (previously known as National Electricity Market Management Company, NEMMCO) for the period from 1 January 1999 to 31 January 2006. Prices are expressed in Australian dollars per megawatt hour ($/MWh). The sample size is 124,224 observations for each region except for TAS1 that joined the pool at a later stage. Figure 33.1 shows the data series for NSW1 with evidence of considerable spiked behavior. Descriptive statistics for the price and return series which are shown in Table 33.1 tend to confirm these outliers.

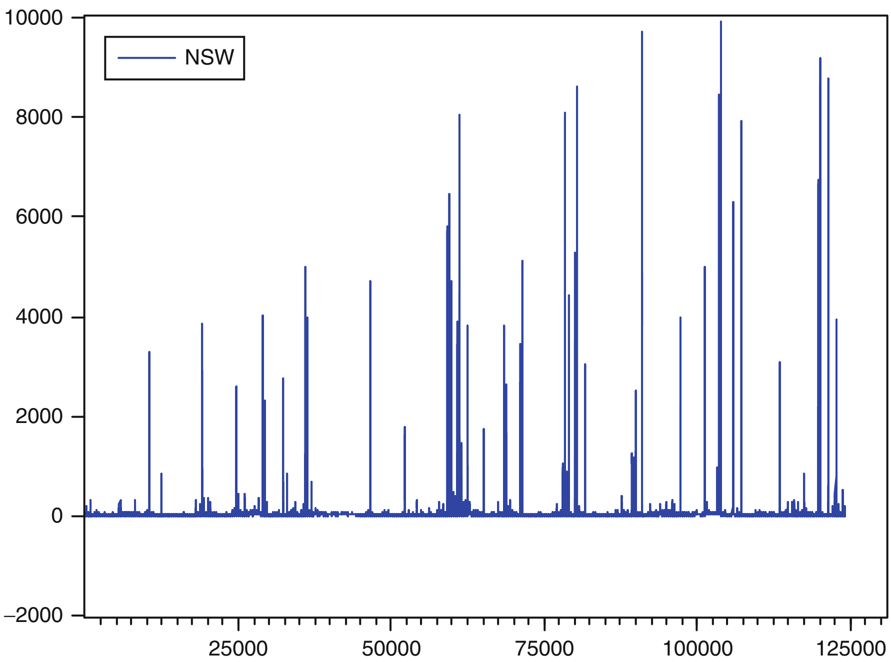


Fig. 33.1 Electricity price series for NSW1 ($/MWh) for the period 01.01.1999–31.01.2006

Table 33.1 The descriptive statistics of spot prices and returns for the different Australian regions

	Range	Minimum	Maximum	Mean	Std. deviation	Skewness	Kurtosis	JB stats
Spot prices								
NSW1	9,912.13	−3.10	9,909.03	34.02	178.11	35.37	1468.54	89701
QLD1	9,098.74	−156.14	8,942.60	36.65	158.45	28.56	1058.17	46527
SA1	9,822.43	−822.45	8,999.98	41.91	155.18	25.40	800.35	26598
VIC1	7,746.07	−329.91	7,416.16	30.19	103.26	35.38	1502.93	93950
TAS1	10,332.00	−332.00	10,000.00	87.30	367.26	21.46	505.16	10584
Returns								
RNSW1	453.00	−222.00	231.00	0.0301	1.36	41.46	15,480.38	9981513
RQLD1	380.49	−11.25	369.24	0.0611	1.74	111.49	19,122.19	15233050
RSA1	423.74	−32.94	390.80	0.0603	1.68	123.00	24,820.63	25665638
RVIC1	351.93	−209.50	142.43	0.0262	1.05	−15.41	17,959.48	13434834
RTAS1	2416.00	−83.50	2,332.50	0.6681	31.42	59.14	3,754.10	586865

We report the range, minimum, maximum, mean, standard deviation, skewness, kurtosis, and Jarque-Bera statistics (JB stats) for each region's price and return series. Mean prices vary between regions from \$30.19 for VIC1 to \$41.91 for SA1 and to \$87.30 for TAS1 as the factor inputs (low-cost coal, high-cost gas and water) vary from one state to another. TAS1 was at the early stage of joining the pool, and these statistics depict and are consistent with the immature stage of the market in this region. The standard deviation of prices is generally high, is widely dispersed across the regions, and is broadly consistent with the pattern of means, ranging from \$103.26 for VIC1 to \$367.26 for TAS1. The lowest maximum price of \$7416.16 is observed in VIC1 and the highest maximum price in TAS1 at \$10,000. No prices above \$10,000 were observed as this was the market price cap during the period of the study. This amount is smaller than other international markets like the state of California in the United States, which recorded over US\$50,000 during a period when no market cap was in force. These extremely high values are the features that we are looking for and are referred to as spikes in this study. These instances are generally viewed as lucrative occurrences by arbitrageurs. The most commonly traded call options on the Australian electricity market have exercise prices of either \$100 or \$300. If we assume that market participants will exercise their rights when the option is in the money, then VIC1 traders will exercise when the spot price is above \$340 (mean price of around \$30 plus three standard deviations away) on average. To ensure that we capture a spike, we adopt a conservative approach whereby we define a spike as the mean plus four standard deviations. We observe a total of 505 spikes in NSW1, QLD1, SA1, and VIC1. QLD1 exhibits the greatest incidence of extreme price spikes by state with 173 occurrences (34 %), followed by SA1 with 159 (31 %) and VIC1 with 96 (19 %). By day of the week, Monday shows the highest incidence with 115 (23 %) tapering gradually to Sunday with 45 occurrences (9 %). June shows the highest incidence by month with 81 (16 %). The highest incidences by year occur in 2002 with 147 spikes (29 %) and 2000 with 116 spikes (23 %), both markedly higher than any other full year in the study period. It should be noted that the incidence of extreme price spikes appears to be declining from 2003 onwards as legislation on collusive behavior was passed in Queensland. There is evidence of a concentration of spikes occurring between the hours 06:30 and approximately 10:00 and between 15:30 and 19:00 h, with a marked increase in frequency concentrated at the 18:00 trading interval.

Another unique characteristic of this dataset is negative price. All five regions exhibit negative minimum prices. As shown in Table 33.1, the negative prices for NSW1, QLD1, SA1, VIC1, and TAS1 are −\$3.10, −\$156.14, −\$822.45, −\$329.91, and −\$332.00, respectively. To understand the possibility of such rare and short-lived occurrence, it is important to understand how the spot price is derived. It is a derived price per trading interval, calculated by a two-step procedure based on the offers to supply made by generators in the pool. The trading day is divided into 48 half-hour "trading intervals" which is then subdivided into 5-min. "dispatch intervals." A "dispatch price" is recorded as the marginal price of supply to meet demand for each 5-min. interval in a given half-hour period. This marginal price is typically the dispatch offer price of the last generator brought into production to meet demand at that interval. The spot price is then calculated as an arithmetic average of the six dispatch prices in a half hour. All generators who are called into production during a given half-hour trading interval receive this spot price for the quantity of electricity delivered during the trading interval. A generator may bid a negative price into the pool for its self-dispatch quantity as a tactical move to ensure that they are among the first to be called in to generate as closing down their operations due to not being called into production may be more costly than producing at negative prices. The implication of a negative price is that the standard return calculation for prices may contain this bias, and to that end, the following return formula is used:

$$RET_t = \frac{(P_t - P_{t-1})}{|P_{t-1}|}. \tag{33.1}$$

where RET_t represents the half-hourly discrete proportionate change in price ("return") at time t, P_t is half-hourly price at time t, and $|P_{t-1}|$ is the absolute value of the previous half-hourly price, i.e., at time $t-1$. The denominator is specified as the absolute value to allow for the presence of negative prices. We prefer a discrete return specification over log returns because the spot market in the AEMO trades at discrete half-hourly intervals – it is not a continuous market in the way of most conventional financial markets. Further, a log return specification will dampen the extreme spike effects we are attempting to capture and is incompatible with negative prices. Descriptive statistics for the half-hourly return series are shown in Table 33.1. Mean half-hourly returns vary widely between regions, from 2.6 % for VIC1 to 6.1 % for QLD1. High maximum returns are observed and are consistent with the spike behavior discussed above.

One of the clear patterns discussed so far is the spike behavior, and, based on the earlier discussion, it is fair to say that these spikes are more likely to occur at some specific periods like 18.00 h, Mondays, and in the year 2002. We develop a method to test if there is seasonality in the return series with half-hourly return as the dependent variable. The explanatory variables consist of seasonal dummy variables, spike behavior, and negative prices. We add lagged returns to the list of independent variables to control for serial correlation in the return series. The model employed in this study is presented as Eq. 33.2:

$$\begin{aligned} RET_{R,t} = a_0 &+ \sum_{i=1}^{5}\beta_{1,i}RET_{R,t-1} + \sum_{j-1}^{6}\beta_{2,j}DAY_j + \sum_{k=1,\neq 9}^{12}\beta_{3,k}MTH_k + \sum_{l=1999,\neq 2001}^{2006}\beta_{4,l}YR_l \\ &+ \sum_{m=1,\neq 24}^{48}\beta_{5,m}HH_m + \sum_{o=1}^{N_{R,S}}\beta_{6,o}SPIKE_{R,o} + \sum_{p=1}^{N_{R,N}}\beta_{7,p}NEG_{R,p} + \varepsilon_t \end{aligned} \tag{33.2}$$

where $RET_{R,t}$ represents the discrete return for region R at time t, α_0 represents the constant term, and DAY_j represents the dummy variable for each day of the week

(j = 1 for Monday, 2 for Tuesday, . . . , 6 for Saturday). The dummy variable for Sunday is dropped on the basis that the returns are relatively lower on that day. We use similar justification when it comes to discarding other periods. MTH_k represents the dummy variable for each month (k = 1 for January, 2 for February, . . . , 12 for December). The dummy variable for September is dropped; YR_l represents the dummy variable for each year included in the sample period (l = 1999, . . . , 2006). The dummy variable for 2001 is dropped; HH_m represents the dummy variable for each half-hourly trading interval (m = 1 for 00:00 h, 2 for 00:30 h, . . . , 48 for 23:30 h). The dummy variable for 11:30 h is dropped; $SPIKE_{R,o}$ represents a set of $N_{R,S}$ dummy variables, one for each extreme spike as previously defined, with $N_{R,S}$ representing the number of extreme returns observed in region R for the period of the study; $NEG_{R,N}$ represents the dummy variable for the return associated with an occurrence of a negative price ($p = 1, \ldots, N_{R,N}$), with $N_{R,N}$ representing the number of occurrences of a negative price for region R during for the period of the study.

The values for the trading interval at 11:30 h, Sunday, September, and the year 2001 were dropped to avoid exact collinearity and to allow comparison of these values with the remaining seasonal coefficients for HH_m, DAY_j, MTH_k, and YR_l. The equation was initially estimated for each region with 20 lagged returns ($RET_{Rt-1}, \ldots, RET_{Rt-20}$). F-tests for redundant variables were performed for all regions, and AIC and SBC values support the finding that lags 1 through 5 were significant. Lags 6 onwards were not found to be significant and were discarded. Standard tests and residual diagnostics revealed no misspecification in the above model.

Results of the regression analysis are presented in Table 33.2. Coefficients and t-statistics (t-stats) are presented for each seasonal dummy variable. The results generated from this model are consistent with Kaminski (1997), Clewlow and Strickland (2000) and De Jong and Huismann (2002) with regard to the seasonal aspects and spike behavior of the price series. Seasonal effects vary between regions, and time-of-day effects are generally more significant than other seasonalities. Positive returns are observed at times of peak population activity in the morning and early evening, and negative returns observed at most other times. In general, significant negative returns are found for the small hours of the morning between 12:30 a.m. and approximately 4:00 a.m. in all regions. NSW1 exhibits an unexplained positive return at 1:30 a.m., reverting to negative returns for the remainder of the early morning. The hours between 5:00 a.m. and 9:30 a.m. inclusively exhibit significant positive effect in all regions, reverting to generally negative returns in the late morning. Negative returns are found in SA1 during midafternoon. Significant positive effects are observed for all regions in the early evening, generally between the hours of 5:00 and 7:00 p.m., reverting to significant negative effects for the remainder of the evening, until positive effects emerge in the late evening at 10:30 p.m. and at midnight. The periods of positive return in the morning and early evening are consistent with peaks in activity in the population. The positive returns observed around 11:00 p.m. are consistent with increased demand for electricity arising from off-peak hot water systems generally switching on at 11:00 p.m.

Table 33.2 The seasonality effects in the different regions

	SA1		VIC1		NSW1		QLD1	
Variable	Coeff	t-stats	Coeff	t-stats	Coeff	t-stats	Coeff	t-stats
Mon	0.0113	**3.36**	0.0279	**4.00**	0.0218	**3.13**	0.0154	**4.05**
Tues	0.0105	**3.14**	0.0249	**3.57**	0.0190	**2.72**	0.0137	**3.61**
Wed	0.0120	**3.56**	0.0246	**3.53**	0.0178	**2.56**	0.0118	**3.09**
Thurs	0.0109	**3.25**	0.0232	**3.33**	0.0213	**3.05**	0.0137	**3.59**
Fri	0.0090	**2.67**	0.0233	**3.35**	0.0176	**2.52**	0.0121	**3.18**
Sat	−0.0018	**−0.53**	0.0137	**1.97**	0.0092	**1.33**	−0.0010	**−0.26**
Jan	0.0003	**0.06**	0.0022	**0.25**	0.0015	**0.16**	0.0172	**3.42**
Feb	0.0063	**1.40**	0.0043	**0.46**	0.0017	**0.18**	0.0102	**1.99**
Mar	−0.0068	**−1.54**	−0.0008	**−0.09**	−0.0015	**−0.16**	0.0220	**4.39**
Apr	0.0019	**0.42**	0.0021	**0.23**	0.0012	**0.13**	0.0029	**0.57**
May	0.0039	**0.87**	0.0051	**0.56**	0.0048	**0.52**	0.0116	**2.31**
June	0.0018	**0.41**	0.0047	**0.51**	0.0064	**0.70**	0.0125	**2.47**
July	0.0089	**2.01**	0.0023	**0.25**	0.0043	**0.47**	0.0134	**2.68**
Aug	0.0074	**1.67**	0.0041	**0.45**	0.0035	**0.38**	0.0124	**2.48**
Oct	0.0016	**0.37**	−0.0289	**−3.16**	−0.0235	**−2.56**	0.0062	**1.24**
Nov	0.0048	**1.07**	−0.0059	**−0.64**	−0.0025	**−0.27**	−0.0020	**−0.40**
Dec	−0.0041	**−0.93**	0.0005	**0.06**	−0.0007	**−0.07**	0.0110	**2.20**
Y1999	0.0406	**12.02**	−0.0128	**−1.83**	−0.0072	**−1.03**	0.0227	**5.91**
Y2000	0.0234	**6.93**	0.0043	**0.61**	0.0136	**1.95**	0.0361	**9.43**
Y2002	0.0043	**1.29**	−0.0008	**−0.11**	0.0085	**1.22**	0.0131	**3.43**
Y2003	0.0001	**0.04**	−0.0036	**−0.51**	0.0002	**0.02**	−0.0087	**−2.28**
Y2004	0.0012	**0.37**	−0.0010	**−0.14**	0.0041	**0.58**	−0.0061	**−1.58**
Y2005	0.0025	**0.75**	−0.0037	**−0.54**	0.0029	**0.41**	−0.0082	**−2.14**
Y2006	0.0210	**2.32**	0.0031	**0.16**	−0.0018	**−0.10**	−0.0080	**−0.78**
H0000	−0.0133	**−1.51**	−0.1022	**−5.61**	−0.1002	**−5.49**	−0.1598	**−16.03**
H0030	−0.0383	**−4.36**	−0.0783	**−4.29**	−0.0587	**−3.22**	−0.1123	**−11.27**
H0100	−0.0402	**−4.58**	−0.0819	**−4.49**	−0.0637	**−3.49**	−0.0930	**−9.33**
H0130	−0.0778	**−8.85**	0.1873	**10.27**	0.0875	**4.80**	−0.0335	**−3.36**
H0200	−0.1838	**−20.93**	−0.1379	**−7.57**	−0.1087	**−5.96**	−0.0922	**−9.25**
H0230	−0.1172	**−13.34**	−0.1147	**−6.29**	−0.0814	**−4.46**	−0.0716	**−7.18**
H0300	−0.1528	**−17.41**	−0.0976	**−5.36**	−0.0650	**−3.56**	−0.0577	**−5.78**
H0330	−0.1384	**−15.76**	−0.0969	**−5.32**	−0.0673	**−3.69**	−0.0511	**−5.13**
H0400	−0.1016	**−11.58**	−0.0619	**−3.39**	−0.0391	**−2.14**	−0.0417	**−4.18**
H0430	−0.0310	**−3.54**	0.0066	**0.36**	0.0226	**1.24**	−0.0096	**−0.96**
H0500	−0.0192	**−2.19**	0.0174	**0.95**	0.0279	**1.53**	−0.0051	**−0.52**
H0530	0.1118	**12.73**	0.1614	**8.86**	0.1441	**7.90**	0.0533	**5.35**
H0600	0.0723	**8.23**	0.0870	**4.77**	0.0769	**4.22**	0.0122	**1.22**
H0630	0.1757	**20.00**	0.2156	**11.82**	0.1587	**8.70**	0.0983	**9.86**
H0700	0.1988	**22.63**	0.1196	**6.56**	0.0752	**4.12**	0.1173	**11.77**
H0730	−0.0456	**−5.19**	−0.0587	**−3.22**	−0.0583	**−3.20**	0.0605	**6.07**
H0800	0.1350	**15.37**	0.1741	**9.55**	0.1550	**8.50**	0.1343	**13.46**

(*continued*)

Table 33.2 (continued)

Variable	SA1 Coeff	SA1 t-stats	VIC1 Coeff	VIC1 t-stats	NSW1 Coeff	NSW1 t-stats	QLD1 Coeff	QLD1 t-stats
H0830	0.1454	**16.55**	0.1281	**7.02**	0.1061	**5.82**	0.1089	**10.91**
H0900	0.0286	**3.26**	0.0284	**1.56**	0.0123	**0.67**	0.0371	**3.72**
H0930	0.0160	**1.83**	0.0733	**4.02**	0.0659	**3.61**	0.0588	**5.89**
H1000	−0.0173	**−1.97**	0.0023	**0.13**	−0.0019	**−0.10**	−0.0256	**−2.56**
H1030	−0.0396	**−4.51**	0.0015	**0.08**	−0.0015	**−0.08**	−0.0228	**−2.29**
H1100	−0.0120	**−1.37**	0.0150	**0.82**	0.0142	**0.78**	−0.0213	**−2.13**
H1200	−0.0219	**−2.50**	0.0067	**0.37**	0.0106	**0.58**	−0.0072	**−0.72**
H1230	0.0111	**1.27**	0.0170	**0.93**	0.0232	**1.27**	−0.0227	**−2.27**
H1300	−0.0020	**−0.23**	−0.0030	**−0.17**	0.0033	**0.18**	−0.0254	**−2.55**
H1330	−0.0020	**−0.23**	0.0432	**2.37**	0.0429	**2.35**	0.0153	**1.53**
H1400	−0.0051	**−0.58**	−0.0046	**−0.25**	−0.0022	**−0.12**	−0.0180	**−1.80**
H1430	−0.0402	**−4.57**	−0.0063	**−0.35**	0.0052	**0.28**	0.0005	**0.05**
H1500	−0.0277	**−3.15**	−0.0011	**−0.06**	−0.0049	**−0.27**	−0.0416	**−4.17**
H1530	−0.0373	**−4.25**	−0.0046	**−0.25**	0.0018	**0.10**	−0.0125	**−1.26**
H1600	−0.0105	**−1.20**	0.0160	**0.88**	0.0171	**0.94**	0.0014	**0.14**
H1630	−0.0403	**−4.59**	−0.0063	**−0.34**	−0.0109	**−0.60**	−0.0081	**−0.81**
H1700	−0.0002	**−0.03**	0.0353	**1.94**	0.0290	**1.59**	0.0436	**4.37**
H1730	0.0355	**4.04**	0.0685	**3.75**	0.0853	**4.67**	0.1127	**11.30**
H1800	0.1622	**18.44**	0.2035	**11.13**	0.2390	**13.06**	0.2770	**27.69**
H1830	0.0675	**7.68**	0.0584	**3.20**	0.0574	**3.14**	0.0789	**7.91**
H1900	−0.0408	**−4.65**	−0.0235	**−1.29**	−0.0237	**−1.30**	0.0380	**3.81**
H1930	−0.1043	**−11.88**	−0.0743	**−4.07**	−0.0820	**−4.49**	−0.1058	**−10.60**
H2000	−0.0473	**−5.39**	−0.0344	**−1.89**	−0.0492	**−2.70**	−0.0593	**−5.94**
H2030	−0.0698	**−7.95**	−0.0336	**−1.84**	−0.0382	**−2.10**	−0.1199	**−12.02**
H2100	−0.0858	**−9.77**	−0.0672	**−3.69**	−0.0679	**−3.72**	−0.0876	**−8.79**
H2130	−0.0493	**−5.61**	−0.0253	**−1.39**	−0.0018	**−0.10**	−0.0524	**−5.26**
H2200	−0.0994	**−11.32**	−0.0998	**−5.47**	−0.0901	**−4.94**	−0.0910	**−9.13**
H2230	0.0502	**5.72**	0.1401	**7.69**	0.1923	**10.54**	0.1460	**14.64**
H2300	−0.0635	**−7.23**	−0.0659	**−3.62**	−0.0690	**−3.79**	−0.0764	**−7.66**
H2330	0.1348	**15.35**	0.2989	**16.40**	0.1890	**10.36**	0.1219	**12.22**
C	0.0145	**1.91**	−0.0206	**−1.31**	−0.0215	**−1.37**	−0.0015	**−0.17**
Return(−1)	−0.0071	**−13.23**	0.0006	**0.34**	0.0047	**3.44**	−0.0048	**−8.25**
Return(−2)			−0.0078	**−4.38**	−0.0062	**−4.54**	−0.0042	**−7.11**
Return(−3)			−0.0077	**−4.34**	−0.0059	**−4.31**	0.0003	**0.47**
Return(−4)							−0.0009	**−1.48**
Return(−5)							0.0085	**14.37**
R-squared	0.9646		0.6119		0.7681		0.9577	
Adj R-squared	0.9645		0.6114		0.7678		0.9576	

[a]F-test was carried to exclude redundant lag returns (from lag 6 to lag 20)
[b]Akaike info criterion and Schwarz criterion were used to determine the optimal lags for the different regions

Day-of-week effects generally appeared stronger for Monday, Tuesday, and Wednesday than other days of the week. Monthly effects are not found to be consistent across regions, nor are yearly effects. There appears to be no clear pattern in monthly effect across regions, although a small but significant negative effect is noted for October in NSW1 and VIC1. Positive effects are noted for winter months in SA1, and QLD demonstrates significant positive effects in spring and summer months. As expected, there is considerable evidence of positive spikes in the return series, and the evidence shows that negative prices within the Australian electricity market cannot be ignored. Such results cast doubts on the application of lognormal prices to the Australian electricity market.

33.3 Australian Demand for Electricity

Given the instantaneous market-clearing nature of prices in the AEMO, a logical extension of this study is to investigate the prevalence of seasonal effects and spike behavior in electricity demand, with a view to examining the extent to which these effects are transmitted from demand to price and how efficiently the spot market absorbs demand-side shocks.

33.3.1 Demand Data

The demand data used in this study are half-hourly observations of total demand. The data is obtained from the same source as the price data and covers the same period and regions. The basic quantity of interest in demand modelling and forecasting is typically the periodic "total system demand" or "total demand." The total demand value reported by AEMO is a derived value, somewhat different from demand (as may be represented by traded volume), as it may be understood in conventional financial markets. Suppliers and distributors lodge schedules and bids for the sale and purchase of electricity with AEMO at 12:30 p.m. on the day prior to actual dispatch of electricity for each interval. AEMO compiles this data and mates it with a short-term forecast of system demand and grid capacity to determine an expected dispatch quantity and dispatch order of generators (Smith 2000).

This study uses AEMO's reported "total demand" values for each region, expressed in megawatts (MW) by half-hour trading interval for the sample period. Total demand is defined by AEMO as the total forecast regional demand against which a dispatch solution is performed. For any particular interval and region, this is determined as described by Eq. 33.3:

$$D_T = \sum_{i=1}^{n} G_i - \sum_{i=1}^{n} L_i + NI_i + \sum_{i=1}^{n} AIL + F(D) + ADE \tag{33.3}$$

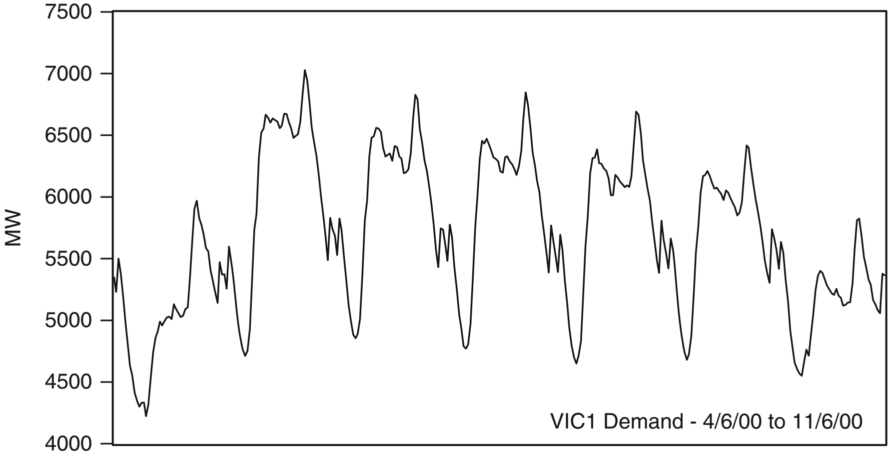


Fig. 33.2 Plot of VIC1 demand for the week commencing Monday 4/6/00, illustrating the regular intraday and daily seasonal patterns in the demand series

where:

D_T is total demand.

G_i is "generator initial MW (SCADA)," the sum of initial MW values for all scheduled generation units within the region, measured at their generator terminals and reported by SCADA.

L_i is "load initial MW (SCADA)," the scheduled base-load generation level for the interval.

NI_i is "net interconnector initial MW into region," the net of all interconnector flows into and out of the region.

AIL is "total allocated interconnector losses" represented by $\sum$(MW losses X regional loss allocation). "MW losses" represent actual power losses due to physical leakage from the transmission system. Regional loss allocation is an NEMMCO predetermined static loss factor for each interconnector.

$F(D)$ is demand forecast, a per-interval demand adjustment that relates the demand at the beginning of the interval to the target at the end of the interval.

ADE is "aggregate dispatch error," an adjustment value used by the NEM to account for disparities between scheduled and actual dispatch for all scheduled generation units in the region.

Figure 33.2 is provided to illustrate the presence of seasonal patterns in the intraday behavior of electricity demand in Victoria over a 10-day period in 2000.

Descriptive statistics for the demand series are shown in Table 33.3.

We report the mean, standard deviation, minimum, maximum, range, skewness, kurtosis, Jarque-Bera statistic, and Augmented Dickey-Fuller[2] statistics for each

[2]Additional statistics are provided for the demand series but not for the price series. This is because the original paper on the price series reports these values.

Table 33.3 Descriptive statistics for demand by region, January 1999 to January 2006

Demand	NSW1	QLD1	SA1	VIC1
Mean[a]	8123.89	5197.74	1451.56	5411.02
S.D.[a]	1306.02	864.33	264.45	757.26
Maximum[a]	12884.15	8231.95	2873.03	8545.39
Minimum[a]	4624.03	2945.96	778.00	2726.88
Skewness	0.03	0.12	0.74	0.04
Kurtosis	2.64	2.67	4.44	2.83
JB stat	700.17	874.81	22026.19	175.72
ADF[b]	−22.83	−17.50	−28.30	−24.77
N	124224	124224	124224	124224

[a]Mean, standard deviation, maximum, and minimum are expressed in megawatts (MW)
[b]Augmented Dickey-Fuller (ADF) statistic rejects the hypothesis of a unit root at the 1 % level of confidence

region's demand series. NSW1 has the highest mean, median, and maximum demand observations of the five regions for the period. New South Wales is Australia's most populous state so we would expect that demand for electric power to be highest in the NSW1 region. The other regions follow generally in order of state population, with VIC1 next highest, followed by QLD1, and SA1. The standard deviation and range of demand levels are generally high, widely dispersed across the regions, and broadly consistent with the pattern of means, ranging from 264 MW for SA1 to 1,306 MW for NSW1. The distributions of demand observations are slightly positively skewed for all regions. NSW1, QLD1, and VIC1 are slightly platykurtic, while SA1 is leptokurtic. Jarque-Bera (JB) statistics reject the hypothesis of normal distribution at the 1 % level of significance for all four regions, and Augmented Dickey-Fuller statistics reject the hypothesis of a unit root at the 1 % level of significance.

33.3.2 Demand Returns

In this section, we consider the proportional changes in demand over a price interval, which for convenience we refer to as "demand returns." The demand return series are of interest because there are a number of over-the-counter and exchange-traded derivative products available for hedgers and speculators in the Australian and overseas electricity markets. Pricing models for derivatives are informed by the behavior of returns on the spot price. The half-hourly pool price and its associated returns exhibit strong seasonal and outlier effects as a result of the occurrence of price spikes. Demand is widely regarded as a major influence on price (and therefore returns), and we are interested in investigating the extent to which the seasonalities observed in half-hourly returns on spot price are present in the equivalent returns on demand. Figure 33.3 shows demand and demand return over a 10-day period and indicates that demand returns appear to exhibit some time-of-day effects but also suggests the presence of sudden and fast-reverting spikes in the demand return series.

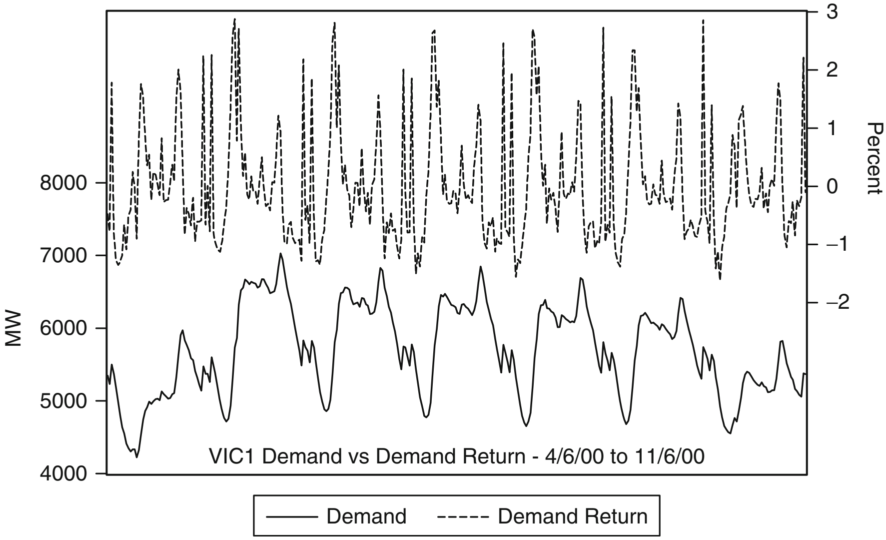


Fig. 33.3 Plot of VIC1 demand and returns on demand for the week commencing 4/6/00. Demand is in MW and returns are percentage returns

In light of the fact that AEMO's total demand is reported at half-hourly intervals in discrete time, the demand return series used in this study were generated as half-hourly discrete returns rather than log returns, according to Eq. 33.4 as follows:

$$RD_t = \frac{(D_t - D_{t-1})}{D_{t-1}} \tag{33.4}$$

where RD_t is discrete demand return at time t, D_t is half-hourly demand at time t, and D_{t-1} is the previous half-hourly total demand, i.e., at time $t-1$. The results of tests for the presence of a unit root give us confidence that the demand and return series are stationary, and we prefer this discrete return specification over log returns, as a log return specification will dampen the spike effects we are attempting to capture.

We define a spike in demand returns as any observed demand return greater than four standard deviations larger than the mean. Tables 33.4 and 33.5 collates the occurrences of spikes as defined. Panel (a) shows the occurrence of spikes by region and in aggregate for weekday, month, and year. Panels (b) and (c) show the occurrence of spikes by half-hourly trading interval.

Table 33.4 Panel (a) shows that in aggregate there are 208 spikes in demand returns observed across all regions during the sample period. VIC1 shows the highest incidence of demand spikes with 92 (44 %) of the 208 observed, followed by NSW1 with 81 (38 %), and SA1 with 22 (10 %) spikes during the sample period.

By day of the week, Monday shows the highest incidence with 67 (32 %) tapering gradually to Sunday with 15 occurrences (7 %). August shows the highest incidence by month with 42 (20 %), and of these 37 occur in NSW1. The next

Table 33.4 Panel (a) Summary of occurrences of extreme demand spikes by region, day of week, month, and year

Interval	NSW1	QLD1	SA1	VIC1	Total
Sun	1	4	4	6	15
Mon	39	3	8	17	67
Tue	12	2	1	21	36
Wed	12	3	4	13	32
Thu	13	0	3	15	31
Fri	4	0	0	15	19
Sat	0	1	2	5	8
Jan	0	0	0	23	23
Feb	0	0	0	17	17
Mar	1	1	5	31	38
Apr	0	0	0	8	8
May	13	0	2	0	15
Jun	20	2	3	0	25
Jul	3	4	4	0	11
Aug	37	1	1	3	42
Sep	4	1	0	1	6
Oct	2	0	1	1	4
Nov	1	4	4	3	12
Dec	0	0	2	5	7
1999	15	4	4	78	101
2000	13	0	2	9	24
2001	17	0	2	4	23
2002	16	3	0	1	20
2003	10	1	0	0	11
2004	7	4	3	0	14
2005	3	1	11	0	15
2006	0	0	0	0	0
Total	***81***	***13***	***22***	***92***	***208***

highest incidences by month are March (38) and June (25), with spikes predominantly in VIC1 for March and in NSW1 in June. The highest incidence by year occurs in 1999 with 101 spikes (49 %), dropping markedly in 2000 (24) and 2001 (23). The incidence of spikes appears to have settled somewhat from 2003 onwards at around 15–16 spikes per year.[3]

Table 33.5 Panel (b) shows the incidence of extreme spikes in demand returns by half-hourly trading interval. There are concentrations of spikes occurring at the 06:30 (91 spikes, of which 78 occur in NSW1) and 23:30 (80 spikes, all of which occur in VIC1). A subperiod analysis of demand returns suggests that sharp peaks in demand

[3]Sample data for 2006 only includes the month of January and is unlikely to be representative of the full year.

Table 33.5 Panel (b) Occurrence of extreme demand spikes by half-hourly trading interval

Interval	NSW1	QLD1	SA1	VIC1	Total
H0000	0	0	1	0	1
H0030	0	0	0	0	0
H0100	0	0	0	0	0
H0130	0	0	0	0	0
H0200	0	0	0	0	0
H0230	0	0	0	0	0
H0300	0	0	0	0	0
H0330	0	0	0	0	0
H0400	0	0	0	0	0
H0430	0	0	0	0	0
H0500	0	1	0	0	1
H0530	1	0	0	2	3
H0600	0	0	0	0	0
H0630	78	7	0	6	91
H0700	0	0	5	0	5
H0730	0	1	2	0	3
H0800	0	1	1	1	3
H0830	0	0	1	0	1
H0900	1	1	2	1	5
H0930	0	0	0	0	0
H1000	0	0	0	0	0
H1030	0	0	0	0	0
H1100	0	0	0	0	0
H1130	0	0	0	0	0
H1200	0	0	0	1	1
H1230	0	0	0	0	0
H1300	0	0	1	0	1
H1330	0	0	2	0	2
H1400	0	0	0	0	0
H1430	0	0	1	0	1
H1500	0	0	0	0	0
H1530	0	0	0	0	0
H1600	0	0	0	0	0
H1630	0	0	0	0	0
H1700	0	0	0	0	0
H1730	0	1	0	0	1
H1800	1	1	1	0	3
H1830	0	0	5	1	6
H1900	0	0	0	0	0
H1930	0	0	0	0	0
H2000	0	0	0	0	0
H2030	0	0	0	0	0

(*continued*)

Table 33.5 (continued)

Interval	NSW1	QLD1	SA1	VIC1	Total
H2100	0	0	0	0	0
H2130	0	0	0	0	0
H2200	0	0	0	0	0
H2230	0	0	0	0	0
H2300	0	0	0	0	0
H2330	0	0	0	80	80

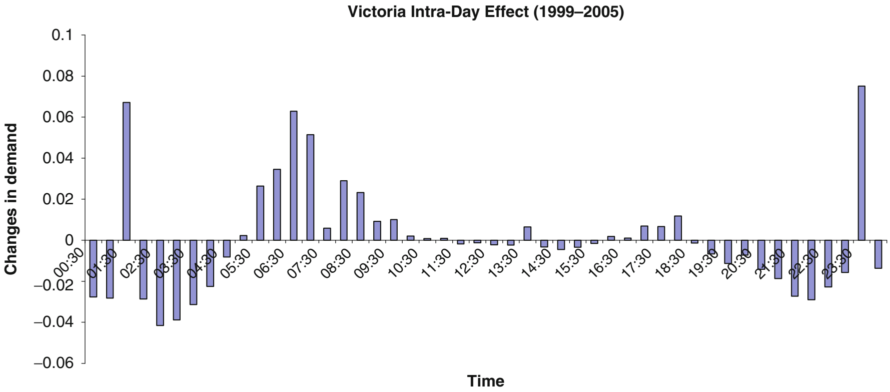


Fig. 33.4 Half-hourly demand returns for VIC1 for the period 1999–2005

returns are persistent throughout the sample period, as illustrated by Fig. 33.4. Figure 33.4 shows the pattern of demand returns for VIC1 for the period 1999–2005 and is illustrative of the pattern in NSW1. The 06:30 peak in demand appears consistent with the commencement of the morning peak in activity in the population. We believe that the 2,330 peak coincides with the activation of off-peak hot water systems set to take advantage of overnight off-peak retail electricity tariffs.

Descriptive statistics for the half-hourly demand return series are shown in Table 33.6. We report the mean, standard deviation, minimum, maximum, range, skewness, kurtosis, and Augmented Dickey-Fuller statistics for each region's demand return series.

Mean, standard deviation, maximum, and minimum are expressed in terms of half-hourly percentage return and are broadly consistent across NSW1, QLD1, SA1, and VIC1. The distributions of demand returns for all four regions demonstrate positive skewness and high positive kurtosis. Jarque-Bera (JB) statistics reject the null hypothesis of normal distribution at the 1 % level of significance for all four regions. This fat-tailed character is consistent with studies on price behavior (see Huisman and Huurman 2003; Higgs and Worthington 2005; Wolack 2000) and appears driven by the presence of spikes in demand returns. Augmented Dickey-Fuller (ADF) statistics robustly reject the hypothesis of a unit root at the 1 % level of significance for all five regions, again consistent with the findings of the earlier studies.

Table 33.6 Descriptive statistics for half-hourly demand returns, by region, January 1999 to January 2006

Returns	NSW1	QLD1	SA1	VIC1
Mean[a]	0.05	0.04	0.05	0.04
S.D.[a]	3.15	2.87	3.31	2.98
Maximum[a]	44.50	52.64	38.41	49.52
Minimum[a]	−26.55	−31.94	−38.97	−33.81
Skewness	1.13	0.94	0.39	1.11
Kurtosis	5.00	5.43	4.45	5.01
JB stat	47,301.83	49,025.43	14,003.32	46,573.67
ADF[b]	−42.57	−43.29	−41.51	−41.51
N	1,24,223	1,24,223	1,24,223	1,24,223

[a]Mean, standard deviation, maximum, and minimum are expressed as percentage values
[b]Augmented Dickey-Fuller (ADF) statistic rejects the hypothesis of a unit root at the 1 % level of confidence

33.3.3 Modelling Seasonality in Demand for Electricity

In this section, we adapt and adjust Eq. 33.2 to model seasonality in demand return. The equation is as follows:

$$RD_{it} = \phi_0 + \beta_1 LR_t + \beta_2 \sum_{n=1,\, i\neq \text{Sun}}^{6} DAY_i + \beta_3 \sum_{n=1,\, i\neq \text{Sep}}^{11} MTH_i + \beta_4 \sum_{n=1999,\, i\neq 2001}^{2006} YR_i + \beta_5 \sum_{n=1,\, i\neq 1,130\ \text{hrs}}^{47} HH_i + \beta_6 \sum_{n=1}^{N_S} DDSPIKE_i + \varepsilon_t \tag{33.5}$$

where RD_{it} represents the discrete demand return for region i at time t, ϕ_0 represents the constant term, LR_{it} represents the lagged demand return for region i at time t, $DDSPIKE_i$ represents the dummy variable set for each occurrence of extreme return as previously defined, and the remaining variables are defined as in Eq. 33.2. The trading interval at 11:30 h, Sunday, September, and the year 2001 were incorporated into the constant term α in the model as the base case for each dummy series. These base cases were selected as the trading interval, day, month, and year, in which demand return activity was consistently lowest in all four regions.

33.4 Empirical Results

Results of the regression analysis are presented in Tables 33.7 and 33.8. Coefficients and t-statistics are presented for each seasonal dummy variable and for lagged returns. In view of the very large number of individual spikes in demand returns (208 spikes identified for the sample period across all regions), coefficients for individual spikes are not explicitly reported though results are discussed.

Table 33.7 Panel (a) – Results Of regression analysis for half-hourly demand return against seasonal dummy variables, by region for day, month, and year

	NSW1		QLD1		SA1		VIC1	
	Coeff	t-stats	Coeff	t-stats	Coeff	t-stats	Coeff	t-stats
C	−.00603	***−8.05***	−.00170	***−3.66***	.00356	***4.02***	−.00228	***−7.43***
Mon	0.00099	***7.88***	0.00084	***3.24***	0.00091	***4.57***	0.00058	***4.24***
Tue	0.00014	***1.41***	0.00020	***0.80***	0.00011	***0.70***	0.00004	***0.30***
Wed	0.00010	***0.96***	0.00014	***0.57***	0.00005	***0.34***	0.00006	***0.42***
Thu	0.00004	***0.30***	−0.00020	***−0.79***	−0.00047	***−2.36***	0.00002	***0.17***
Fri	−0.00099	***−7.27***	−0.00105	***−4.05***	−0.00092	***−4.28***	−0.00062	***−4.58***
Sat	−0.00025	***−1.87***	−0.00033	***−1.28***	−0.00018	***−0.85***	−0.00021	***−1.58***
Jan	−0.00018	***−0.38***	0.00001	***0.03***	0.00009	***0.12***	−0.00008	***−0.46***
Feb	0.00005	***0.10***	−0.00005	***−0.13***	0.00016	***0.23***	−0.00004	***−0.20***
Mar	0.00003	***0.05***	−0.00008	***−0.22***	0.00006	***0.09***	−0.00018	***−0.98***
Apr	−0.00009	***−0.19***	−0.00025	***−0.68***	−0.00014	***−0.20***	−0.00011	***−0.61***
May	0.00009	***0.20***	0.00009	***0.24***	0.00022	***0.31***	0.00004	***0.24***
Jun	0.00003	***0.07***	0.00010	***0.27***	0.00038	***0.54***	0.00003	***0.18***
Jul	0.00011	***0.23***	0.00012	***0.34***	0.00016	***0.23***	0.00002	***0.12***
Aug	−0.00003	***−0.06***	0.00005	***0.13***	0.00040	***0.61***	0.00000	***0.00***
Oct	0.00000	***−0.01***	0.00001	***0.02***	0.00009	***0.14***	0.00001	***0.03***
Nov	0.00008	***0.18***	0.00001	***0.02***	0.00013	***0.19***	0.00004	***0.22***
Dec	0.00012	***0.26***	−0.00010	***−0.28***	0.00041	***0.59***	0.00000	***0.03***
1999	−0.00009	***−0.20***	−0.00006	***−0.24***	−0.00019	***−0.31***	−0.00031	***−2.26***
2000	0.00003	***0.06***	0.00007	***0.25***	0.00002	***0.03***	0.00000	***−0.01***
2002	−0.00002	***−0.05***	−0.00001	***−0.03***	−0.00004	***−0.06***	0.00001	***0.10***
2003	0.00001	***0.02***	0.00001	***0.03***	0.00003	***0.05***	0.00002	***0.14***
2004	−0.00005	***−0.12***	−0.00001	***−0.04***	0.00004	***0.07***	0.00002	***0.15***
2005	0.00002	***0.05***	0.00001	***0.03***	0.00009	***0.15***	0.00003	***0.18***
2006	−0.00013	***−0.13***	−0.00008	***−0.11***	−0.00129	***−0.86***	0.00005	***0.15***
R^2	0.87		0.83		0.82		0.82	0.87
Adj R^2	0.87		0.83		0.82		0.82	0.87

We find that coefficients are relatively small for NSW1, QLD1, SA1, and VIC1. Day-of-week effects are positive and significant for Monday in NSW1, QLD1, SA1, and VIC1. Thursday shows significant negative effect in demand returns in SA1 only. There is significant negative Friday effect in all regions. There appears to be no clear pattern evident for monthly effect across regions. There is no clear pattern evident in yearly effect, although significant negative effect is observed for 1999 in VIC1. Half-hourly time-of-day effects offer more interesting results and are broadly more consistent across regions than the seasonal effects previously discussed. In general, significant negative demand returns are found for the small hours of the morning from 12:30 a.m. until approximately 4:00–5:00 a.m. in all regions. VIC1 exhibits an unexplained highly significant positive return at 1:30 a.m., reverting to negative returns until for the remainder of the early morning. There is wide variation in the pattern of demand returns during the waking day. NSW1 and VIC1 show broadly

Table 33.8 Panel (b): Results of regression analysis for demand return against seasonal dummy variables, by half−hourly trading interval by region, 0000–2,330 h

	NSW1		QLD1		SA1		VIC1	
	Coeff	t-stats	Coeff	t-stats	Coeff	t-stats	Coeff	t-stats
H0000	−0.013	***−14.30***	−0.040	***−91.44***	0.044	***49.76***	−0.055	***−139.42***
H0030	−0.008	***−8.91***	−0.057	***−114.74***	−0.013	***−14.63***	−0.017	***−48.79***
H0100	−0.014	***−15.48***	−0.047	***−89.84***	0.004	***4.89***	−0.010	***−26.90***
H0130	−0.014	***−15.96***	−0.043	***−87.43***	−0.062	***−70.67***	0.086	***238.13***
H0200	−0.020	***−22.39***	−0.036	***−74.58***	−0.064	***−68.82***	−0.066	***−169.17***
H0230	−0.019	***−21.03***	−0.029	***−61.46***	−0.053	***−55.84***	−0.023	***−62.27***
H0300	−0.009	***−10.63***	−0.019	***−41.36***	−0.040	***−42.74***	−0.012	***−33.13***
H0330	−0.005	***−5.91***	−0.010	***−23.47***	−0.031	***−33.90***	−0.006	***−17.25***
H0400	0.005	***5.71***	−0.004	***−9.95***	−0.023	***−25.48***	−0.002	***−5.04***
H0430	0.015	***17.58***	0.005	***12.12***	−0.013	***−15.02***	0.007	***20.52***
H0500	0.020	***22.54***	0.013	***29.54***	−0.004	***−4.49***	0.009	***26.51***
H0530	0.048	***54.16***	0.034	***76.97***	0.005	***5.53***	0.027	***77.09***
H0600	0.039	***43.14***	0.045	***93.50***	0.022	***24.88***	0.021	***59.57***
H0630	0.049	***53.66***	0.079	***157.08***	0.028	***31.42***	0.045	***123.52***
H0700	0.045	***48.87***	0.084	***133.88***	0.052	***58.70***	0.017	***44.80***
H0730	0.011	***12.11***	0.082	***133.22***	0.041	***44.66***	−0.022	***−57.89***
H0800	0.036	***40.06***	0.058	***94.63***	0.023	***25.87***	0.028	***78.38***
H0830	0.020	***22.05***	0.035	***67.65***	0.031	***35.51***	0.009	***24.48***
H0900	0.003	***3.91***	0.018	***38.06***	0.019	***21.19***	−0.002	***−6.33***
H0930	0.015	***17.25***	0.014	***30.12***	0.007	***8.13***	0.007	***20.14***
H1000	0.001	***1.24***	0.004	***8.62***	0.005	***5.81***	−0.001	***−3.97***
H1030	0.002	***2.45***	−0.001	***−1.36***	−0.002	***−3.36***	0.002	***5.85***
H1100	0.002	***3.03***	−0.002	***−6.49***	−0.001	***−0.98***	0.003	***7.98***
H1200	0.004	***5.80***	−0.001	***−5.05***	−0.001	***−2.53***	0.002	***6.15***
H1230	0.003	***3.82***	−0.004	***−10.98***	0.002	***2.08***	0.001	***2.39***
H1300	0.002	***2.03***	−0.005	***−10.64***	0.000	***0.23***	0.001	***3.86***
H1330	0.004	***5.03***	−0.001	***−3.23***	−0.002	***−2.92***	0.010	***28.51***
H1400	0.002	***2.14***	−0.001	***−2.88***	0.006	***6.97***	−0.005	***−12.89***
H1430	0.005	***5.82***	0.000	***0.89***	−0.010	***−11.42***	0.000	***−0.61***
H1500	0.005	***5.44***	−0.003	***−6.23***	−0.006	***−6.46***	0.002	***4.52***
H1530	0.007	***7.57***	−0.001	***−1.98***	−0.010	***−11.66***	0.003	***7.70***
H1600	0.011	***12.34***	0.004	***9.22***	−0.009	***−9.88***	0.005	***14.29***
H1630	0.008	***8.69***	0.008	***17.99***	−0.004	***−4.93***	0.002	***6.33***
H1700	0.019	***22.22***	0.016	***36.28***	−0.003	***−3.59***	0.009	***23.96***
H1730	0.014	***15.77***	0.021	***46.66***	0.005	***5.65***	0.005	***13.52***
H1800	0.027	***30.31***	0.027	***58.50***	0.010	***11.86***	0.010	***28.01***
H1830	−0.001	***−1.18***	0.013	***28.47***	0.015	***16.97***	−0.006	***−16.80***
H1900	−0.001	***−1.01***	0.006	***14.25***	0.002	***2.06***	−0.004	***−10.01***
H1930	−0.004	***−5.07***	−0.010	***−21.73***	−0.008	***−8.85***	−0.005	***−13.97***
H2000	−0.002	***−2.52***	−0.014	***−32.78***	−0.012	***−14.03***	0.002	***5.10***

(*continued*)

Table 33.8 (continued)

	NSW1		QLD1		SA1		VIC1	
	Coeff	t-stats	Coeff	t-stats	Coeff	t-stats	Coeff	t-stats
H2030	−0.010	***−11.60***	−0.023	***−51.22***	−0.010	***−11.32***	−0.007	***−20.68***
H2100	−0.008	***−8.61***	−0.022	***−48.41***	−0.017	***−19.81***	−0.008	***−23.12***
H2130	0.004	***4.81***	−0.021	***−47.15***	−0.021	***−23.79***	−0.014	***−39.32***
H2200	−0.012	***−13.21***	−0.026	***−57.72***	−0.021	***−24.30***	−0.011	***−29.98***
H2230	0.028	***31.77***	−0.016	***−35.83***	−0.024	***−27.00***	−0.004	***−10.35***
H2300	−0.014	***−15.49***	−0.019	***−42.48***	0.000	***−0.10***	0.000	***−0.52***
H2330	0.000	***0.14***	−0.005	***−11.30***	−0.009	***−10.29***	0.085	***234.83***

similar intraday patterns, with significant positive demand returns dominating between 04:30 and 18:00 and with minor variation reverting to negative demand returns for the remainder of the evening. QLD1 and SA1 are broadly similar, demonstrating significant positive effect between 04:30 and 10:30 and predominantly negative effect from 10:30 to 15:30, reverting to positive effect from 16:00 to 19:30 when we see reversion to negative effect for the remainder of the evening. The periods of positive return in the early morning and early evening are consistent with peaks in activity in the population. We would expect to observe positive demand returns arising from off-peak hot water systems generally switching on at 11:00 p.m., but curiously positive returns around that hour are evident only in NSW1.

33.5 Conclusions

The current literature establishes that electricity time series differ from traditional financial data having greater incidence of spikes than is generally observed in financial data and this results in extreme volatility. The work done overseas is of limited practical application to the Australian market as our market structure is different. The lesson learned from our study is that even within the same country, variation exists as different states reflect variation in geographic differences. The implication is that specific models must be developed for each geographic region. Our work is innovative in that we develop an econometric time series seasonal model that can be applied in areas where seasonality is suspected. For instance, the model was initially applied to price and return series and then later on to demand return series. This model has the capability to be applied in other financial time series models.

Developing models to explain and to predict electricity prices is a significant task. This is important both for the market participants who operate in the physical market and for those trading in electricity derivatives. Modelling electricity prices and trading in this market is challenging, so much so that investment banks avoid trading electricity derivatives because of the incidence of spikes. This provides a strong incentive for this research and future research into the electricity market. A better understanding of the time series nature of the electricity prices may help in the development of more efficient forecasting models which will help to lower risk management costs and thus reduce the cost of managing electricity price exposures. Electricity generators, electricity retailers, transmission and distribution network suppliers, electricity retailer companies, and electricity consumers will all benefit from development of more accurate models in this area.

Acknowledgment The authors wish to acknowledge the financial support of the Australian Centre for Financial Studies, via grant 18/2005. The research on the seasonality within the spot price has been published in 2011, "Seasonal factors and outlier effects in rate of return in Australia's National Electricity Market," *Applied Economics*, Vol. 43, Issue 3; pp. 355–369. Binesh Seetanah provided invaluable research assistance. Any remaining errors are the responsibility of the authors.

References

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy, 81*(3), 637.

Bunn, D., & Karakatsani, N. (2003). *Forecasting electricity prices*. Working paper, London Business School.

Clewlow, L., & Strickland, C. (2000). *Valuing energy options in a one-factor model fitted to forward prices*. Working paper, University of Sydney.

Eichler, M., Grothe, O., Manner, H., & Tuerk, D. (2012). *Modeling spike occurrences in electricity spot prices for forecasting*. Working paper, RM/12/029, Maastricht University.

Engle, R., Granger, C., & Hallman, J. (1989). Merging short and long-run forecasts: An application of seasonal cointegration to monthly electricity sales forecasting. *Journal of Econometrics, 50*, 297–327.

Escribano, A. J., Peaea J., & Villaplana, P. (2002). *Modelling electricity prices: International evidence*. Working paper, Universidad Carlos III de Madrid.

Goto, M., & Karolyi, G. A. (2004). *Understanding electricity price volatility within and across markets*. Dice Centre working paper 2004–12, pp. 1–41.

Harvey, A., & Koopman, S. J. (2002). Forecasting hourly electricity demand using time-varying splines. *Journal of the American Statistical Association, 88*(424), 1228–1236.

Huisman, R., & Huurman, C. (2003). Fat tails in power prices. ERIM report series reference no. ERS-2003-059-F&A. SSRN: http://ssrn.com/abstract=450992

Johnson, B., & Barz, G. (1999). Selecting stochastic processes for modelling electricity prices. In *Energy modelling and the management of uncertainty*. Risk Publications.

Kaminski, V. (1997). The challenge of pricing and risk managing electricity derivatives. In *The U.S. Power Market* (pp. 149–71).

Knittel, C. R., & Roberts, M. (2001). An empirical examination of deregulated electricity prices. Working paper, PWP-087, University of California Energy Institute.

Montero, J. M., Garcia, M. C., & Fernandez-Aviles, G. (2011). Modelling the volatility of the Spanish wholesale electricity spot market. Asymmetric GARCH models vs. Threshold ARSV model. *Revista de Estudios de Economía Aplicada, 29*(2), 1–20.

Pardo, A., Meneu, V., & Valor, E. (2002). Temperature and seasonality influences on Spanish electricity load. *Energy Economics, 24*, 55–70.

Smith, M. (2000). Modelling and short-term forecasting of New South Wales electricity system load. *Journal of Business & Economic Statistics, 18*(4), 465–478.

Thomas, S., Ramiah, V., Heaney, R., & Mitchell, H. (2011). Seasonal factors and outlier effects in rate of return in Australia's national electricity market. *Applied Economics, 43*(3), 355–369.

Pricing Commercial Timberland Returns in the United States

34

Bin Mei and Michael L. Clutter

Contents

B. Mei (✉) • M.L. Clutter
Warnell School of Forestry and Natural Resources, University of Georgia, Athens, GA, USA
e-mail: bmei@uga.edu; mclutter@warnell.uga.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_34,

Abstract

Commercial timberland assets have attracted more attention in recent decades. One unique feature of this asset class roots in the biological growth, which is independent of traditional financial markets. Using both parametric and non-parametric approaches, we evaluate private- and public-equity timberland investments in the United States. Private-equity timberland returns are proxied by the NCREIF Timberland Index, whereas public-equity timberland returns are proxied by the value-weighted returns on a dynamic portfolio of the US publicly traded forestry firms that had or have been managing timberlands. The results from parametric analysis reveal that private-equity timberland investments outperform the market and have low systematic risk, whereas public-equity timberland investments fare similarly as the market. The nonparametric stochastic discount factor analyses reveal that both private- and public-equity timberland assets have higher excess returns.

Static estimations of the capital asset pricing model and Fama-French three-factor model are obtained by ordinary least squares, whereas dynamic estimations are by state space specifications with the Kalman filter. In estimating the stochastic discount factors, linear programming is used.

Keywords

Alternative asset class • Asset pricing • Evaluation • Fama-French three-factor model • Nonparametric analysis • State space model • Stochastic discount factor • Timberland investments • Time series • Time-varying parameter

34.1 Introduction

Timberland investments have been unprecedentedly active in the United States the past 30 years. A number of factors have motivated public attention toward timberland assets. On the supply side, because of the internal subsidies from timber divisions to processing mills, timberland properties managed by traditional vertically integrated forest products firms have been undervalued by the Wall Street. To fix this mispricing, these firms began divesting their timberlands as a strategic move. For example, International Paper, a global leading forest products firm, has disposed of most of its timberlands and focused on its core business of paper and packaging products manufacturing in recent decades. Currently, almost no forest products firms in the United States still own timberlands (Harris et al. 2010). On the demand side, institutional investors, e.g., organizations with fiduciary obligations such as pension funds, university endowments, foundations, and trusts, have diversified into nonfinancial asset classes such as commercial timberlands on the passage of Employee Retirement Income Security Act (ERISA) in 1974.

There are several ways to invest in commercial timberlands. High-net-wealth families and individuals can participate in commingled (pooled) funds, or they can own and manage timberlands directly. Others can buy stocks and bonds of publicly traded forestry firms that focus on timberland business. Most institutional investors hold timberland properties via timberland investment management organizations (TIMOs). TIMOs manage their institutional assets in either separately managed accounts (individually managed accounts) or pooled funds. A separately managed account holds timberland properties of one investor in a single portfolio, whereas a pooled fund collects capital from a number of investors and allocates it to a portfolio of timberland properties. Investors tend to have more discretion with separate accounts than pooled funds (Zinkhan and Cubbage 2003). In 2008, there were about 30 TIMOs in the United States, and the total value of their timberland assets exceeded $35 billion (Zinkhan 2008).

Since the public recognition of commercial timberland as an alternative asset class, a number of studies have been conducted to assess the financial performance of timberland investments. The major findings of previous research can be summarized as follows. (1) Timberland has countercyclical returns or low (even negative in some cases) correlation with the financial assets (Binkley et al. 1996; Cascio and Clutter 2008; Redmond and Cubbage 1988; Washburn and Binkley 1990; Zinkhan 1988). (2) Timberland can be an effective hedge against unexpected inflation (Fortson 1986; Washburn and Binkley 1993). (3) If timberland investors can exploit the biological growth of timber thus time the market, they can get higher and better returns (Caulfield 1998; Conroy and Miles 1989; Haight and Holmes 1991). (4) Relative inefficiency tends to exist in timberland markets (Caulfield 1998), although this situation has been alleviated through time (Washburn 2008; Zinkhan 2008). (5) Among a variety of forestry-related investment vehicles, institutional timberland investments and timberland limited partnerships have low-risk levels but excess returns (Sun and Zhang 2001). (6) In the long run, timber and/or timberland returns are cointegrated with other nontimber financial instruments (Heikkinen 2002; Liao et al. 2009).

Almost all of the above studies are based on the single-period capital asset pricing model (CAPM). Sun and Zhang (2001) extended the literature in timberland investments by employing the arbitrage pricing theory (APT), and Heikkinen (2002) and Liao et al. (2009) expanded the literature by using cointegration analysis. Nevertheless, all those methods are parametric in nature. This study has several contributions. First, timberland assets are considered separately in private and public markets, and their returns are compared. Second, supplementary to the ordinary least squares (OLS) estimation of the CAPM and the Fama-French three-factor model, a state space model with the Kalman filter is employed to examine the time-varying risk-adjusted excess return (α) and systematic risk (β). Finally, the nonparametric stochastic discount factor (SDF) approach is introduced for pricing timberland returns.

The major results are that private-equity timberland investments have significant excess returns but low systematic risk, whereas public-equity timberland investments fare similarly as the market, and that intertemporal consumption decisions affect the intertemporal marginal rate of substitution of timberland investors and thus impact the rational pricing of timberland assets. These results can further our understanding of the financial aspects of commercial timberland assets in the United States. The next two sections describe the methodologies and the data. Section 34.4 explains the empirical results, and the last section makes some concluding remarks.

34.2 Methods

For the parametric method, an explicit model is needed. Two candidate models prevalent in the finance literature are the CAPM and the Fama-French three-factor model. The parametric method is often criticized for the "joint hypothesis tests" problem, i.e., testing the asset pricing model and the abnormal performance (market efficiency) simultaneously. The nonparametric method does not require such an explicit model specification and is therefore not subject to these critiques. The SDF approach is a general, nonparametric asset pricing approach and is complement to the parametric approaches.

34.2.1 CAPM

Built on Markowitz's (1952) groundwork of mean-variance efficient portfolio, Sharpe (1964) and Lintner (1965) developed its economy-wide implications – the CAPM. The CAPM states that the expected return on an asset or a portfolio $E[R_i]$ equals a risk-free rate R_f plus a premium that depends on the asset's β_i and the expected risk premium on the market portfolio $E[R_m] - R_f$, i.e.,

$$E[R_i] = R_f + \beta_i\left(E[R_m] - R_f\right). \tag{34.1}$$

In empirical regression analysis, the CAPM is estimated in the excess return form

$$R_i - R_f = \alpha_i + \beta_i\left(R_m - R_f\right) + \mu_i, \tag{34.2}$$

where ex post realized returns R_i and R_m rather than ex ante expected returns $E[R_i]$ and $E[R_m]$ are used. The intercept α_i is called Jensen's (1968) alpha. A positive α suggests that the individual asset outperforms the market and earns a higher than risk-adjusted return, whereas a negative α suggests that the individual asset underperforms the market and earns a lower than risk-adjusted return. Therefore, Jensen's alpha has become a commonly used measure of abnormal performance, and testing whether it is zero has been widely used in the empirical asset pricing literature.

34.2.2 Fama-French Three-Factor Model

Given the empirical evidence that small size stocks outperform large size stocks, and value (high book-to-market) stocks outperform growth (low book-to-market) stocks on average, Fama and French (1993) develop a model that includes these extra two factors to adjust for risk:

$$E[R_i] - R_f = \beta_{RMRF,i} E[R_{RMRF}] + \beta_{SMB,i} E[R_{SMB}] + \beta_{HML,i} E[R_{HML}], \tag{34.3}$$

where $R_{RMRF} = R_m - R_f$ is the same market factor as in the CAPM, representing the market risk premium; $R_{SMB} = R_{\text{small}} - R_{\text{big}}$ is the size factor, representing the return difference between a portfolio of small stocks and a portfolio of large stocks (SMB stands for "small minus big"); $R_{HML} = R_{\text{highBM}} - R_{\text{lowBM}}$ is the book-to-market factor, representing the return difference between a portfolio of high book-to-market stocks and a portfolio of low book-to-market stocks (HML stands for "high minus low"); and β's are called factor loadings, representing each asset's sensitivity to these factors. When estimating Fama-French three-factor model, ex post realized returns are used, as in the case of the CAPM, and an intercept is added to capture the abnormal performance,

$$R_i - R_f = \alpha_i + \beta_{RMRF,i} R_{RMRF} + \beta_{SMB,i} R_{SMB} + \beta_{HML,i} R_{HML} + \varepsilon_i. \tag{34.4}$$

34.2.3 CAPM and Fama-French Three-Factor Model Under the State Space Framework

The CAPM (Eq. 34.2) and the Fama-French three-factor model (Eq. 34.4) are usually estimated by OLS, possibly with some correction for the autocorrelations in the errors. One restrictive nature of the OLS method is that the coefficients in the regression are imposed to be constant. This condition may be unrealistic in real asset pricing modeling. For instance, one would suspect that both α's and β's should be time varying. To solve this problem, we can estimate the CAPM and the Fama-French three-factor model in the state space framework with the Kalman filter (Appendix 1). Using the CAPM as an example, in the state space framework, the system of equations is specified as

$$\begin{aligned} R_{i,t} - R_{f,t} &= \alpha_{i,t} + \beta_{i,t}\left(R_{m,t} - R_{f,t}\right) + \mu_{i,t} \\ \alpha_{i,t+1} &= \alpha_{i,t} + \xi_t \\ \beta_{i,t+1} &= \beta_{i,t} + \tau_t \end{aligned} \tag{34.5}$$

where $\mu_{i,t}$, ξ_t, and τ_t are normally and independently distributed mean-zero error terms. In the state space model, the first equation in (34.5) is called the observation or measurement equation, and the second and third equations are called the state equations. In this particular case, each state variable follows a random walk.

One advantage of the state space approach with time-varying parameters is that it can incorporate external shocks, such as policy and regime shifts, economic reforms, and political uncertainties, into the system, especially when the shocks are diffuse in nature (Sun 2007). This approach has been applied to a variety of issues, including demand systems (e.g., Doran and Rambaldi 1997), aggregate consumptions (e.g., Song et al. 1996), policy analysis (e.g., Sun 2007), and price modeling and forecasting (e.g., Malaty et al. 2007).

34.2.4 Stochastic Discount Factor Approach

The single-period asset pricing models ignore the consumption decisions. In effect, investors make their consumption and portfolio choices simultaneously in an intertemporal setting. In the framework of an exchange economy in which an investor maximizes the expectation of a time-separable utility function (Lucas 1978), it can be proved that (Appendix 2)

$$E_t\left[\left(1+R_{i,t+1}\right)M_{t+1}\right]=1, \tag{34.6}$$

where $R_{i,t+1}$ is the return on asset i in the economy and M_{t+1} is known as the stochastic discount factor, or intertemporal marginal rate of substitution, or pricing kernel (e.g., Campbell et al. 1997).

Hansen and Jagannathan (1991) demonstrated how to identify the SDF from a set of basis assets, i.e., the derivation of the volatility bounds. These bounds are recognized as regions of admissible mean-standard deviation pairs of the SDF. Their major assumptions are the law of one price and the absence of arbitrage opportunities. Accordingly, there are two particular solutions for the SDF: the law of one price SDF and the no-arbitrage SDF. The process of retrieving the reverse-engineered law of one price SDF is equivalent to the following constrained optimization problem:

$$\begin{aligned}
&\underset{M_t}{\text{Min}}\,\sigma_{M_t}=\left[\frac{1}{T-1}\sum_{t=1}^{T}\left(M_t-v\right)\right]^{1/2}\\
&s.t.\frac{1}{T}\sum_{t=1}^{T}M_t=v\\
&\quad\ \frac{1}{T}\sum_{t=1}^{T}M_t\left(1+R_{i,t}\right)=1
\end{aligned} \tag{34.7}$$

for a range of selected v (mean of M_t) and for all assets $i=1,2,\cdots,N$. Under the stronger condition of no arbitrage, another positivity constraint on M_t is needed. Therefore, the only difference between the law of one price SDF and the no-arbitrage SDF is whether M_t is allowed to be negative. In this study, no-arbitrage SDF is used. Following Hansen and Jagannathan (1991), nonnegativity instead of positivity restriction $M_t \geq 0$ is added to retrieve the no-arbitrage SDF. Last, sample size T should be sufficiently large such that the time-series version of law of large numbers applies; that is, the sample moments on a finite record converge to their population counterparts as the sample size becomes large (Hansen and Jagannathan 1991).

Provided the existence of a risk-free asset, it can be shown that

$$E_t\left[\left(R_{i,t+1} - R_f\right)M_{t+1}\right] = 0. \tag{34.8}$$

This equation presents the basis for testing the risk-adjusted performance of a portfolio (Chen and Knez 1996). Namely, one can test whether

$$\alpha_i = E_t\left[\alpha_{i,t}\right] = E_t\left[\left(R_{i,t+1} - R_f\right)M_{t+1}\right] = 0. \tag{34.9}$$

Ahn et al. (2003) pointed out that this measure generalizes Jensen's alpha and does not count on a specific asset pricing model. Based on this method, they reassess the profitability of momentum strategies and found that their non-parametric risk adjustment explains almost half of the anomalies.

34.3 Data

34.3.1 Timberland Returns

Returns for both private- and public-equity timberland investments are analyzed. Although TIMOs have become the major timberland investment management entities for institutional investors as well as high-net-wealth families and individuals, their financial data are rarely publicly available. To provide a performance benchmark, several TIMOs, together with National Council of Real Estate Investment Fiduciaries (NCREIF) and the Frank Russell Company, initiated the NCREIF Timberland Index in early 1992 (Binkley et al. 2003) (Appendix 3). NCREIF members can be divided into data contribution members, professional members, and academic members. Data contribution members include investment managers and plan sponsors who own or manage real estate in a fiduciary setting. Professional members include providers of accounting, appraisal, legal, consulting, or other services to the data contribution members. Academic members include full-time professors of real estate. Data contribution members submit their data on a quarterly basis for computation of the NCREIF Property Index. Regarding the NCREIF Timberland Index, it is some TIMOs that are the major data contribution members. The quarterly NCREIF Timberland Index is reported at both regional (the South, the Northeast, and the Pacific Northwest) and national levels, and extends back to 1987. In this study, the national-level NCREIF Timberland Index (1987Q1–2010Q4) is used as a return proxy for the US private-equity timberland investments.

Returns on public-equity timberland investments are proxied by the value-weighted returns on a dynamic portfolio of the US publicly traded forestry firms that had or have been managing timberlands. These firms include Deltic Timber, the Timber Co, IP Timberlands Ltd., Plum Creek, Pope Resources, Potlatch, Rayonier, and Weyerhaeuser. Deltic Timber and Pope Resources are natural resources companies focused on the ownership and management of timberland; the Timber Co and IP Timberlands Ltd. are subsidiaries of Georgia-Pacific and International Paper that track the value and performance of their timberland properties; Plum Creek, Potlatch, and Rayonier are publicly traded real estate investment trusts (REITs) that are engaged in timberland management; and Weyerhaeuser is a forest products firm that has a significant portion of its business in timberlands. The market value of each firm is calculated as the product of stock price and total shares outstanding at the end of each quarter. Financial data for these forestry firms are obtained from the Center for Research in Security Prices (CRSP). To be consistent with the NCREIF Timberland Index, the sample spans from 1987Q1 to 2010Q4.

34.3.2 Basis Assets

To mimic the complete investment opportunity set that is available to investors, a parsimonious set of basis assets needs to be specified. King (1966) proved that industry groupings maximize intragroup correlation and minimize intergroup correlation and concluded that market and industry factors capture most of the common variation in stock returns. Following Hansen and Jagannathan (1991), we construct the reference set by forming industry portfolios according to SIC code. In this study, two sets of basis assets are chosen – one is the five-industry portfolios plus long-term treasury bonds, and the other is the ten-industry portfolios plus long-term treasury bonds. The industry groups are derived from stocks listed on NYSE, AMEX, and NASDAQ based on their four-digit SIC codes. The five industries are classified as consumer goods, manufacturing, Hi-Tech, healthcare, and others, whereas the ten industries are classified as consumer nondurables, consumer durables, manufacturing, energy, Hi-Tech, telephone and television transmission, shops, healthcare, utilities, and others. Value-weighted returns on the industry portfolios are obtained from Kenneth R. French's website, and returns on the portfolio of long-term treasury bonds are obtained from CRSP. Presuming that the basis assets are rationally priced, the SDF can be retrieved.

34.3.3 Other Indices

Market returns are approximated by the value-weighted returns on all NYSE, AMEX, and NASDAQ stocks from CRSP. Risk-free rate, as approximated by the 1-month treasury bill rate from Ibbotson Associates, Inc., and Fama-French factors are available on Kenneth R. French's website.

34.4 Empirical Results

34.4.1 Estimation of the CAPM and the Fama-French Three-Factor Model

Panel A of Table 34.1 presents the OLS estimation of the CAPM and the Fama-French three-factor model using the quarterly NCREIF Timberland Index after adjustment for the seasonality. A significant positive α from the CAPM suggests that private-equity timberland investments have a risk-adjusted excess return of about 8.20 % (2.05 % $\times$ 4) per year. This excess return is slightly larger after accounting for Fama-French factors. Market β's from both models are insignificantly different from zero, but significantly less than one. This means that private-equity timberland investments are not only weakly correlated with the market but also less risky than the market. The small magnitudes with high *p*-values of the coefficients for SMB and HML signify that these two extra factors add limited explanatory power to the CAPM in pricing private-equity timberland returns.

In contrast, the CAPM and the Fama-French three-factor model fit the returns on the dynamic portfolio of forestry firms much better, as implied by the higher R^2 values (Panel A of Table 34.2). This is within our expectation since these forestry firms are publicly traded and are more exposed to the market. However, α's are insignificant albeit positive, indicating no abnormal performance. Market β's are significantly different from zero, but not from one. In addition, β's for SMB and HML in Fama-French three-factor model are significant at the 5 % level or better, meaning these factors capture some variations in the portfolio returns that are not explained by the market premium. As a result, the abnormal performance (α) has dropped by 53 %. The magnitudes of β's indicate that the dynamic portfolio is dominated by mid-large firms with middle book-to-market ratios.

34.4.2 State Space Estimation of the CAPM and the Fama-French Three-Factor Model

Panel B of Table 34.1 presents the state space estimation of the CAPM and the Fama-French three-factor model using the NCREIF Timberland Index. Those OLS coefficient estimates are used as the starting values. Only α is specified as a state variable (stochastic level) in that little time variation is observed in β, and both AIC and SBC favor the deterministic-β model. Back to the model specification in system (Eq. 34.5), this is equivalent to restrict $\tau_t = 0$. The AIC and SBC are marginally larger than those for the OLS estimation because of the relatively small sample size. Figure 34.1 depicts the evolution of the risk-adjusted excess returns of the NCREIF Timberland Index estimated from the CAPM. For most time in the last 22 years, the NCREIF Timberland Index has achieved positive abnormal returns with an average of 10.20 % per year (calculated from the estimated α series). Nevertheless, in certain years (2001–2003), the α is low and even negative,

Table 34.1 Estimation of the CAPM and the Fama-French three-factor model using the NCREIF Timberland Index (1987Q1–2010Q4)

CAPM			FF3		
Coefficient	Estimate	*p*-value	Coefficient	Estimate	*p*-value
Panel A: OLS estimation					
α	2.05	0.001	α	2.11	0.001
β	0.01	0.773	β_{RMRF}	0.02	0.675
			β_{SMB}	−0.04	0. 634
			β_{HML}	−0.05	0.318
H_0: $\beta = 1$		0.000	H_0: $\beta_{RMRF} = 1$		0.000
R^2	0.14		R^2	0.15	
Log likelihood	−251.24		Log likelihood	−250.59	
S.E. of regression	3.78		S.E. of regression	3.82	
Durbin-Watson stat.	1.94		Durbin-Watson stat.	1.96	
Akaike info. criterion	5.53		Akaike info criterion	5.56	
Schwarz criterion	5.61		Schwarz criterion	5.69	
F-stat.	7.09		*F*-stat.	3.82	
Panel B: State space estimation					
α	0.54	0.692	α	0.79	0.549
β	−0.02	0.919	β_{RMRF}	0.01	0.856
			β_{SMB}	−0.11	0.263
			β_{HML}	−0.05	0.581
H_0: $\beta = 1$		0.000	H_0: $\beta_{RMRF} = 1$		0.000
Log likelihood	−275.00		Log likelihood	−273.59	
Akaike info. criterion	5.86		Akaike info criterion	5.80	
Schwarz criterion	5.94		Schwarz criterion	5.94	

(1) OLS estimates after correction for the fourth-order autocorrelation in the residuals. (2) Only α is specified stochastic under the state space framework, while β is specified deterministic due to its lack of variation and AIC criterion

indicating no abnormal performance. Although not reported here, the time-varying α's estimated from Fama-French three-factor model exhibit similar patterns.

For the dynamic portfolio, however, only β is specified to be stochastic since little time variation is observed in α, and both AIC and SBC favor the deterministic-α model. The time-varying β of the dynamic portfolio of forestry firms is plotted in Fig. 34.2. Overall, there is a decreasing trend in the market β. The average β over the sample period is 1.05, which is not significantly different from the market risk.

34.4.3 Abnormal Performance Measured by the SDF Approach

The mean of the no-arbitrage SDF M_t is specified in the selected range of [0.9750, 1] with an increment step of 0.0025. When the five-industry portfolios plus the long-term treasury bonds are used as the basis assets, the global minimum variance of M_t is identified at $v = 0.9800$; when the ten-industry portfolios plus

Table 34.2 Estimation of the CAPM and the Fama-French three-factor model using returns on a dynamic portfolio of publicly traded forestry firms in the United States (1987Q1–2010Q4)

CAPM			FF3		
Coefficient	Estimate	p-value	Coefficient	Estimate	p-value
Panel A: OLS estimation					
α	0.59	0.521	α	0.28	0.750
β	0.95	0.000	β_{RMRF}	0.87	0.000
			β_{SMB}	0.39	0.031
			β_{HML}	0.31	0.001
H_0: $\beta = 1$		0.330	H_0: $\beta_{RMRF} = 1$		0.249
R^2	0.48		R^2	0.54	
Log likelihood	−344.00		Log likelihood	−337.27	
S.E. of regression	8.80		S.E. of regression	8.29	
Durbin-Watson stat.	2.19		Durbin-Watson stat.	2.24	
Akaike info. criterion	7.21		Akaike info criterion	7.11	
Schwarz criterion	7.26		Schwarz criterion	7.22	
F-stat.	85.16		F-stat.	36.58	
Panel B: State space estimation					
α	0.59	0.500	α	0.24	0.786
β	0.95	0.000	β_{RMRF}	0.84	0.000
			β_{SMB}	0.39	0.021
			β_{HML}	0.32	0.001
H_0: $\beta = 1$		0.390	H_0: $\beta_{RMRF} = 1$		0.446
Log likelihood	−355.38		Log likelihood	−348.61	
Akaike info. criterion	7.47		Akaike info criterion	7.37	
Schwarz criterion	7.55		Schwarz criterion	7.50	

Only β is specified stochastic under the stochastic framework, while α is specified deterministic due to its lack of variation and AIC criterion

the long-term treasury bonds are used instead, the global minimum variance of M_t is identified at $v = 0.9750$.

The SDF performance measures for both the NCREIF Timberland Index, and the returns on the dynamic portfolio of publicly traded timber firms are reported in Table 34.3. The α values for both return indices have increased, and the latter has become marginally significant. This indeed implies that intertemporal consumption decisions play a key role in pricing timberland assets. In a word, there is clear evidence of statistically as well as economically significant excess returns for the NCREIF Timberland Index, but only some evidence of economically significant excess returns for the portfolio of publicly traded timber firms.

34.5 Conclusion

Using both parametric and nonparametric techniques, in this study, we reexamined the financial performance of timberland investments. Private-equity timberland

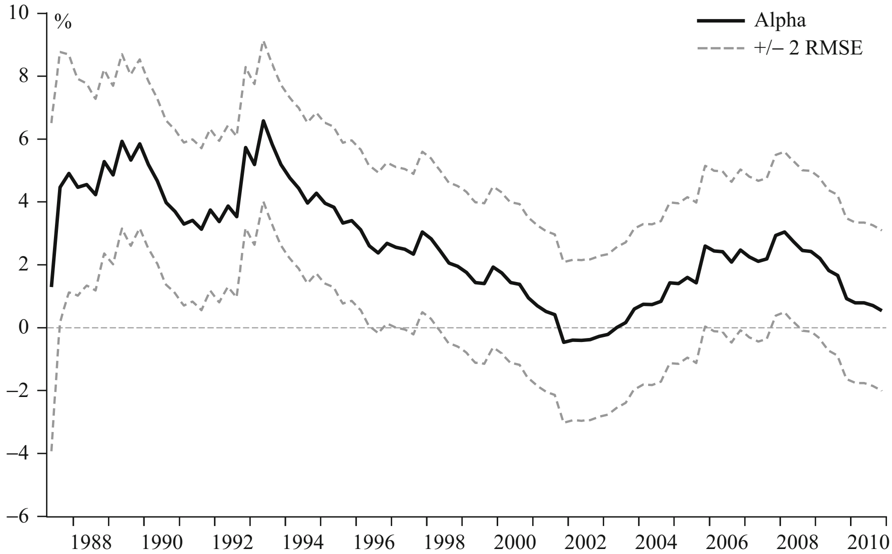


Fig. 34.1 Evolution of α over time from the state space estimation of the CAPM using the NCREIF Timberland Index (1987Q1–2010Q4). Note: The time-varying α estimated from Fama-French three-factor model exhibits similar patterns, thus is not shown separately. The graph is available from the authors upon request

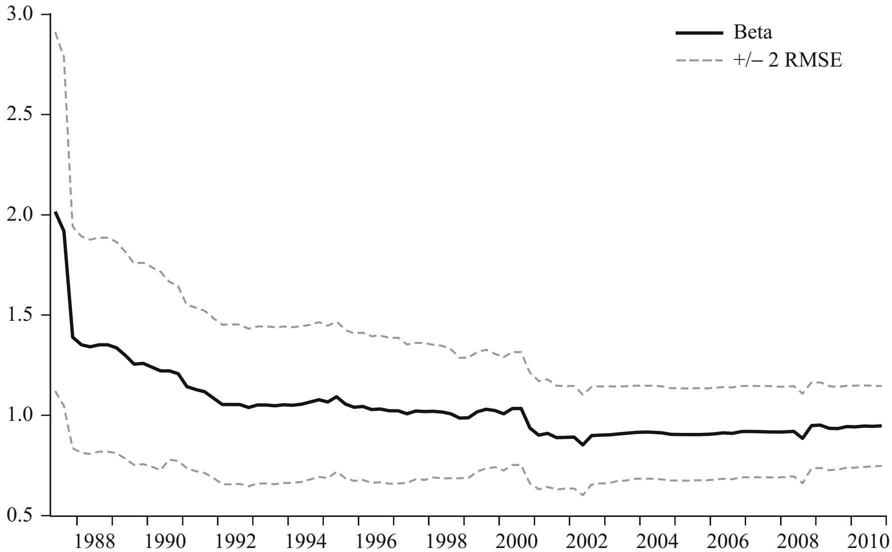


Fig. 34.2 Evolution of β over time from the state space estimation of the CAPM using returns on a dynamic portfolio of publicly traded forestry firms in the United States (1987Q1–2010Q4). Note: The time-varying β estimated from Fama-French three-factor model exhibits similar patterns, thus is not shown separately. The graph is available from the authors upon request

Table 34.3 Performance measures of timberland returns by the nonparametric SDF approach (1987Q1–2010Q4)

		Performance measure (α)		p-value (one tail)	
Mean of M_t (v)	S.D. of M_t (σ_{M_t})	(1)	(2)	(1)	(2)
Panel A: Five-industry portfolios plus long-term T-bonds					
0.9775	0.199	2.63	1.59	0.000	0.119
0.9800	0.176	2.60	1.36	0.000	0.156
0.9825	0.217	2.57	1.13	0.000	0.202
Panel B: Ten-industry portfolios plus long-term T-bonds					
0.9725	0.244	2.76	2.03	0.000	0.056
0.9750	0.237	2.75	1.82	0.000	0.082
0.9775	0.255	2.77	1.60	0.000	0.116

Column (1) is for the NCREIF Timberland Index, and Column (2) is for the returns on a dynamic portfolio of the US publicly traded forestry firms that had or have been managing timberlands

returns are approximated by the NCREIF Timberland Index, whereas public-equity timberland returns are approximated by the value-weighted returns on a dynamic portfolio of the US publicly traded timber firms. The parametric analyses reveal that private-equity timberland assets outperform the market but have low systematic risk, whereas public-equity timberland assets perform similarly as the market. Therefore, inclusion of private-equity timberland properties can improve the efficient frontier, albeit such potential is limited for public-equity timberland properties. Unlike the parametric methods, the nonparametric SDF approach does not rely on any specific asset pricing models and hence are not subject to the "joint hypothesis test" criticisms. The results from the SDF approach suggest higher excess returns for both private- and public-equity timberland investments, which in turn signify the important role of intertemporal consumption decisions in rational pricing of timberland assets.

The positive α of private-equity timberland returns may be associated with the patience of institutional investors toward embedded strategic options for timberlands (Zinkhan 2008). If a timberland property has potential for higher and better use such as residential or commercial development opportunities, or if it is suitable for conservation easements, or if it has mineral or gas opportunities, it may have extra income sources, and the land value can be dramatically higher. The positive α may also be related to the liquidity risk that institutional investors bear since a typical TIMO has an investment time horizon of 10–15 years or even longer. In contrast, stocks of publicly traded timber firms can be easily traded on the stock exchanges. Moreover, initiation of a TIMO-type separately managed account usually requires a capital commitment of \$25–\$50 million, while participation in a TIMO-type pooled fund generally requires a minimum capital commitment of \$1–\$3 million (Zinkhan and Cubbage 2003). The large capital amount may enable the investors to achieve some degree of diversification.

The lower excess returns of the NCREIF Timberland Index around 2001–2003 may be associated with its relative weak performance during that time. In 2001Q4, the NCREIF Timberland Index fell by 6.5 %, the largest drop it ever had, which was primarily caused by the capital loss from the shrinking timberland values. In the same period, the S&P 500 index went up by 7.8 %. The overall decreasing trend in β for the dynamic portfolio of forestry firms may be related to the massive restructurings of these firms. For instance, Plum Creek, Potlatch, and Rayonier have converted themselves into timber REITs in recent years. With improved tax efficiency and increased concentration on timberland management, these timber REITs are expected to be less risky.

Another interesting fact noted in this study is that, despite the current economic downturn triggered by the subprime residential mortgage blowup, private-equity timberland returns remain relatively strong. While the CRSP market index went down 39 % in 2008, the NCREIF Timberland Index achieved a 9 % return, or on the risk-adjusted basis, an excess return of 10 % (calculated using the estimated α series in 2008). In contrast, the portfolio value of publicly traded timber firms fell 39 % just like the market. However, it should be noted that most of those forestry firms do have non-timberland business, such as paper and lumber mills, which may be more sensitive to the overall economic conditions. A close examination of the three publicly traded timber REITs reveals that they were less affected by the gloomy market. Looking ahead, global economic crisis will last for some time, multiple factors will affect timberland returns, and the net effect on timberland properties has yet to be observed (Washburn 2008).

It should be noted that there have been some concerns about the data and method consistency of the NCREIF Timberland Index. As pointed out by Binkley et al. (1996), there are no standardized appraisal and valuation practice in forestry, so heterogeneity may exist in the data. In addition, due to lack of quarterly appraisals for many properties in the NCREIF Timberland Index, quarterly return series may be less useful than the annual ones. Finally, the NCREIF Timberland Index is a composite performance measure of a very large pool of commercial forestland properties acquired in the private market for investment purposes. Hence, individual investors should use cautions when interpreting the NCREIF Timberland Index.

Appendix 1: State Space Model with the Kalman Filter

The multivariate time-series model can be represented by the following state space form:

$$y_t = Z_t\alpha_t + \varepsilon_t, \quad \varepsilon_t \sim NID(0, H_t) \tag{34.10}$$

$$\alpha_{t+1} = T_t\alpha_t + R_t\eta_t, \quad \eta_t \sim NID(0, Q_t) \tag{34.11}$$

for $t = 1, \cdots, N$, where y_t is $p \times 1$ vector of observed values at time t, Z_t is a $p \times m$ matrix of variables, α_t is $m \times 1$ state vector, T_t is called the transition matrix of order $m \times m$, and R_t is an $m \times r$ selection matrix with $m \geq r$. The first equation is called the observation or measurement equation, and the second is called state equation. The parameters α_t, H_t, and Q_t in the system of equations can be estimated jointly by the maximum likelihood method with the recursive algorithm Kalman filter. The intention of filtering is to update the information of the system each time a new observation y_t is available, and the filtering equations are

$$\begin{aligned} v_t &= y_t - Z_t a_t, \\ F_t &= Z_t P_t Z_t^{'} + H_t, \\ K_t &= T_t P_t Z_t^{'} F_t^{-1}, \\ L_t &= T_t - K_t Z_t, \\ a_{t+1} &= T_t a_t + K_t v_t, \\ P_{t+1} &= T_t P_t L_t^{'} + R_t Q_t R_t^{'}, \end{aligned} \tag{34.12}$$

For $t = 1, \cdots, N$. The mean vector a_1 and the variance matrix P_1 are known for the initial state vector α_1 (Durbin and Koopman 2001; Harvey 1989).

Appendix 2: Heuristic Proof of Equation 34.6

In a pure exchange economy with identical consumers, a typical consumer wishes to maximize the expected sum of time-separable utilities

$$\begin{aligned} &\underset{C_t}{\text{Max}}\, E_t\left[\sum_{i=0}^{\infty}\beta^i U(C_{t+i})\right] \\ &s.t. \sum_{j=1}^{N} x_t^j p_t^j + C_t = W_t + \sum_{j=1}^{N} x_{t-1}^j\left(p_t^j + d_t^j\right) \end{aligned} \tag{34.13}$$

where x_t^j is the amount of security j purchased at time t, p_t^j is the price of security j at time t, W_t is the individual's endowed wealth at time t, C_t is the individual's consumption at time t, d_t^j is the dividend paid by security j at time t, and β is time discount. Express C_t in terms of x_t^j, and differentiate the objective function with respect to x_t^j, then we can get the following first-order condition:

$$E_t\left[U'(C_t)p_t^j\right] = E_t\left[\beta U'(C_{t+1})\left(p_{t+1}^j + d_{t+1}^j\right)\right] \tag{34.14}$$

for all j. After rearranging the terms, we can reach Eq. 34.6, where

$$\begin{aligned} M_t &= \frac{\beta U'(C_{t+1})}{U'(C_t)} \\ R_{t+1} &= \frac{p_{t+1}^j + d_{t+1}^j}{p_t^j} - 1. \end{aligned} \tag{34.15}$$

Appendix 3: NCREIF Timberland Index

The NCREIF Timberland Index has two components, the income return and the capital return. The income return is also known as EBITDDA return, which represents earnings before interest expenses, income taxes, depreciation, depletion, and amortization. The capital return is derived from land appreciation. The formulas to calculate these returns are

$$IR_t = \frac{EBITDDA_t}{MV_{t-1} + 0.5(CI_t - PS_t + PP_\mathrm{t} - EBITDDA_t)} \tag{34.16}$$

$$CR_t = \frac{MV_t - MV_{t-1} - CI_t + PS_t - PP_t}{MV_{t-1} + 0.5(CI_t - PS_t + PP_t - EBITDDA_t)} \tag{34.17}$$

where IR_t and CR_t are the income return and capital return, respectively; $EBITDDA_t$ equals the net operating revenue obtained from the tree farm (primarily from timber sales); CI_t equals the capitalized expenditures on the tree farm (e.g., forest regeneration and road construction); PS_t equals the net proceeds from sales of land from the tree farm; PP_t equals the gross costs of adding land to the tree farm; and MV_t equals the market value of the tree farm (Binkley et al. 2003).

Appendix 4: EViews Code for Estimating the CAPM and the Fama-French Three-Factor Model

Texts after the single quotation marks are notations.

```
' Specify the file location
CD "C:\Mei Bin\Publication\Handbook of FES"

' Create a workfile in EViews and read in quarterly data 1987Q1-2010Q4
Workfile Timberland_Jul2011 q 1987 2010
' 7 is the total number of series to be read in
Read(t = dat, s) quarterly.csv 7
' Group the 7 series
group quarterly NCREIF MktRf SMB HML RF port
' Estimate the CAPM for NCREIF, adjusted for autocorrelation
' LS means OLS. Equation given by dependent variable followed by a list of independent variables
' In excess returns
equation CAPM1.LS (NCREIF-RF) C MktRf AR(4)
' Estimate the Fama-French three-factor model for the NCREIF Timberland Index
equation FF31.LS (NCREIF-RF) C MktRf SMB HML AR(4)
```

```
' Estimate the CAPM for the portfolio, AR(4) term dropped
due to its insignificance.
  equation CAPM2.LS (port-RF) C MktRf
  ' Estimate the Fama-French three-factor model for the
portfolio
  equation FF32.LS (port-RF) C MktRf SMB HML

  ' State space estimation of time-varying parameters for
the CAPM
  ' Define a state space model for the NCREIF Timberland Index
  sspace KFcapm1
  ' Signal equation, the CAPM
  ' Define alpha to be time-varying (state variable)
  KFcapm1.append @signal (NCREIF-RF) = sv1 + c(1)*MktRf +
[var = exp(c(2))]
  ' State equation as a random walk
  KFcapm1.append @state sv1 = sv1(-1) + [var = exp(c(3))]
  ' Starting values for the state space model. Values come
the OLS estimation
  KFcapm1.append @param c(1) 0.01 c(2) 0 c(3) 0
  ' Maximum likelihood estimation
  KFcapm1.ml(showopts, m = 500, c = 0.0001, m)
  Show KFcapm1.output
  ' Save the time-varying alphas and its RMSEs
  KFcapm1.makestates(t = filt) CAPM1filt*
  KFcapm1.makestates(t = filtse) CAPM1filtse*

  ' Generate the graph of time-varying alphas with the 95 %
confidence intervals
  series CAPM1_a_bandplus = CAPM1filtsv1 + 2*CAPM1filtsesv1
  series CAPM1_a_bandminus = CAPM1filtsv1 - 2*CAPM1filtsesv1
  ' Group the series to be shown in a graph
  Group CAPM1_a_curves CAPM1filtsv1 CAPM1_a_bandplus
CAPM1_a_bandminus

  ' State space model for the portfolio of public forest
products firms
  sspace KFcapm2
  ' Define beta to be time-varying (state variable)
  KFcapm2.append @signal (port-RF) = c(1) + sv1*MktRf +
[var = exp(c(2))]
  KFcapm2.append @state sv1 = sv1(-1) + [var = exp(c(3))]
  KFcapm2.append @param c(1) 0.59 c(2) 4.3 c(3) -30
  KFcapm2.ml(showopts, m = 500, c = 0.0001, m)
```

```
Show KFcapm2.output

KFcapm2.makestates(t = filt) CAPM2filt*
KFcapm2.makestates(t = filtse) CAPM2filtse*
series CAPM2_b_bandplus = CAPM2filtsv1 + 2*CAPM2filtsesv1
series CAPM2_b_bandminus = CAPM2filtsv1 - 2*CAPM2filtsesv1

Group CAPM2_b_curves CAPM2filtsv1 CAPM2_b_bandplus
CAPM2_b_bandminus

sspace KFff1
' Time-varying alpha in the Fama-French three-factor
model
KFff1.append @signal (NCREIF-RF) = sv1 + c(1)*MktRf + c(2)
*SMB + c(3)*HML + [var = exp(c(4))]
KFff1.append @state sv1 = sv1(-1) + [var = exp(c(5))]
KFff1.append @param c(1) 0.02 c(2) -0.04 c(3) -0.05 c(4) 0
c(5) 0
KFff1.ml(showopts, m = 500, c = 0.0001, m)
Show KFff1.output

sspace KFff2
' Time-varying beta in the Fama-French three-factor model
KFff2.append @signal (port-RF) = c(1) + sv1*MktRf + c(2)
*SMB + c(3)*HML + [var = exp(c(4))]
KFff2.append @state sv1 = sv1(-1) + [var = exp(c(5))]
KFff2.append @param c(1) 0 c(2) 0 c(3) 3 c(4) 3 c(5) 0
KFff2.ml(showopts, m = 500, c = 0.0001, m)
Show KFff2.output
' Save the workfile
wfsave Timberland_Jul2011
```

Appendix 5: Steps for the SDF Approach Using Excel Solver

First, choose minimizing the standard deviation of the SDFs as the objective function.

Second, set the mean of the SDFs equal to a predetermined value. This is constraint No.1.

Third, for each basis asset (industry group) in the industry portfolio, add one constraint according to $\frac{1}{T}\sum_{t=1}^{T} M_t\left(1 + R_{i,t}\right) = 1$. That is, add five more constraints when using the five-industry portfolio, whereas add ten more constraints when using the ten-industry portfolio. Fourth, specify the solutions to be nonnegative and solve for the SDFs. Fifth, use the SDFs to price timberland returns according to Eq. 34.9. Repeat steps 1–5 with a different value as the given mean of the SDFs.

References

Ahn, D.-H., Conrad, J., & Dittmar, R. F. (2003). Risk adjustment and trading strategies. *Review of Financial Studies, 16*, 459–485.

Binkley, C. S., Raper, C. F., & Washburn, C. L. (1996). Institutional ownership of U.S. Timberland: History, rationale, and implications for forest management. *Journal of Forestry, 94*, 21–28.

Binkley, C. S., Washbrun, C. L., & Aronow, M. E. (2003). *The NCREIF timberland property index: Research notes 2003*. Boston: Hancock Timber Resource Group.

Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (1997). *The econometrics of financial markets*. Princeton: Princeton University Press.

Cascio, A. J., & Clutter, M. L. (2008). Risk and required return assessments of equity Timberland investments in the United States. *Forest Products Journal, 58*, 61–70.

Caulfield, J. P. (1998). Timberland return drivers and investing styles for an asset that has come of age. *Real Estate Finance, 14*, 65–78.

Chen, Z., & Knez, P. J. (1996). Portfolio performance measurement: Theory and applications. *The Review of Financial Studies, 9*, 511–555.

Conroy, R., & Miles, M. (1989). Commercial forestland in the pension portfolio: The biological beta. *Financial Analysts Journal, 45*, 46–54.

Doran, H. E., & Rambaldi, A. N. (1997). Applying linear time-varying constraints to econometric models: With an application to demand systems. *Journal of Econometrics, 79*, 83–95.

Durbin, J., & Koopman, S. J. (2001). *Time series analysis by state space methods*. New York: Oxford University Press.

Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*, 3–56.

Fortson, J. C. (1986). Factors affecting the discount rate for forestry investments. *Forest Products Journal, 36*, 67–72.

Haight, R. G., & Holmes, T. P. (1991). Stochastic price models and optimal tree cutting: Results for Loblolly pine. *Natural Resource Modeling, 5*, 424–443.

Hansen, L. P., & Jagannathan, R. (1991). Implications of security market data for models of dynamic economies. *Journal of Political Economy, 99*, 225–262.

Harris, T., Baldwin, S., & Siry, J. (2010). *United States Timberland markets 2000 to mid-2010: Transactions, values & market research*. Athens: Timber Mart-South.

Harvey, A. C. (1989). *Forecasting, structural time series models and the Kalman filter*. Cambridge, UK: Cambridge University Press.

Heikkinen, V. P. (2002). Co-integration of timber and financial markets-implications for portfolio selection. *Forest Science, 48*, 118–128.

Jensen, M. C. (1968). The performance of mutual funds in the period 1945–1964. *The Journal of Finance, 23*, 389–416.

Kenneth, R. French's website. http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

King, B. F. (1966). Market and industry factors in stock price behavior. *The Journal of Business, 39*, 139–190.

Liao, X., Zhang, Y., & Sun, C. (2009). Investments in timberland and softwood timber as parts of portfolio selection in the United States: A cointegration analysis and capital asset pricing model. *Forest Science, 55*, 471–479.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics, 47*, 13–37.

Lucas, R. E., Jr. (1978). Asset prices in an exchange economy. *Econometrica, 46*, 1429–1445.

Malaty, R., Toppinen, A., & Viitanen, J. (2007). Modeling and forecasting Finnish pine sawlog stumpage prices using alternative time-series methods. *Canadian Journal of Forest Research, 37*, 178–187.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance, 7*, 77–91.

Redmond, C. H., & Cubbage, F. W. (1988). Portfolio risk and returns from timber asset investments. *Land Economics, 64*, 325–337.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance, 19*, 425–442.

Song, H., Liu, X., & Romilly, P. (1996). A time varying parameter approach to the Chinese aggregate consumption function. *Economics of Planning, 29*, 185–203.

Sun, C. (2007). Variation of federal cost-share programs in the United States and the inducement effects on tree planting. *Journal of Forest Economics, 12*, 279–296.

Sun, C., & Zhang, D. (2001). Assessing the financial performance of forestry-related investment vehicles: Capital Asset Pricing Model vs. Arbitrage Pricing Theory. *American Journal of Agricultural Economics, 83*, 617–628.

Washburn, C. (2008). *Forestry investments in a slowing global economy*. London/Athens: International Forest Investment/Center for Forest Business.

Washburn, C. L., & Binkley, C. S. (1990). On the use of period-average stumpage prices to estimate forest asset pricing models. *Land Economics, 66*, 379–393.

Washburn, C. L., & Binkley, C. S. (1993). Do forest assets hedge inflation? *Land Economics, 69*, 215–224.

Zinkhan, F. C. (1988). Forestry projects, modern portfolio theory, and discount rate selection. *Southern Journal of Applied Forestry, 12*, 132–135.

Zinkhan, F. C. (2008). History of modern timberland investments. In *International forest investment: An executive education conference* (Munich, Germany), Center for Forest Business, Athens.

Zinkhan, F. C., & Cubbage, F. W. (2003). Financial analysis of timber investments. In E. O. Sills & K. L. Abt (Eds.), *Forests in a market economy*. Dordrecht: Kluwer.

Optimal Orthogonal Portfolios with Conditioning Information

35

Wayne E. Ferson and Andrew F. Siegel

Contents

Abstract

Optimal orthogonal portfolios are a central feature of tests of asset pricing models and are important in active portfolio management problems. The portfolios combine with a benchmark portfolio to form ex ante mean variance efficient portfolios. This paper derives and characterizes optimal orthogonal portfolios in the presence of conditioning information in the form of a set of lagged instruments. In this setting, studied by Hansen and Richard (1987), the conditioning information is used to optimize with respect to the unconditional moments. We present an

W.E. Ferson (✉)
University of Southern California, Los Angeles, CA, USA
e-mail: ferson@marshall.usc.edu; wayne.ferson@marshall.usc.edu

A.F. Siegel
University of Washington, Seattle, WA, USA
e-mail: asiegel@uw.edu; asiegel@u.washington.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_35,

empirical illustration of the properties of the optimal orthogonal portfolios. From an asset pricing perspective, a standard stock market index is far from efficient when portfolios trade based on lagged interest rates and dividend yields. From an active portfolio management perspective, the example shows that a strong tilt toward bonds improves the efficiency of equity portfolios.

The methodology in this paper includes regression and maximum likelihood parameter estimation, as well as method of moments estimation. We form maximum likelihood estimates of nonlinear functions as the functions evaluated at the maximum likelihood parameter estimates. Our analytical results also provide economic interpretation for test statistics like the Wald test or multivariate F test used in asset pricing research.

Keywords

Asset pricing tests • Conditioning information • Minimum-variance efficiency • Optimal portfolios • Predicting returns • Portfolio management • Stochastic discount factors • Generalized method of moments • Maximum likelihood • Parametric bootstrap • Sharpe ratios

35.1 Introduction

The optimal orthogonal portfolio, also known as the most mispriced portfolio or the active portfolio, is a central concept in asset pricing tests and in modern portfolio management. In asset pricing problems, it represents the difference between the performance of a benchmark portfolio and the maximum potential performance in a sample of assets (Jobson and Korkie 1982). In modern portfolio management, it shows how to actively tilt away from a given benchmark portfolio to achieve portfolio efficiency (Gibbons et al. 1989; Grinold and Kahn 1992).

Optimal orthogonal portfolios are studied by Roll (1980), MacKinlay (1995), Campbell et al. (1987), and others. However, these studies restrict the analysis to a setting where the portfolio weights are fixed over time. In contrast, studies in asset pricing use predetermined variables to model conditional expected returns, correlations, and volatility. Portfolio weights may be functions of the predetermined variables, and they will generally vary over time. Quantitative portfolio managers routinely use conditioning information in optimized portfolio strategies. Therefore, it is important to understand optimal orthogonal portfolios in a conditional setting.

This paper derives, characterizes, and illustrates optimal orthogonal portfolios in a conditional setting. The setting is one where the conditional means and variances of returns are time varying and optimal time-varying portfolio weights achieve unconditional mean variance efficiency with respect to the information, as described by Hansen and Richard (1987) and Ferson and Siegel (2001).[1] Ferson

[1]An alternative is to study conditional efficiency, where the weights minimize the conditional variance. This may be handled by simply reinterpreting the classical analysis.

and Siegel (2001) argue that this setting is interesting from the perspective of active portfolio management, where the client cannot observe the information that a portfolio manager may have. Ferson and Siegel (2003, 2009) show that this setting is also interesting from the perspective of testing asset pricing models, but they do not develop the optimal orthogonal portfolio.

We show that the optimal orthogonal portfolio has time-varying weights, and we derive the weights in a closed form. The portfolio weights for unconditionally efficient portfolios in the presence of conditioning information are derived by Ferson and Siegel (2001). They consider the case with no risk-free asset and the case with a fixed risk-free asset whose return is constant over time. We generalize these solutions to the case with a "conditionally" risk-free asset whose return is known at the beginning of the period and is thus included in the lagged conditioning information and may vary over time. We derive solutions for the optimal orthogonal portfolios with conditioning information, including cases where there is no risk-free asset, a constant risk-free rate, or a time-varying conditionally risk-free rate. We show that a "law of conservation of squared Sharpe ratios" holds, implying that the optimal orthogonal portfolio's squared unconditional Sharpe ratio is the difference between that of the benchmark portfolio and the maximum unconditional squared Sharpe ratio that is possible using the assets and conditioning information. Empirical examples illustrate the performance of the optimal orthogonal portfolios with conditioning information and the behavior of the portfolio weights.

Section 35.2 briefly reviews optimal orthogonal portfolios in the classical case with no conditioning information. Section 35.3 describes the setting for our analysis with conditioning information. Section 35.4 presents the main theoretical results, and Sect. 35.5 presents our empirical examples. Section 35.6 concludes. Appendix 1 includes the proofs of the main results, and Appendix 2 describes the methodology of our empirical examples in detail, including a general description of the parametric bootstrap.

35.2 Optimal Orthogonal Portfolios: The Classical Case

In the classical case, portfolio weights are fixed constants over time, and there is no conditioning information. Optimal orthogonal portfolios are tied to mean variance efficiency. Mean variance efficient portfolios maximize the expected return, given the variance of the return. Since Markowitz (1952) and Sharpe (1964), such portfolios have been at the core of financial economics.

The mean variance efficiency of a given portfolio can be described using a system of time-series regressions. If $r_t = R_t - \gamma_0$ is the vector of N *excess* returns at time t, measured in excess of a given risk-free or zero-beta return, γ_0, and $r_{pt} = R_{pt} - \gamma_0$ is the excess return on a benchmark portfolio, the regression system is

$$r_t = \alpha + \beta r_{pt} + u_t; \quad t = 1, \ldots, T, \tag{35.1}$$

where T is the number of time-series observations, β is the N-vector of regression slopes or betas, and α is the N-vector of intercepts or alphas. The tested portfolio r_{pt} is represented among the returns in r_t, so the covariance matrix of the residuals in Eq. 35.1 is singular (r_{pt} might be included explicitly or might be a fixed-weight portfolio of the assets in r_t). The portfolio r_p is minimum-variance efficient and has the given zero-beta return only if $\alpha = 0$.[2] The mean variance efficiency of a given portfolio is of normative investment interest, as an efficient portfolio maximizes a concave utility function defined solely over the mean and variance of the portfolio return, as would follow from normally distributed returns in a single-period model, for example. Equation 35.1 may be interpreted as referring to multiple factor portfolios, where r_p is a K-vector and β is an $N \times K$ matrix. Then, the benchmark portfolio is a linear combination of the K returns in the vector r_p (e.g., Shanken 1987; Gibbons et al. 1989).

Definition The *most mispriced (or optimal orthogonal) portfolio with respect to* r_p, *when there is no conditioning information, has excess return* $r_c = x'_c r$, *where the weights* x_c *satisfy.*

$$x_c = \underset{x}{\text{Arg Max}} \frac{(x'\alpha)^2}{\text{Var}(x'r)}. \tag{35.2}$$

It is clear from the definition in Eq. 35.2 why the portfolio is referred to as the most mispriced. The vector α captures the "mispricing" of the tested asset returns in Eq. 35.1 when evaluated using the benchmark r_p, and the portfolio x_c has the largest squared alpha relative to its variance. This interpretation also reveals why the portfolio is of central interest in active portfolio management. Given a benchmark portfolio r_p, an active portfolio manager places bets by deviating from the portfolio weights that define the benchmark. The manager is rewarded for bets that deliver higher returns and penalized for increasing the volatility. The portfolio in (2) describes the active bets that achieve the largest amount of extra return for the variance. Thus, the solution is also referred to as the *active* portfolio by Gibbons et al. (1989). (See Grinold and Kahn (1992) for an in-depth treatment of modern portfolio management.)

In the classical case of fixed portfolio weights, the solution to Eq. 35.2 is given by $x_c = \left(\underline{1}'V^{-1}\alpha\right)^{-1}V^{-1}\alpha$, where $\underline{1}$ is an N-vector of ones and $V = \text{Cov}(r)$, the covariance matrix of the returns. Using this solution, several well-known properties

[2]Note the distinction between *minimum variance efficient* portfolios, which minimize the variance for the given mean return, and *mean variance efficient*, which maximize the mean return given its variance. The latter set of portfolios is a subset of the former, typically depicted as the positively sloped portion of the minimum variance efficient boundary when graphed with mean return on the y-axis and standard deviation or variance of return on the x-axis. The portfolio r_p is mean variance efficient when $\alpha = 0$ and $\text{E}(r_p)$ exceeds the expected excess return of the global minimum variance portfolio.

of the optimal orthogonal portfolio follow.[3] For example, a combination of the portfolio r_c and the benchmark portfolio r_p is *optimal*, that is, minimum-variance efficient (Jobson and Korkie 1982). The portfolio is *orthogonal* to the benchmark portfolio in the sense that $\mathrm{Cov}(x_c' r, r_p) = 0$.

The optimal orthogonal portfolio is central for the interpretation of tests of portfolio efficiency. Classical test statistics for the hypothesis that $\alpha = 0$ in Eq. 35.1 can be written in terms of squared Sharpe ratios (e.g., Jobson and Korkie 1982). Consider the Wald statistic:

$$W = T\hat{\alpha}'[\mathrm{Cov}(\hat{\alpha})]^{-1}\hat{\alpha} = T\left(\frac{\hat{S}^2(R) - \hat{S}^2(R_p)}{1 + \hat{S}^2(R_p)}\right) \dot{\sim} \chi^2(N) \tag{35.3}$$

where $\hat{\alpha}$ is the OLS or maximum likelihood (ML) estimator of α (after removing r_p or another asset from the vector r to avoid singularity of the covariance matrix) and $\mathrm{Cov}(\hat{\alpha})$ is its asymptotic covariance matrix. Upper case R's refer to gross returns, and lower case r's refer to returns in excess of the zero-beta rate. The term $\hat{S}^2(R_p)$ is the sample value of the squared Sharpe ratio of R_p when the zero-beta rate is γ_0 so that $S^2(R_p) = [E(r_p)/\sigma(r_p)]^2$. The term $\hat{S}^2(R)$ is the sample value of the maximum squared Sharpe ratio that can be obtained by portfolios of the assets in R (including R_p):

$$S^2(R) = \max_x \left\{ \frac{[E(x'r)]^2}{\mathrm{Var}(x'r)} \right\}. \tag{35.4}$$

The Wald statistic has an asymptotic chi-squared distribution with N degrees of freedom, under the null hypothesis that R_p is efficient with the given zero-beta return. Scaled with a degrees of freedom adjustment, the statistic has an F distribution under normally distributed returns (Gibbons et al. 1989).

It can be shown that the squared Sharpe ratios can be decomposed using the optimal orthogonal portfolio as $S^2(R) = S^2(R_p) + S^2(R_c)$. A similar decomposition holds at the sample values. This decomposition, a "law of conservation of squared Sharpe ratios," is used by Jobson and Korkie (1982) to derive the second equality in (35.3). Since the Sharpe ratio is the slope of a line in the mean-standard deviation space, Eq. 35.3 suggests a graphical representation for the statistic in the sample mean-standard deviation space. It measures the distance between the sample mean-standard deviation frontier and the location of the tested portfolio, inside the frontier. This distance is proportional to the squared Sharpe ratio of the optimal orthogonal portfolio. Kandel (1984), Roll (1985), Gibbons et al. (1989), and Kandel and Stambaugh (1987, 1989) further develop this interpretation.

[3]See Roll (1980), Gibbons et al. (1989), MacKinlay (1995), and Campbell et al. (1987) for analyses of optimal orthogonal portfolios in the classical case with no conditioning information.

35.3 The Conditional Setting

We use conditioning information in a setting similar to that of Hansen and Richard (1987) and Ferson and Siegel (2001), where minimum-variance efficiency is defined in terms of the unconditional means and variances of the portfolios that result from the use of conditioning information. Ferson and Siegel (2009) refer to this as efficiency *with respect to the information*, Z. This setting has proven useful in asset pricing tests (Ferson and Siegel 2003, 2009; Bekaert and Liu 2004), in forming hedging portfolios (Ferson et al. 2006), and in portfolio management problems (Ahkbar et al. 2007; Chiang 2012). We study the optimal orthogonal portfolio in this setting. The distinction between mean variance efficiency and minimum-variance efficiency, as in the classical setting, applies in this setting as well.

Consider a portfolio of N assets with gross returns, R_{t+1}, where the weights that determine the portfolio at time t are functions of the information, Z_t. The gross return on such a portfolio with weight $x(Z_t)$ is $x'(Z_t)R_{t+1}$. The restrictions on the portfolio weight function are that the weights must sum to 1 (almost surely in Z_t), and that the unconditional expected value and second moments of the portfolio return are well defined. Consider now all portfolio returns that may be formed, for a given set of asset returns R_{t+1} and given conditioning information, Z_t, with well-defined first and second moments. This set determines a mean-standard deviation frontier, as shown by Hansen and Richard (1987). This frontier depicts the *unconditional* means versus the *unconditional* standard deviations of the portfolio returns. A portfolio is defined to be efficient with respect to the information Z_t if and only if it is on this mean-standard deviation frontier.

Ferson and Siegel (2001) derive solutions for efficient-with-respect-to-Z portfolios in closed form. They consider the case with no risk-free asset and the case with a fixed risk-free asset whose return is constant over time. In Theorem 1 of Appendix 1, we derive the solution for the case with a risk-free asset whose return is known at the beginning of the period and is thus included in the information Z and may vary over time. In this case the variation in the risk-free rate over time affects the unconditional variance of the portfolio return.

Ferson and Siegel (2001) argue that efficiency with respect to the information is especially relevant in a portfolio management context. It is reasonable to assume that the portfolio manager has more information about asset returns than the client. Assume that the client desires an unconditionally mean-variance efficient portfolio. The manager observes conditioning information that is relevant about future returns, and by conditioning on this information, he or she can expand the investor's opportunity set. The manager maximizes the investor's mean variance opportunity set by using his information to maximize the unconditional mean for a given unconditional variance. The efficient-with-respect-to-Z strategy is therefore the strategy that the investor would wish the portfolio manager to use.

Ferson and Siegel (2009) show how asset pricing theories make statements about portfolios that are efficient with respect to information Z and develop tests of the hypothesis that a portfolio is efficient with respect to Z. The optimal orthogonal portfolio with respect to information Z is a useful concept in these portfolio efficiency tests. We begin the analysis with a result from Hansen and Richard (1987).

Proposition 1 (Hansen and Richard 1987, Corollary 3.1.) Given N asset gross returns, R_{t+1}, a given portfolio with gross return $R_{p,t+1}$ is *minimum-variance efficient with respect to the information* Z_t if and only if Eq. 35.5 is satisfied (equivalently, there exists constants γ_0 and γ_1 such that Eq. 35.6 is satisfied) for all $x(Z_t)$ such that $x'(Z_t)\underline{1} = 1$ almost surely, where the relevant unconditional expectations exist and are finite:

$$\mathrm{Var}\left(R_{p,t+1}\right) \le \mathrm{Var}\left[x'(Z_t)R_{t+1}\right] \text{ if } E\left(R_{p,t+1}\right) = E\left[x'(Z_t)R_{t+1}\right] \tag{35.5}$$

$$E\left[x'(Z_t)R_{t+1}\right] = \gamma_0 + \gamma_1 \,\mathrm{Cov}\left[x'(Z_t)R_{t+1}, R_{p,t+1}\right]. \tag{35.6}$$

Equation 35.5 is the *definition* of efficiency with respect to Z. It states that $R_{p,\,t+1}$ is on the minimum-variance boundary formed by all possible portfolios that use the assets in R and the conditioning information. Equation 35.6 states that the familiar expected return-covariance relation from Fama (1973) and Roll (1977) must hold with respect to the efficient portfolio. In Eq. 35.6, the coefficients γ_0 and γ_1 are fixed scalars that do not depend on the functions $x(\cdot)$ or the realizations of Z_t.

35.4 The Main Results

The optimal orthogonal portfolio with conditioning information plays roles analogous to the classical setting with no conditioning information. Thus, for example, restricting the maximization in Eq. 35.4 to *fixed-weight* portfolios where x is a constant vector, we obtain efficiency in the classical case. In contrast, an efficient portfolio with respect to the information Z maximizes the squared Sharpe ratio over *all portfolio weight functions*, $x(Z)$. Maximizing over a larger set of weights expands the investment opportunity set and produces a larger maximum Sharpe ratio.

With conditioning information, the optimal orthogonal portfolio's weight function is time varying, and we derive this portfolio weight for three cases. First, with no risk-free asset in which case a fixed unconditional "zero-beta" rate γ_0 is arbitrarily chosen. By varying the zero-beta rate, the solutions can describe any point on the efficient-with-respect-to-Z frontier. Second, we consider a conditionally time-varying risk-free asset whose return $R_f = R_f(Z)$ is measureable and thus known as part of the information set Z so that $\mathrm{Var}[R_f(Z)|Z] = 0$, but which is unconditionally risky in the sense that $\mathrm{Var}[R_f(Z)] > 0$. Here, we again choose an arbitrary zero-beta rate γ_0 to describe the frontier. In the third case, there exists an unconditional risk-free asset with fixed return $R_f = \gamma_0$. In this case, the efficient-with-respect-to-Z frontier becomes a line passing through γ_0 (at risk zero) and the point representing the mean and standard deviation of a particular portfolio strategy's return.

We consider portfolios formed from the risky assets using weights $x_q = x_q(Z)$ where the weights must sum to 1 (for all Z) when there is no risk-free asset. This constraint is relaxed when there is a conditional or unconditional risk-free asset (where the implicit weight in the risk-free asset is then set at 1 minus the sum of the weights in the risky assets). When there is no risk-free asset, the portfolio return is $R_{q,t+1} = x_q'(Z_t)R_{t+1}$, and when there is a risk-free asset, the portfolio return is $R_{q,t+1} = R_f(Z_t) + x_q'(Z_t)\left[R_{t+1} - R_f(Z_t)\underline{1}\right]$ whether or not $R_f(Z)$ is constant. In either case, we denote the (unconditional) portfolio mean $\mu_q = E(R_q)$ and variance $\sigma_q^2 = \mathrm{Var}(R_q)$. When there exists a conditional time-varying risk-free asset, the portfolio takes advantage of the ability to adapt both the percentage invested in risky assets and their portfolio weights in response to the information Z. This may be interpreted as "market timing" and security selection, respectively. We define the optimal orthogonal portfolio with respect to a given benchmark portfolio P formed from the risky assets using (possibly) time-varying weights $x_p = x_p(Z)$.

Definition The *most mispriced (or optimal orthogonal) portfolio, R_c, with respect to the benchmark portfolio R_p and conditioning information Z*, with portfolio weight denoted $x_c(Z)$, uses the conditioning information to maximize α_c^2/σ_c^2 where the unconditional variance of R_c is σ_c^2, the unconditional mean is $\mu_c = E(R_c)$, the unconditional alpha of R_c with respect to R_p is $\alpha_c = \mu_c - [\gamma_0 + (\mu_p - \gamma_0)\sigma_{cp}/\sigma_p^2]$, the zero-beta rate is γ_0, and the unconditional covariance is $\sigma_{cp} = \mathrm{Cov}(R_c, R_p)$.

Proposition 2 The unique most mispriced (or optimal orthogonal) portfolio R_c with respect to a given benchmark portfolio R_p (with weights x_p and expected return μ_p), conditioning information Z and given zero-beta rate, γ_0, has the following portfolio weight in each of the following cases. If there is no risk-free rate, then the weights conditionally sum to 1 as defined by

$$x_c(Z) = A\left\{\frac{\Lambda\underline{1}}{\underline{1}'\Lambda\underline{1}} + [(c+1)\mu_s + b]\left(\Lambda - \frac{\Lambda\underline{1}\,\underline{1}'\Lambda}{\underline{1}'\Lambda\underline{1}}\right)\mu(Z)\right\} + Bx_p, \tag{35.7}$$

$R_f(Z)$ (which may be either constant or time varying, but if constant,[4] then we must choose $\gamma_0 = R_f$ along with any $\mu_s \neq R_f$), then the solution is

$$x_c(Z) = A\left[(c+1)\mu_s + b - R_f\right]Q\left[\mu(Z) - R_f\underline{1}\right] + Bx_p, \tag{35.8}$$

where

$$A = \frac{(\mu_s - \gamma_0)/\sigma_s^2}{(\mu_s - \gamma_0)/\sigma_s^2 - (\mu_p - \gamma_0)/\sigma_p^2}, \tag{35.9}$$

[4]Equation 35.11 cannot be used to determine μ_s when $R_f(Z)$ is almost surely constant due to division by zero, and, in this case, every choice $\mu_s \neq R_f$ uses the same (rescaled) portfolio of risky assets $Q\left[\mu(Z) - R_f\underline{1}\right]$ in the formation of an efficient portfolio $x_s(Z)$.

$$B = -\frac{(\mu_p - \gamma_0)/\sigma_p^2}{(\mu_s - \gamma_0)/\sigma_s^2 - (\mu_p - \gamma_0)/\sigma_p^2}, \tag{35.10}$$

$$\mu_s = -(a + b\gamma_0)/(b + c\gamma_0), \tag{35.11}$$

$$\sigma_s^2 = a + 2b\mu_s + c\mu_s^2, \tag{35.12}$$

$$\begin{aligned} Q = Q(Z) &\equiv \left\{ E\left[\left(R - R_f\underline{1}\right)\left(R - R_f\underline{1}\right)' \middle| Z \right] \right\}^{-1} \\ &= \left\{ \left[\mu(Z) - R_f\underline{1}\right]\left[\mu(Z) - R_f\underline{1}\right]' + \Sigma_\varepsilon(Z) \right\}^{-1}, \end{aligned} \tag{35.13}$$

$$\Lambda = \Lambda(Z) \equiv \left\{ E\left[RR' \middle| Z \right] \right\}^{-1} = \left[\mu(Z)\mu'(Z) + \Sigma_\varepsilon(Z) \right]^{-1}. \tag{35.14}$$

The constants a, b, and c are defined in Appendix 1 in Theorem 1 (when there exists a risk-free rate) and in Theorem 2 (when there is no risk-free rate).

Proof: See Appendix 1.

The term μ_s represents the unconditional expected return of the efficient-with-respect-to-Z portfolio, R_s, that maximizes the squared unconditional Sharpe ratio in Eq. 35.4 over all portfolio weight functions with respect to the given value of γ_0. When there exists a risk-free rate $R_f(Z)$ (that may be time varying because its value is included in the information set Z at the beginning of the period), the conditional mean variance boundary, the tangency intercept $R_f(Z)$, and thus the location of the conditionally mean variance efficient portfolio may vary over time as $R_f(Z)$ varies. When there is no risk-free rate, or when the risk-free rate is time varying, we use the parameter γ_0 to determine a fixed location on the (curved) unconditionally efficient-with-respect-to-Z boundary; however, when there is a fixed risk-free rate, this boundary is a degenerate hyperbola, and every portfolio on the upper line is efficient. We next show that the optimal orthogonal portfolio R_c can be formed by combining the benchmark portfolio R_p with the efficient-with-respect-to-Z portfolio R_s, and from this result, it then follows that R_p and R_c can be combined to produce the efficient-with-respect-to-Z portfolio R_s.

Proposition 3 The most mispriced or optimal orthogonal portfolio R_c may be found as a fixed linear combination of the benchmark portfolio R_p and the efficient-with-respect-to-Z portfolio, R_s (that maximizes the squared Sharpe ratio for the given zero-beta rate, γ_0), as follows:

$$R_c = \frac{\left[(\mu_s - \gamma_0)/\sigma_s^2\right]R_s - \left[\left(\mu_p - \gamma_0\right)/\sigma_p^2\right]R_p}{\left[(\mu_s - \gamma_0)/\sigma_s^2\right] - \left[\left(\mu_p - \gamma_0\right)/\sigma_p^2\right]} \tag{35.15}$$

Or

$$R_c = \frac{\sigma_p^2 R_s - \sigma_{ps} R_p}{\sigma_p^2 - \sigma_{ps}} = AR_s + BR_p \text{ with } A + B = 1, \tag{35.16}$$

where we assume that R_s and R_p are not perfectly correlated and that $\sigma_p^2 \neq \sigma_{sp}$.

Proof: See Appendix 1.

Propositions 2 and 3 extend the concept of the active or optimal orthogonal portfolio to the setting of efficiency with respect to conditioning information. Given a benchmark portfolio R_p, its optimal orthogonal portfolio with respect to Z shows how to tilt away from the benchmark weights to obtain efficiency with respect to Z. The portfolio R_c has weights that depend on Z. Thus, the optimal tilt away from a benchmark uses the manager's information Z in a dynamic way.

Equation 35.16 shows how the optimal orthogonal portfolio R_c can be formed by combining an efficient-with-respect-to-Z portfolio R_s with R_p. The portfolio R_c is the regression error of R_s projected on R_p, normalized so that the weights sum to 1, as can be seen by solving Eq. 35.16 for R_s. Thus, the portfolio R_c is uncorrelated with R_p.

We defined the most mispriced portfolio with conditioning information as maximizing the squared alpha relative to the unconditional variance of R_c. Since R_c is orthogonal to R_p, its residual variance in regression (Eq. 35.1) is the same as its total variance. Thus, we can think of the optimal orthogonal portfolio as maximizing alpha given its residual variance among all orthogonal portfolios.

Given a benchmark portfolio with return R_p, the optimal orthogonal portfolio with conditioning information is useful for active portfolio management. It might seem natural for a manager with information Z to simply reinterpret the classical analysis, where all the moments are the conditional moments given Z. This, in fact, is the interpretation that much of the literature on active portfolio management has used (e.g., Jorion 2003; Roll 1982). This approach produces *conditionally* mean variance efficient portfolios given Z. However, as shown by Dybvig and Ross (1985), a conditionally efficient portfolio is likely to be seen as inefficient from the (unconditional) perspective of a client without access to the information Z. The optimal orthogonal portfolio describes the active portfolio bets that deliver optimal performance from the client's perspective.

Let $S^2(R)$ be the maximum squared Sharpe ratio obtained by the efficient-with-respect-to-Z portfolio and let R_c be the optimal orthogonal portfolio with respect to R_p and information Z.

Proposition 4 *Law of conservation of squared Sharpe ratios*. For a given zero-beta or risk-free rate, if $S_s^2 = S^2(R)$ is the maximum squared Sharpe ratio obtained by portfolios $x(Z)$ and R_c is the optimal orthogonal portfolio with respect to R_p and information Z at that zero-beta rate, then $S_s^2 = S_p^2 + S_c^2$, where S_i^2 denotes the squared Sharpe ratio of portfolio i.

Proof: See Appendix 1.

Proposition 4 shows that if we test the hypothesis that a portfolio is efficient with respect to Z using versions of the test statistic in Eq. 35.3, as developed in Ferson and Siegel (2009), then the role of the optimal orthogonal portfolio with information Z is analogous to the role of the optimal orthogonal portfolio in the case of the classical test statistics. The Sharpe ratio of the optimal orthogonal portfolio with conditioning information indicates how far the tested benchmark portfolio is from the efficient-with-respect-to-Z frontier. The squared Sharpe ratio of the optimal orthogonal portfolio is the numerator of the test statistic. This numerator and thus the test statistic is zero only if the tested portfolio is efficient in the sample, and it grows larger as the tested portfolio is further from efficiency.

35.5 Empirical Examples

We present empirical examples using data on portfolios of common stocks, where the firms are grouped according to conventional criteria. We use the returns of common stocks sorted according to market capitalization and book-to-market ratios, focusing on value-weighted decile portfolios of small capitalization stocks, value stocks (high book/market), and growth stocks (low book/market), as provided on Ken French's website. We also include a long-term government bond return, splicing the Ibbotson Associates 20-year US government bond return series for 1931–1971, with the CRSP greater than 120 month US government bond return after 1971. The market portfolio, measured as the CRSP value-weighted stock return index, is the benchmark or tested portfolio, R_p. The risk-free return is the return from rolling over 1-month Treasury bills from CRSP. We use its sample average, 3.8 % per year, as the fixed zero-beta rate in all of the examples. As conditioning information in Z, its return is lagged by 1 year. All of the returns are discretely compounded annual returns, and the sample period is 1931–2007.

The lagged instruments, Z, are the lagged Treasury return and the log of the market price/dividend ratio at the end of the previous year. In calculating the price/dividend ratio, the stock price is the real price of the S & P 500 Index, and the dividend is the real dividends accruing to the index over the past year. These data are from Robert Shiller's website.

We treat the risk-free asset in three distinctly different ways to highlight the three different versions of our solutions for the optimal orthogonal portfolios. In the first case, the risk-free rate is assumed to be a fixed constant. Here, we do not include the Treasury return as a lagged instrument, and we set $\gamma_0 = 3.8$ % to be the average Treasury return during the sample. The target mean μ_s of the efficient-with-respect-to-Z portfolio is set equal to the sample mean return of the market index in this case, which determines the amount of leverage the portfolio uses at the fixed risk-free rate.

In the second case, no risk-free asset exists. Here, we again set $\gamma_0 = 3.8\ \%$ to pick a point on the mean variance boundary, and we do not allow the portfolio to take a position in a risk-free asset. We do allow the lagged Treasury return as conditioning information in Z, which highlights the information in lagged Treasury returns about the future risky asset returns. In the third case, there is a conditionally risk-free return that is contained in Z. Here, we use the lagged Treasury return in the conditioning information, and we allow the portfolio to trade the subsequent Treasury return in addition to the other risky assets. The subsequent Treasury return is not really known *ex ante* as the formula assumes, but its correlation with the lagged return is 0.92 during our sample, and this allows us to implement the time-varying risk-free rate example in a somewhat realistic way. (In reality, there is no *ex ante* risk-free asset given the importance of inflation risk.) In practical terms, this example highlights the effects of "market timing," or varying the amount of "cash" in the portfolio, in addition to varying the allocation among the risky assets.

Table 35.1 summarizes the results. The rows show results for the benchmark index (Market), three equity portfolios, and the government bond return. The CAPM α refers to the intercept in the regression (1), of the portfolio returns in excess of the Treasury bill returns, on the excess return of the market index. The small stock portfolio has the largest alpha, at 3.95 % per year, while the growth stock portfolio has a negative alpha. The symbol σ_u refers to the standard deviation of the regression residuals. The small stock portfolio has the largest σ_u or nonmarket risk, at more than 25 % per year.

The bottom four rows of Panel A summarize the optimal orthogonal portfolios when the market index is the benchmark. The fixed-weight portfolio R_c uses no conditioning information. Its alpha is larger than any of the separate assets, at 4.66 % per year, and its residual standard deviation is also relatively large, at 16.5 % per year. Since the portfolio is orthogonal to the market index, its residual standard deviation is the same as its total standard deviation, or volatility of return. The ratio of the alpha to the residual volatility is known as the appraisal ratio (Treynor and Black 1973) or the information ratio (Grinold and Kahn 1992). Optimal orthogonal portfolios try to maximize the square of this ratio. The fixed-weight portfolio R_c delivers an information ratio of 0.282, substantially larger than those of the small stock or value portfolios, at 0.155 and 0.177, respectively, and also larger than the bond portfolio, which has an information ratio of 0.183 by virtue of its relatively small volatility.

Table 35.1 summarizes performance statistics for the optimal orthogonal portfolios with conditioning information. There are three versions with (1) a fixed risk-free rate, (2) no risk-free rate, and (3) a time-varying conditional risk-free rate. The information ratios in all three cases are larger than those of the individual assets or the fixed-weight active portfolio, which illustrates the potential value of using conditioning information explicitly in a portfolio management context (see also, Chiang 2009). The improvement over fixed weights is modest for the case with no risk-free asset. However, the information ratio is about three times as large, at 0.955, in the example with a time-varying risk-free rate. This illustrates the

Table 35.1 Optimal orthogonal portfolios

Panel A: summary statistics

					Average active portfolio weights Avg $[X_c(Z)]$		
Asset	CAPM α	σ_u	Information Ratio	Fixed R_c	Fixed R_f	No R_f	Varying R_f
Market index	0	0	0.000	−1.260	0.179	−0.210	0.308
Small stocks	3.95	25.5	0.155	0.353	−0.262	0.133	0.097
Value stocks	3.14	17.8	0.177	0.345	0.137	0.079	0.142
Growth stocks	−1.00	8.3	−0.121	0.155	0.028	0.098	0.020
Bonds	1.66	9.1	0.183	1.410	0.918	0.900	0.432
Fixed R_c	4.66	16.5	0.282	1.000			
$X_c(Z)$ portfolios:							
Fixed R_f	8.66	25.5	0.340				
No R_f	3.12	10.5	0.297				
Varying R_f	8.77	9.2	0.955				

Panel B: squared Sharpe ratios

	$S^2(R_p)$	$S^2(R_c)$	Sum
Fixed R_c	0.189	0.079	0.267
$X_c(Z)$ portfolios:			
Fixed R_f	0.189	0.115	0.304
No R_f	0.189	0.088	0.277
Varying R_f	0.182	0.909	1.191

Annual returns on portfolios of common stocks and long-term US government bonds cover the 1931–2007 period. The CAPM α refers to the intercept in the regression Eq. 35.1, of the portfolio return in excess of a zero-beta parameter, on that of the broad value-weighted stock market index (market index). The symbol σ_u refers to the standard deviation of the regression residuals. The fixed portfolio, R_c, ignores the conditioning information. The alphas and residual standard deviations are annual percentage units. The information ratio is the ratio $\alpha/\ \sigma_u$

potential usefulness of interest rate information and the ability to hold "cash" as a function of market conditions, in a portfolio management context.

The averages over time of the optimal orthogonal portfolios' weights on the risky assets are shown in the right-hand columns of Panel A. The weights are normalized to sum to 1.0. The weights x_c of the optimal orthogonal portfolios combine with the benchmark (whose weights, x_p, are 100 % in the market index) to determine an efficient portfolio. The overall efficient portfolio weights x_s therefore vary over time and depend on how x_c and x_p are combined, which is determined by the coefficients A and B in Eqs. 35.9 and 35.10. The estimated weights are as follows. In the fixed risk-free rate case, $x_s = 0.60\,x_c + 0.40\,x_p$. In the no risk-free rate case, $x_s = 0.23\ x_c + 0.77\ x_p$. In the time-varying risk-free rate case, $x_s = 0.44\ x_c + 0.56\ x_p$. Thus, the efficient portfolio is formed as a convex combination of the benchmark and the active portfolio, with reasonable weights in each case.

The four right-hand columns of Panel A show that two of the active portfolios take short positions in the market benchmark, indicating an optimal tilt away from the market index. The fixed-weight portfolio R_c takes an extreme short position of -126 %, while on average the portfolio using the conditioning information but no risk-free asset takes a position of -21 % in the market index. These short positions finance large long positions in the US government bond and also long positions (in most cases) in small stocks, value stocks, and growth stocks. It is interesting that all the portfolios tilt positively, although by small amounts, into growth stocks even though growth stocks have negative CAPM alphas. This occurs because of the correlations among the asset classes.

All the versions of the optimal orthogonal portfolio suggest strong tilts into government bonds. The bond tilt is the most extreme for the fixed-weight solution, at 141 %, and is relatively modest, at 43.2 %, for the portfolio assuming a time-varying conditional risk-free rate. This makes sense, as that portfolio can hold short-term Treasuries in addition to government bonds. The large weights in bonds reflect various features of the data and the value of the zero-beta rate. With larger values for the zero-beta rate, provided that the rate remains below the mean of the global minimum-variance portfolio of risky assets, the optimal orthogonal portfolios become more aggressive as the target expected return of the efficient-with-respect-to-Z portfolio increases.

The squared Sharpe ratios in Panel B of Table 35.1 indicate how far the stock market index is from efficiency. The squared Sharpe ratio for the market is 0.189, measured relative to the fixed zero-beta rate of 3.8 %. The market portfolio's squared Sharpe ratio is slightly smaller, at 0.182, for returns measured in excess of a time-varying risk-free rate. This is because the negative covariance between the risk-free rate and stock returns increases the variance of the excess return. For the fixed-weight orthogonal portfolio R_c, the squared Sharpe ratio is 0.079, and for the portfolios using conditioning information, it varies between 0.088 and 0.909.

According to the *law of conservation of squared Sharpe ratios* in Proposition 4, the sum of the index and optimal orthogonal portfolios' squared Sharpe ratios is the squared slope of the tangency from the given zero-beta rate to the relevant

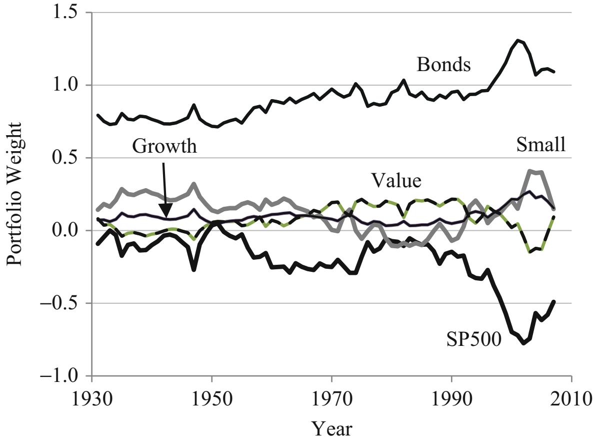


Fig. 35.1 Optimal orthogonal portfolio weights: the no risk-free rate case

mean variance frontier. The sum is 0.267 when no conditioning information is used and is 0.304–1.091 when the information is used. With a fixed risk-free rate, we condition only on the lagged dividend yield, which is a relatively weak predictor for stock returns (see Ferson et al. 2003). However, in the case with no risk-free rate, the portfolio is not allowed to hold short-term Treasuries, which substantially weakens the performance. In the time-varying risk-free rate case, the portfolio strategy is allowed to "market time" by holding short-term Treasuries. In the other two cases, we use the information in the lagged Treasury rate, which is a relatively strong predictor, and the efficient-with-respect-to-Z boundary is far above the mean variance boundary that ignores the conditioning information. (Ferson and Siegel 2009, present an analysis of the statistical significance of differences like these.)

Table 35.1 suggests that the portfolio weights of the fixed-weight R_c portfolio are extreme and would not be realistic in practical portfolio management settings. This reflects well-known issues with mean classical variance optimal solutions in practice (e.g., Michaud 1989; Siegel and Woodgate 2007). To obtain practical portfolio weights in the classical mean variance problem, it is generally necessary to constrain the weights (e.g., Frost and Savarino 1988) or shrink them toward a benchmark e.g., Jorion (2003) or Jagannathan and Ma (2003). In this context, note that the optimal orthogonal portfolios using Z take less extreme positions on average than the fixed-weight solution, yet still are able to generate larger information ratios.

Figures 35.1 and 35.2 present time-series plots of the weights for the optimal orthogonal portfolios using Z. Like the fixed-weight R_c case, the weights in the case with a fixed risk-free rate assumption appear too volatile and noisy to be of practical interest, likely reflecting the poor predictive ability based solely on the dividend yield. We do not plot them here to save space. The weights in the other two examples are shown in Figs. 35.1 and 35.2. They generally vary relatively smoothly over time, suggesting that they would not involve prohibitive trading costs in practice.

Figure 35.1 depicts the weights for the optimal orthogonal portfolio in the case with no risk-free asset. The market index weights are negative through much of the

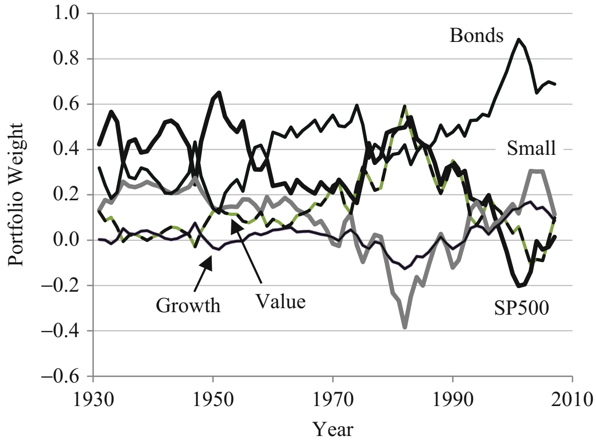


Fig. 35.2 Optimal orthogonal portfolio weights: the time-varying risk-free rate case

sample and near −25 %, indicating that the overall efficient-with-respect-to-Z portfolio keeps about 71 % of its money in the market index for much of the sample, computed as (0.23)(−0.25) + 0.77 using the estimated weights for the no risk-free rate case, $x_s = 0.23\ x_c + 0.77\ x_p$, as reported earlier. Starting in the mid-1990s, the market weights decrease to around −50 % for the optimal orthogonal portfolio weight and 66 % for the overall efficient-with-respect-to-Z portfolio. The weight of this optimal orthogonal portfolio in small stocks is positive for much of the sample, but turns slightly negative in the early 1970s, then positive again in the early 1990s. The strategy shorts value stocks in the 1930s and 1940s, but holds positive positions through most of the rest of the sample. The government bond gets the largest weight, starting near 80 % and growing sharply in the 1990s to near 100 % in the latter parts of the sample.

Figure 35.2 depicts the weights for the optimal orthogonal portfolio in the case with a time-varying, conditional risk-free rate. This strategy keeps positive weights in the market index until 1999, then it holds small short positions for most of the rest of the sample. The weight in small stocks is positive for much of the sample, turning slightly negative in the early 1970s and then positive again in the early 1990s. The strategy holds value stocks long until the early 2000s and shorts growth stocks during much of the 1980s. The government bond again gets the largest weight, starting at 32 % and growing to almost 70 % at the end of the sample.

35.6 Conclusions

This chapter derives, characterizes, and illustrates optimal orthogonal portfolios in the presence of conditioning information in the form of a set of lagged instruments. Optimal orthogonal portfolios combine with a benchmark portfolio to form mean variance efficient portfolios. We generalize previously published solutions for optimal orthogonal portfolios with information to include the case of a time-varying, but conditionally known, risk-free asset return.

In a portfolio management context, it is reasonable to assume that the portfolio manager has more information about asset returns than the client. The manager observes information about future returns and, by conditioning on this information, can expand the investor's opportunity set. Starting from a given benchmark portfolio, the optimal orthogonal portfolio with conditioning information describes the active portfolio bets that, when combined with the benchmark, deliver mean variance optimal performance from the uninformed client's perspective.

We present empirical examples using a broad stock market index as the benchmark and portfolios featuring small capitalization, value and growth stocks, and a long-term US government bond return. We examine three versions of the optimal orthogonal portfolio with (1) a fixed risk-free rate, (2) no risk-free rate, and (3) a time-varying conditional risk-free rate. The optimal orthogonal portfolios with conditioning information have larger information ratios than the orthogonal portfolio that does not use the conditioning information. At the same time, they take less extreme positions than the fixed-weight solution.

From an asset pricing perspective, a standard stock market index is far from efficient when portfolios trade based on lagged interest rates and dividend yields. From an active portfolio management perspective, the example shows that a strong tilt toward bonds improves the efficiency of equity portfolios. Our results should be useful in future asset pricing and portfolio management applications.

Appendix 1: Theorems and Proofs

Efficient Portfolio Solutions

Portfolio weights for efficient portfolios in the presence of conditioning information are derived by Ferson and Siegel (2001). They consider the case with no risk-free asset and the case with a fixed risk-free asset whose return is constant over time. In Theorem 1, we generalize to consider the case with a risk-free asset whose return is known at the beginning of the period, and thus is included in the information Z, and may vary over time. We then, in Theorem 2, reproduce the case with no risk-free asset from Ferson and Siegel (2001) for future reference.

Consider N risky assets with returns R. In $N \times 1$ column vector notation, we have

$$R = \mu(Z) + \varepsilon.$$

The noise term ε is assumed to have conditional mean zero given Z and nonsingular conditional covariance matrix $\Sigma_\varepsilon(Z)$. The conditional expected return vector is $\mu(Z) = E(R|Z)$. Let the $1 \times N$ row vector $x'(Z) = (x_1(Z), \ldots, x_N(Z))$ denote the portfolio share invested in each of the N risky assets, investing (or borrowing) at the risk-free rate the amount $1 - x'(Z)\underline{1}$, where $\underline{1} \equiv (1, \ldots, 1)'$ denotes the column vector of ones. We allow for a conditional risk-free asset returning $R_f = R_f(Z)$.

The return on the portfolio is $R_s = R_f + x'(Z)\left(R - R_f\underline{1}\right)$, with unconditional expectation and variance as follows:

$$\mu_s = E\left(R_f\right) + E\left\{x'(Z)\left[\mu(Z) - R_f\underline{1}\right]\right\},$$
$$\sigma_s^2 = E\left(R_s^2\right) - \mu_s^2 = E\left[E\left(R_s^2|Z\right)\right] - \mu_s^2,$$
$$\sigma_s^2 = E\left(R_f^2\right) + E\left[x'(Z)Q^{-1}x(Z)\right] + 2E\left\{R_f x'(Z)\left[\mu(Z) - R_f\underline{1}\right]\right\} - \mu_s^2, \tag{35.17}$$

where we have defined the $N \times N$ matrix

$$Q = Q(Z) \equiv \left\{E\left[\left(R - R_f\underline{1}\right)\left(R - R_f\underline{1}\right)'\middle|Z\right]\right\}^{-1}$$
$$= \left\{\left[\mu(Z) - R_f\underline{1}\right]\left[\mu(Z) - R_f\underline{1}\right]' + \Sigma_\varepsilon(Z)\right\}^{-1}.$$

Also, define the constants:

$$\zeta \equiv E\left\{\left[\mu(Z) - R_f\underline{1}\right]' Q\left[\mu(Z) - R_f\underline{1}\right]\right\},$$

$$\varphi \equiv E\left\{R_f\left[\mu(Z) - R_f\underline{1}\right]' Q\left[\mu(Z) - R_f\underline{1}\right]\right\},$$

and

$$\psi \equiv E\left\{R_f^2\left[\mu(Z) - R_f\underline{1}\right]' Q\left[\mu(Z) - R_f\underline{1}\right]\right\}.$$

Theorem 1 Given a target unconditional expected return μ_s, N risky assets, instruments Z, and a conditional risk-free asset with rate $R_f = R_f(Z)$ that may vary over time, the unique portfolio having minimum unconditional variance is determined by the weights

$$\begin{aligned} x_s(Z) &= \left(\frac{\mu_s - E\left(R_f\right) + \varphi}{\zeta} - R_f\right)Q\left[\mu(Z) - R_f\underline{1}\right] \\ &= \left[(c+1)\mu_s + b - R_f\right]Q\left[\mu(Z) - R_f\underline{1}\right], \end{aligned} \tag{35.18}$$

and the optimal portfolio variance is

$$\sigma_s^2 = a + 2b\mu_s + c\mu_s^2$$

where $a = E\left(R_f^2\right) + \frac{\left[E\left(R_f\right) - \varphi\right]^2}{\zeta} - \psi$, $b = \frac{\varphi - E\left(R_f\right)}{\zeta}$, and $c = \frac{1}{\zeta} - 1$. When the risk-free asset return is constant, then these formulas simplify to Theorem 2 of Ferson and Siegel (2001) with

$$x_s(Z) = \frac{\mu_s - R_f}{\zeta} Q\left[\mu(Z) - R_f \underline{1}\right]$$

and with optimal portfolio variance $\sigma_s^2 = \frac{1-\zeta}{\zeta}\left(\mu_s - R_f\right)^2$.

Proof Our objective is to minimize, over the choice of $x_s(Z)$, the portfolio variance $\mathrm{Var}(R_s)$ subject to $E(R_s) = \mu_s$, where $R_s = R_f + x'(Z)\left(R - R_f\underline{1}\right)$ and the variance is given by Eq. 35.17. We form the Lagrangian:

$$L[x(Z)] = E\left[x'(Z)Q^{-1}x(Z)\right] + 2E\left\{R_f x'(Z)\left[\mu(Z) - R_f\underline{1}\right]\right\}$$
$$+2\lambda E\left\{\mu_s - R_f - x'(Z)\left[\mu(Z) - R_f\underline{1}\right]\right\}$$

and proceed using a perturbation argument. Let $q(Z) = x(Z) + dy(Z)$, where $x(Z)$ is the conjectured optimal solution, $y(Z)$ is any regular function of Z, and d is a scalar. Optimality of $x(Z)$ follows when the partial derivative of $L[q(Z)]$ with respect to d is identically zero when evaluated at $d = 0$. Thus,

$$0 = E\left(y'(Z)\left\{Q^{-1}x(Z) + \left(R_f - \lambda\right)\left[\mu(Z) - R_f\underline{1}\right]\right\}\right)$$

for all functions $y(Z)$, which implies that $Q^{-1}x(Z) + \left(R_f - \lambda\right)\left[\mu(Z) - R_f\underline{1}\right] = 0$ almost surely in Z. Solve this expression for x(Z) to obtain Eq. 35.18, where the Lagrange multiplier λ is evaluated by solving for the target mean, μ_s. The expression for the optimal portfolio variance follows by substituting the optimal weight function into Eq. 35.17. Formulas for fixed R_f then follow directly. QED.

When the risk-free asset's return is time varying and contained in the information set Z at the beginning of the portfolio formation period, the *conditional* mean variance efficient boundary varies over time with the value of $R_f(Z)$ along with the conditional asset means and covariances . In this case, a zero-beta parameter, γ_0, may be chosen to fix a point on the *unconditionally* efficient-with-respect-to-Z boundary. The choice of the zero-beta parameter corresponds to the choice of a target unconditional expected return μ_s. For a given value of γ_0, the target mean maximizes the squared Sharpe ratio $(\mu_s - \gamma_0)^2/\sigma_s^2$ along the mean variance boundary, which implies $\mu_s = -(a + b\gamma_0)/(b + c\gamma_0)$.

When there is a risk-free asset that is constant over time, the unconditionally efficient-with-respect-to-Z boundary is linear (a degenerate hyperbola) and reaches the risk-free asset at zero risk. In this case, we use $\gamma_0 = R_f$ and can obtain any μ_s larger or smaller than R_f, levering the efficient portfolio up or down with positions in the risk-free asset.

When there is no risk-free asset, we define portfolio s by letting $x' = x'(Z) = [x_1(Z), \ldots, x_N(Z)]$ denote the shares invested in each of the N risky assets, with the constraint that the weights sum to 1.0 almost surely in Z. The return on this portfolio, $R_s = x'(Z)R$, has expectation and variance as follows:

$$\mu_s = E\left[x'(Z)\mu(Z)\right],$$

$$\sigma_s^2 = E\left\{x'(Z)\Lambda^{-1}x(Z)\right\} - \mu_s^2,$$

where we have defined the $N \times N$ matrix

$$\Lambda = \Lambda(Z) \equiv \left\{E\left[RR'|Z\right]\right\}^{-1} = \left[\mu(Z)\mu'(Z) + \Sigma_{\varepsilon}(Z)\right]^{-1}.$$

Also, define the constants:

$$\delta_1 = E\left(\frac{1}{\underline{1}'\Lambda\underline{1}}\right),$$

$$\delta_2 = E\left(\frac{\underline{1}'\Lambda\mu(Z)}{\underline{1}'\Lambda\underline{1}}\right),$$

and

$$\delta_3 = E\left[\mu'(Z)\left(\Lambda - \frac{\Lambda\underline{1}\,\underline{1}'\Lambda}{\underline{1}'\Lambda\underline{1}}\right)\mu(Z)\right].$$

Theorem 2 (Ferson and Siegel 2001, Theorem 3) Given N risky assets and no risk-free asset, the unique portfolio having minimum unconditional variance and unconditional expected return μ_s is determined by the weights:

$$\begin{aligned} x_s'(Z) &= \frac{\underline{1}'\Lambda}{\underline{1}'\Lambda\underline{1}} + \frac{\mu_s - \delta_2}{\delta_3}\mu'(Z)\left(\Lambda - \frac{\Lambda\underline{1}\,\underline{1}'\Lambda}{\underline{1}'\Lambda\underline{1}}\right) \\ &= \frac{\underline{1}'\Lambda}{\underline{1}'\Lambda\underline{1}} + [(c+1)\mu_s + b]\mu'(Z)\left(\Lambda - \frac{\Lambda\underline{1}\,\underline{1}'\Lambda}{\underline{1}'\Lambda\underline{1}}\right), \end{aligned} \tag{35.19}$$

and the optimal portfolio variance is

$$\sigma_s^2 = a + 2b\mu_s + c\mu_s^2,$$

where $a = \delta_1 + \delta_2^2/\delta_3$, $b = -\delta_2/\delta_3$, and $c = (1 - \delta_3)/\delta_3$.

The efficient-with-respect-to-Z boundary is formed by varying the value of the target mean return μ_s in Eq. 35.19. Note that the second term on the right-hand side of Eq. 35.19 is proportional to the vector of weights of an excess return, or zero net investment portfolio (post multiplying that term by a vector of ones implies that the weights sum to zero). The first term in Eq. 35.19 is the weight of the global minimum conditional second moment portfolio. Thus, Eq. 35.19 illustrates two-fund separation: any efficient-with-respect-to-Z portfolio can be found as a combination of the global minimum conditional second moment portfolio and some weight on the unconditionally efficient excess return described by the second term.

Proofs of Propositions

Proof of Propositions 2 and 3

We begin with Proposition 3. To see that Eqs. 35.15 and 35.16 are equivalent, use the fact that efficiency of R_s implies that $\mu_p = \gamma_0 + \frac{\sigma_{ps}}{\sigma_s^2}(\mu_s - \gamma_0)$ and thus $\frac{\mu_s - \gamma_0}{\sigma_s^2} = \frac{\mu_p - \gamma_0}{\sigma_{ps}}$. Next, given any portfolio $R_q \neq R_c$, we will show that $\alpha_q^2/\sigma_q^2 < \alpha_c^2/\sigma_c^2$. Beginning with Eq. 35.16, we compute:

$$\sigma_{cs} = \frac{\sigma_p^2\sigma_s^2 - \sigma_{ps}^2}{\sigma_p^2 - \sigma_{ps}}$$

and

$$\sigma_c^2 = \frac{\sigma_p^2\left(\sigma_p^2\sigma_s^2 - \sigma_{ps}^2\right)}{\left(\sigma_p^2 - \sigma_{ps}\right)^2}.$$

The efficiency of R_s implies that $\mu_c = \gamma_0 + \frac{\sigma_{cs}}{\sigma_s^2}(\mu_s - \gamma_0)$, $\mu_p = \gamma_0 + \frac{\sigma_{ps}}{\sigma_s^2}(\mu_s - \gamma_0)$, and $\mu_q = \gamma_0 + \frac{\sigma_{q\,s}}{\sigma_s^2}(\mu_s - \gamma_0)$. Substituting these expressions and using the fact that $\sigma_{cp} = 0$, which follows from (35.16), we compute:

$$\begin{aligned}
\frac{\alpha_c^2}{\sigma_c^2} - \frac{\alpha_q^2}{\sigma_q^2} &= \frac{\left(\mu_c - \left[\gamma_o + \left(\mu_p - \gamma_o\right)\sigma_{cp}/{\sigma_p}^2\right]\right)^2}{\sigma_c^2} - \frac{\left(\mu_q - \left[\gamma_o + \left(\mu_p - \gamma_o\right)\sigma_{qp}/{\sigma_p}^2\right]\right)^2}{\sigma_q^2} \\
&= (\mu_s - \gamma_0)^2\left(\frac{\sigma_{cs}^2}{\sigma_s^4\sigma_c^2} - \frac{\left(\sigma_{qs}\sigma_p^2 - \sigma_{ps}\sigma_{qp}\right)^2}{\sigma_s^4\sigma_p^4\sigma_q^2}\right) \\
&= (\mu_s - \gamma_0)^2\left(\frac{\sigma_p^2\sigma_s^2 - \sigma_{ps}^2}{\sigma_s^4\sigma_p^2} - \frac{\left(\sigma_{qs}^2\sigma_p^4 + \sigma_{ps}^2\sigma_{qp}^2 - 2\sigma_{qs}\sigma_p^2\sigma_{ps}\sigma_{qp}\right)}{\sigma_s^4\sigma_p^4\sigma_q^2}\right) \\
&= \frac{(\mu_s - \gamma_0)^2}{\sigma_s^4\sigma_p^4\sigma_q^2}\left(\sigma_p^4\sigma_s^2\sigma_q^2 - \sigma_p^2\sigma_q^2\sigma_{ps}^2 - \sigma_{qs}^2\sigma_p^4 - \sigma_{ps}^2\sigma_{qp}^2 + 2\sigma_p^2\sigma_{qs}\sigma_{ps}\sigma_{qp}\right).
\end{aligned}$$

We now use the fact that $\sigma_{sp}^2 < \sigma_s^2\sigma_p^2$ (because, by assumption, R_p and R_s are not perfectly correlated) to see that

$$\frac{\alpha_c^2}{\sigma_c^2} - \frac{\alpha_q^2}{\sigma_q^2} \geq \frac{(\mu_s - \gamma_0)^2}{\sigma_p^4 \sigma_q^2 \sigma_s^4} \left[\sigma_p^4 \sigma_q^2 \sigma_s^2 - \sigma_p^2 \sigma_q^2 \sigma_{sp}^2 - \sigma_p^4 \sigma_{qs}^2 - \sigma_{qp}^2 \left(\sigma_s^2 \sigma_p^2 \right) + 2\sigma_p^2 \sigma_{qs} \sigma_{qp} \sigma_{ps} \right]$$
$$= \frac{(\mu_s - \gamma_0)^2}{\sigma_p^2 \sigma_q^2 \sigma_s^4} \left(\sigma_p^2 \sigma_q^2 \sigma_s^2 - \sigma_q^2 \sigma_{sp}^2 - \sigma_p^2 \sigma_{qs}^2 - \sigma_{qp}^2 \sigma_s^2 + 2\sigma_{qs} \sigma_{qp} \sigma_{ps} \right) \geq 0$$

where the final inequality follows from recognizing that the variance-covariance terms in parentheses are equal to the determinant of the (necessarily nonnegative definite) covariance matrix of (R_p, R_q, R_s). This establishes the maximal property of R_c.

To show uniqueness, note further that the inequality will be strict (and we will have $\alpha_c^2/\sigma_c^2 - \alpha_q^2/\sigma_q^2 > 0$) unless we have both of the following conditions corresponding to the two inequalities in the final calculation: (1) $\sigma_{qp} = 0$ so that R_q and R_p are orthogonal, and (2) the covariance matrix of (R_p, R_q, R_s) is singular so that R_q is a linear combination of R_p and R_s. However, there is only one portfolio orthogonal to R_p that can be formed as a linear combination $\lambda R_p + (1 - \lambda) R_s$, and this solution is R_c. This establishes Proposition 3, which holds in the case of both Theorem 1 and Theorem 2, that is, whether or not there is a conditionally risk-free asset.

The expressions in Proposition 2 for the optimal weights, $x_c(Z)$, follow from substituting portfolio weights from Theorem 1 (if there exists a conditional risk-free asset that may be time varying) or Theorem 2 (otherwise) into Eq. 35.15 and noting that the constants A and B represent the combining portfolio weights implied by (35.15) as $R_c = AR_s + BR_p$. Substituting, we see that Eq. 35.15 implies that the portfolio weight function $Ax_s(Z) + Bx_p(Z)$ generates returns R_c, completing the proof of Proposition 2. QED.

Proof of Proposition 4

Let R_s denote the efficient-with-respect-to-Z portfolio corresponding to zero-beta rate γ_0. The efficiency of R_s implies that we may substitute $\mu_p - \gamma_0 = (\mu_s - \gamma_0)\frac{\sigma_{ps}}{\sigma_s^2}$ and $\mu_c - \gamma_0 = (\mu_s - \gamma_0)\frac{\sigma_{cs}}{\sigma_s^2}$ to find

$$S_p^2 + S_c^2 = \frac{(\mu_p - \gamma_0)^2}{\sigma_p^2} + \frac{(\mu_c - \gamma_0)^2}{\sigma_c^2} = \frac{(\mu_s - \gamma_0)^2}{\sigma_s^4} \left(\frac{\sigma_{ps}^2}{\sigma_p^2} + \frac{\sigma_{cs}^2}{\sigma_c^2} \right).$$

Next, substituting for σ_{cs} and σ_c^2 from the above expressions, we verify that this expression reduces to S_s^2. QED.

Appendix 2: Methodology

We estimate the conditional mean functions, $\mu(Z)$, by ordinary least squares regressions of the returns on the lagged values of the conditioning variables. On the assumption that the conditional mean returns are linear functions of Z, these are the optimal generalized method of moments (GMM, see Hansen 1982) estimators. The covariance matrix of the residuals is used as the estimate of $\Sigma_\varepsilon(Z)$, which is assumed to be constant. These are the maximum likelihood estimates (MLE) under joint normality of (R,Z). In general, the conditional covariance matrix of the returns given Z will be time varying as a function of Z, as in conditional heteroskedasticity. Ferson and Siegel (2003) model conditional heteroskedasticity in alternative ways and find using parametric bootstrap simulations that this increases the tendency of the efficient-with-respect-to-Z portfolio weights to behave conservatively in the face of extreme realizations of Z.

The optimal orthogonal portfolio weights in Table 35.1 and Fig. 35.1 are estimated from Eqs. 35.7 to 35.8 in the text where, in the time-varying risk-free rate case, the Treasury bill return is assumed to be conditionally risk-free in Eq. 35.8. The benchmark portfolio x_p is a vector with a 1.0 in the place of the market index and zeros elsewhere. The matrix Q is estimated by using the MLE estimates of $\mu(Z)$ and $\Sigma_\varepsilon(Z)$ in the function given by Eq. 35.13. The parameter μ_p is estimated as the sample excess return on the market index, and γ_0 is the sample mean of the Treasury return, 3.8 %.

The Parametric Bootstrap

The parametric bootstrap is a special case of the simple, or nonparametric, bootstrap, itself an example of a resampling scheme. Introduced by Efron (1979), the bootstrap is useful when we wish to conduct statistical inferences, but when we either don't have an analytical formula for the sampling variation of a statistic, don't wish to assume normality or some other convenient distribution that allows for an analytical formula or have a sample too small to trust asymptotic distribution theory. The basic idea is to build a sampling distribution by resampling from the data at hand. In the simplest example, we have some statistic that we have estimated from a sample, and we want to know its sampling distribution. We resample from the original data, randomly with replacement, to generate an artificial sample of the same size, and we compute the statistic on the artificial sample. Repeating this many times, the histogram of the statistics computed on the artificial samples is an estimate of the sampling distribution for the original statistic. This distribution can be used to estimate standard errors, confidence intervals, etc. We can think of the bootstrap samples as being related to the original sample as the original sample is to the population. There are many variations on the bootstrap, and a good overview is provided by Efron and Tibshirani (1993).

In the simple, or nonparametric, bootstrap, no assumptions are made about the form of the distribution. It is assumed, however, that the sample accurately reflects the underlying population distribution, and this is critical for reliable inferences. For example, suppose that the true distribution was a uniform on [0, M]. In a sample drawn from this distribution, the maximum value is likely to be smaller than M, so that the bootstrap will likely understate the true variability of the data. This problem is obviously worse if the original sample has fewer observations. If the data are contaminated with measurement errors, in contrast, the extent of the true variability can be overstated. Even with large sample sizes, the bootstrap can be unreliable. For example, if the true distribution has infinite variance, the bootstrap distribution for the sample mean is inconsistent (Athreya 1987).

With a parametric bootstrap, we can sometimes do better than with a nonparametric bootstrap, where "do better" means, for example, obtain more accurate confidence intervals (e.g., Andrews et al. 2006). The idea of the parametric bootstrap is to use some of the parametric structure of the data. This might be as simple as assuming the form of the probability distribution. For example, assuming that the data are independent and normally distributed, we can generate artificial samples from a normal distribution using the sample mean and variance as the parameters. This is not exactly the right thing to do, because we should be sampling from a population with the true parameter values, not their estimated values. But, if the estimates of the mean and variance are good enough, we should be able to obtain reliable inferences.

To illustrate and further suggest the flexibility of the parametric bootstrap, consider an example, similar to the setting in our paper, where we have a regression of stock returns on a vector of lagged instruments, Z, which are highly persistent over time. Obviously, sampling from the Z's randomly with replacement would destroy their strong time-series dependence. Time-series dependence can be accommodated in a nonparametric way by using a block bootstrap. Here, we sample randomly a block of consecutive observations, where the block length is set to capture the extent of memory in the data.

In order to capture the time-series dependence of the lagged Z in a parametric bootstrap, we can model the lagged instruments as vector AR(1), for example, retaining the estimator of the AR(1) coefficient and the model residual, which we call the shocks, U_z. Regressing the future stock returns on the lagged instruments, we retain the regression coefficient and the residuals, which we call the shocks, U_r. We generate a sample of artificial data, with the same length as the original sample, as follows. We concatenate the shocks as $v = (U_z, U_r)$. Resampling rows from v, randomly with replacement, retains the covariances between the stock return shocks and the instrument Z shocks. This can be important for capturing features like the lagged stochastic regressor bias described by Stambaugh (1999). Drawing an initial row from v, we take the U_z shock and add it to the initial value of the Z (perhaps, drawn from the unconditional sample distribution) to produce the first lagged instrument vector, Z_{t-1}. We draw another row from v and construct the first observation of the remaining data as follows. The stocks' returns are formed by adding the U_r shock to βz_{t-1}, where β is the "true" regression coefficient that defines the conditional expected return. The contemporaneous values of the Zs are formed by multiplying the VAR coefficient by Z_{t-1} and adding the shock, U_z. The next observation is generated by taking the previous contemporaneous value of the Z as Z_{t-1} and repeating the process. In this way, the Z values are built up recursively, which captures their strong serial correlation.

References

Abhyankar, A., Basu, D., & Streme, A. (2007). Portfolio efficiency and discount factor bounds with conditioning information: An empirical study. *Journal of Banking and Finance, 31*, 419–437.

Andrews, D. W. K., Lieberman, O., & Marmer, V. (2006). Higher order improvements of the parametric bootstrap for long-memory gaussian processes. *Journal of Econometrics, 133*, 673–702.

Athreya, K. B. (1987). Bootstrap of the mean in the infinite variance case. *Annals of Statistics, 15*, 724–731.

Bekaert, G., & Liu, J. (2004). Conditioning information and variance bounds on pricing kernels. *Review of Financial Studies, 17*, 339–378.

Campbell, J., Lo, A., & MacKinlay, A. C. (1987). *The econometrics of financial markets*. Princeton: Princeton University Press.

Chiang, E. (2012). *Modern portfolio management with conditioning information*. Working paper. University of North Carolina, Charlotte.

Dybvig, P., & Ross, S. (1985). Differential information and performance measurement using a security market line. *Journal of Finance, 40*, 383–399.

Efron, B. (1979). Bootstrap methods: Another looks at the jackknife. *Annals of Statistics, 7*, 1–26.

Efron, B., & Tibshirani, R. J. (1993). *Introduction to the bootstrap*. New York: Chapman and Hall.

Fama, E. F. (1973). A note on the market model and the two-parameter model. *Journal of Finance, 28*, 1181–1185.

Ferson, W. E., & Siegel, A. F. (2001). The efficient use of conditioning information in portfolios. *Journal of Finance, 56*, 967–982.

Ferson, W. E., & Siegel, A. F. (2003). Stochastic discount factor bounds with conditioning information. *Review of Financial Studies, 16*, 567–595.

Ferson, W. E., & Siegel, A. F. (2009). Testing portfolio efficiency with conditioning information. *Review of Financial Studies, 22*, 2735–2758.

Ferson, W. E., Sarkissian, S., & Simin, T. (2003). Spurious regressions in financial economics? *Journal of Finance, 58*, 1393–1414.

Ferson, W. E., Siegel, A. F., & Xu, P. (2006). Mimicking portfolios with conditioning information. *Journal of Financial and Quantitative Analysis, 41*, 607–635.

Frost, P. A., & Savarino, J. E. (1988). For better performance: Constrain portfolio weights. *Journal of Portfolio Management, 15*, 29–34.

Gibbons, M. R., Ross, S. A., & Shanken, J. (1989). A test of the efficiency of a given portfolio. *Econometrica, 57*, 1121–1152.

Grinold, R. C., & Kahn, R. N. (1992). *Active portfolio management* (2nd ed.). New York: McGraw Hill.

Hansen, Lars P. (1982). Large Sample Properties of Generalized Method of Moments Estimators. *Econometrica*, *50*, 1029–1054.

Hansen, L. P., & Richard, S. F. (1987). The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. *Econometrica, 55*(3), 587–613.

Jagannathan, R., & Ma, T. (2003). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *Journal of Finance, 58*, 1651–1684.

Jobson, J. D., & Korkie, B. (1982). Potential performance and tests of portfolio efficiency. *Journal of Financial Economics, 10*, 433–466.

Jorion, P. (2003). Portfolio optimization with tracking error constraints. *Financial Analysts Journal, 59*, 70–82.

Kandel, S. A. (1984). The likelihood ratio test statistic of mean variance efficiency without a riskless asset. *Journal of Financial Economics, 13*, 575–592.

Kandel, S. A., & Stambaugh, R. F. (1987). On correlations and inferences about mean-variance efficiency. *Journal of Financial Economics, 18*, 61–90.

Kandel, S. A., & Stambaugh, R. F. (1989). A mean-variance framework for tests of asset pricing models. *Review of Financial Studies, 2*, 125–156.
MacKinlay, A. C. (1995). Multifactor models do not explain deviations from the CAPM. *Journal of Financial Economics, 38*, 3–28.
Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance, 7*, 77–91.
Michaud, R. O. (1989). The Markowitz optimization enigma: Is 'optimized' optimal? *Financial Analysts Journal, 45*, 31–42.
Roll, R. R. (1977). A critique of the asset pricing theory's tests – part 1: On past and potential testability of the theory. *Journal of Financial Economics, 4*, 129–176.
Roll, R. R. (1980). Orthogonal portfolios. *Journal of Financial and Quantitative Analysis, 15*, 1005–1023.
Roll, R. R. (1982). A mean variance analysis of tracking error. *Journal of Portfolio Management, 18*, 13–22.
Roll, R. R. (1985). A note on the geometry of Shanken's CSR T^2 test for mean/variance efficiency. *Journal of Financial Economics, 14*, 349–357.
Shanken, J. (1987). Multivariate proxies and asset pricing relations: Living with the Roll critique. *Journal of Financial Economics, 18*, 91–110.
Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance, 19*, 425–442.
Siegel, A. F., & Woodgate, A. (2007). Performance of portfolios optimized with estimation error. *Management Science, 53*, 1005–1015.
Stambaugh, R. F. (1999). Predictive regressions. *Journal of Financial Economics, 54*, 375–421.
Treynor, J. L., & Black, F. (1973). How to use security analysis to improve portfolio selection. *Journal of Business, 46*, 66–86.

Multifactor, Multi-indicator Approach to Asset Pricing: Method and Empirical Evidence

36

Cheng-Few Lee, K. C. John Wei, and Hong-Yi Chen

Contents

Abstract

This paper uses a multifactor, multi-indicator approach to test the capital asset pricing model (CAPM) and the arbitrage pricing theory (APT). This approach is able to solve the measuring problem in the market portfolio in testing CAPM; and it is also able to directly test APT by linking the common factors to the macroeconomic indicators. Our results from testing CAPM support

C.-F. Lee (✉)
Department of Finance and Economics, Rutgers Business School, Rutgers, The State University of New Jersey, Piscataway, NJ, USA

Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
e-mail: cflee@business.rutgers.edu; lee@business.rutgers.edu

K.C.J. Wei
Hong Kong University of Science and Technology, Kowloon, Hong Kong
e-mail: johnwei@ust.hk

H.-Y. Chen
Department of Finance, National Central University, Taoyuan, Taiwan
e-mail: fnhchen@ncu.edu.tw

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_36,

Stambough's (*Journal of Financial Economics*, *10*, 237–268, 1982) argument that the inference about the tests of CAPM is insensitive to alternative market indexes.

We propose a MIMIC approach to test CAPM and APT. The beta estimated from the MIMIC model by allowing measurement error on the market portfolio does not significantly improve the OLS beta, while the MLE estimator does a better job than the OLS and GLS estimators in the cross-sectional regressions because the MLE estimator takes care of the measurement error in beta. Therefore, the measurement error problem on beta is more serious than that on the market portfolio.

Keywords

Capital asset pricing model, CAPM • Arbitrage pricing theory • Multifactor multi-indicator approach • MIMIC • Measurement error • LISREL approach • Ordinary least square, OLS • General least square, GLS • Maximum-likelihood estimation, MLE

36.1 Introduction

Roll (1977) has shown that the capital asset pricing model (CAPM) can never be tested unless the market portfolio is capable of being measured and identified. However, the market portfolio is actually unobservable. Stated differently, since the market portfolio is subject to measurement error, Sharpe (1964), Lintner (1965), and Mossin's (1966) type of CAPM can never be tested directly. In contrast, the test of Ross's (1976, 1977) arbitrage pricing theory (APT) does not rely upon the identifications of the market portfolio or the true factors. Nevertheless, Shanken (1982) argues that Ross's contention that APT is inherently more easily tested is questionable. If we can directly link these unobservable factors to some observable indicators, the Shanken criticism of the test of APT can be avoided or reduced.

Fortunately, a multiple indicators and multiple causes (MIMIC) model, proposed by Zellner (1970), Goldberger (1972a, b), Joreskog and Goldberger (1975), and others, is an attractive methodology in dealing with the problem of unobservable variables. Goldberger (1972b) conceptually described that structural equation model is a combination of factor analysis and econometrics model. Goldberger (1972b) and Joreskog and Goldberger (1975) develop a structural equation model with multiple indicators and multiple causes of a single latent variable, MIMIC model, and obtain maximum likelihood estimates of parameters. The MIMIC model displays a mixture of econometric and factor analysis themes. This concept has successfully been used to test economic models, such as the structure of price expectation and inflationary expectation (see Turnovsky 1970; Lahiri 1976). In addition, structural equation model has also been used in finance-related studies. Titman and Wessels (1988), Chang et al. (2009), and Yang et al. (2009) apply structural equation models (e.g., LISREL approach and MIMIC model) in determining capital structure decision. Maddala and

Nimalendran (1996) use structural equation model to examine the effect of earnings surprises on stock prices, trading volumes, and bid-ask spread. However, the structural equation model, especially MIMIC model, has not been used in capital asset pricing determination.

The purpose of this paper is twofold: (i) to use the MIMIC model to reexamine CAPM and (ii) to use the MIMIC model to investigate the relationship between the factors in APT and the macroeconomic indicators directly. APT is attractive to both academicians and practitioners, because the model allows more than one factor. However, to date the practical applications of APT are still limited since previous studies in testing the model do not directly link the factors to the indicators. If the linkage between the factors and the indicators can be derived, practical applications will be much improved.

The outline of this paper is as follows. In Sect. 36.2, the MIMIC model is reviewed and CAPM and APT in terms of the MIMIC model are demonstrated. Section 36.3 shows how MIMIC can be used to test CAPM, and Sect. 36.4 investigates the MIMIC applied to test APT. Finally, a brief summary is contained in Sect. 36.5.

36.2 The MIMIC Model and the Tests of CAPM and APT

36.2.1 The MIMC Model

Suppose that a system has k unobservable latent variables $z = (z_1, \ldots, z_k)'$, p observable exogenous indicators $x = (x_1, \ldots, x_p)'$, and m observable endogenous variables $y = (y_1, \ldots, y_m)'$.[1] The specification of this extended MIMIC model of Jöreskog and Goldberger (1975) is as follows. The latent factors z are linearly determined, subject to disturbances $e = (e_1, \ldots, e_k)'$, by observable exogenous indicators x:[2]

$$z = ax + e, \tag{36.1}$$

where

$$a = \begin{bmatrix} a_{11}, & \cdots, & a_{1p} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ a_{k1}, & \cdots, & a_{kp} \end{bmatrix} \quad \text{is a } k \times p \text{ matrix.}$$

[1]Fogler et al. (1981) and Chen et al. (1983) indirectly link the factors extracted from the APT to economic indicators. Jöreskog and Goldberger (1975) have shown that this kind of indirect estimation method is not as efficient as the direct estimation method to be explored in this section.

[2]Here, we use different terminologies in defining the factors and indicators compared with those used in traditional MIMIC model.

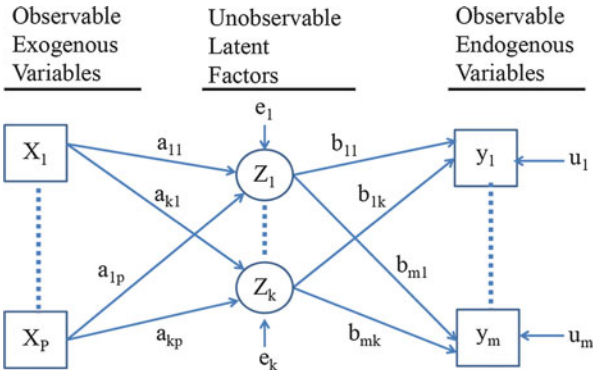


Fig. 36.1 Multiple causes and indicators of unobservable variables. This figure shows the path diagram illustrating MIMIC model of the structural equation system. In this figure, observable variables $X_1, \ldots, X_P$ are causes of the latent variables $Z_1, \ldots, Z_k$, while $Y_1, \ldots, Y_m$ are indicators of $Z_1, \ldots, Z_k$

In addition, the latent factors z linearly determine the components of endogenous variables y subject to disturbances $u = (u_1, \ldots, u_m)'$:

$$y = bz + u, \tag{36.2}$$

where

$$b = \begin{bmatrix} b_{11}, & \cdots, & b_{1k} \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ b_{m1}, & \cdots, & b_{mk} \end{bmatrix} \text{ is an } m \times k \text{ matrix.}$$

The disturbances are assumed to be mutually independent and normally disturbed with mean zero, namely, $e \sim N(0,\Sigma)$, $u \sim N(0,\theta^2)$, where $\Sigma = \text{diag}(\sigma_1^2, \ldots, \sigma_k^2)$ and $\theta^2 = \text{diag}(\theta_1^2, \ldots, \theta_m^2)$. For convenience, all variables are taken to have mean zero. The system of Eqs. 36.1 and 36.2 are shown in Fig. 36.1.

Solving the equation systems of Eqs. 36.1 and 36.2, we have the following reduced form connecting the observable variables:

$$y = bax + be + u = hx + v, \tag{36.3}$$

where the reduced-form coefficient matrix is

$$h = ba. \tag{36.4}$$

The reduced-form disturbance matrix is

$$v = be + u, \tag{36.5}$$

which has a covariance matrix of

$$\Omega = E(vv') = E\left[(be + u)(be + u)'\right] = b\Sigma b' + \theta^2, \tag{36.6}$$

where E presents expectation operator.

There are two types of restrictions on the reduced form: (i) the $m \times p$ regression coefficient matrix h has rank k, the $m \times p$ components of h being expressed in terms of $kp + mk$ elements of a and b, and (ii) the $m \times m$ residual covariance matrix Ω satisfies a factor analysis model with k common factors, the $m(m + 1)/2$ distinct elements of Ω being expressed in terms of the $k + km + m$ elements of σ^2, b, and θ^2. The first restriction, which is the same as the simultaneous equation model, is familiar to econometricians. The second restriction, which is the same as the factor analysis model, is familiar to psychometricians. In Eq. 36.5, e, b, and u are regarded as the common factors, the factor loadings, and the unique disturbances in the factor analysis model, respectively.

We observe that the reduced-form parameters remain unchanged, when any column, say j, of b is multiplied by a scalar and the j th row of a and σ_j are both divided by the same scalar. To remove this indeterminacy of the model, we can normalize the model through (i) σ^2, or (ii) b, or (iii) a. After normalization, the maximum likelihood estimation (MLE) procedure can be used to obtain consistent estimators of the elements in parameters a, b, and θ^2 (see Attfield 1983; Chen 1981; Joreskog and Goldberger 1975; and others). In the following, we demonstrate how to apply the MIMIC model to test CAPM and APT.

36.2.2 The Testing Model of CAPM by the MIMIC Approach

The CAPM can be rewritten, in terms of MIMIC model, as follows:

$$\begin{aligned} \widetilde{r}_i &= \beta_i \widetilde{r}^*_m + \widetilde{u}_i \\ \widetilde{r}^*_m &= \widetilde{r}_m + \widetilde{e}_m, \end{aligned} \qquad i = 1, \ldots, N, \tag{36.7}$$

where

$\widetilde{r}_i$ = the realized excess return (total return less risk-free rate) on security i in a deviation form

$\widetilde{r}_m$ = the realized excess return of the NYSE Composite Index

$\widetilde{r}^*_m$ = the unobservable excess return on the market portfolio

In this special one-factor case, we remove the indeterminacy by setting the coefficient on $\widetilde{r}_m$ equal to one. Equation 36.7 is a simultaneous equation model, in which there are N equations linking the individual security (or portfolio) return to the unobservable true market return and one equation linking the unobservable true market portfolio return to the realized return of the NYSE composite index. After obtaining the estimated betas from simultaneous equation system of (36.7), a cross-sectional regression of the security return against its risk (β) will be used to test CAPM or to estimate the riskless rate and the market risk premium as follows:

$$r_{it} = \hat{a}_{0t} + \hat{a}_{1t}\beta_i, \tag{36.8}$$

where
r_{it} = the excess return on security i at time t
$\hat{a}_{0t}$ = the estimate of the intercept which is supposed to be zero
$\hat{a}_{1t}$ = the estimate of the market risk premium

Four different estimation procedures will be used to estimate Eq. 36.8. They are: (i) stationary OLS, (ii) nonstationary OLS, (iii) GLS, and (iv) MLE.[3]

36.2.3 The Testing Model of APT by the MIMIC Approach

The testing model of APT, in terms of the MIMIC model, can be rewritten as follows:

$$\begin{array}{ll} \widetilde{r}_i = b_{i1}\widetilde{f}_1 + \ldots + b_{ik}\widetilde{f}_k + \widetilde{u}_i, & (i = 1, \ldots, N) \\ \widetilde{f}_j = a_{j1}\widetilde{I}_1 + \ldots + a_{jp}\widetilde{I}_p + \widetilde{e}_j, & (j = 1, \ldots, k) \end{array}, \tag{36.9}$$

where
$\widetilde{f}_j$ = the jth unobservable factor
$\widetilde{I}_h$ = the hth macroeconomic indicator, $h = 1, \cdots, p$

For convenience and easy explanation, each factor $\widetilde{f}$ is assumed to have different set of explained indicators $\widetilde{I}$'s. Note that there are N return equations plus k factor equations in the simultaneous equation system (36.9). The LISREL computer program of Jöreskog and Sörbon (1981) is used to estimate the parameters, a and b, in Eq. 36.9. A cross-sectional regression is also used to test APT and to estimate the riskless rate and the factor risk premia by regressing the security return against its risks, b's. The a's coefficients in Eq. 36.9 can be used to explain the relationship between factors and indicators.

36.3 Test of CAPM

This section tests CAPM using the market model and the MIMIC model described in Sect. 36.2. The objective is to investigate whether the MIMIC method yields a different inference from the market model. Nineteen industry common stock portfolios are formed with the same manner used by Schipper and Thompson (1981), and Stambough (1982).[4] The return on a portfolio is the arithmetic average of returns for firms on the CRSP monthly tape with the appropriate two-digit SEC code for the given month. The tests use returns from the period 1963–1982, and this total period is divided into two equal subperiods: (i) subperiod 1, 1963–1972 and

[3]The terminologies stationary OLS and nonstationary OLS have been used by Friend and Westfield (1980). The GLS and MLE methods have been discussed by Litzenberger and Ramaswamy (1979).

[4]See Chen (1981).

(ii) subperiod 2, 1973–1982. Portfolios are formed primarily because they provide a convenient way to limit the computational dimensions of the MIMIC method. As mentioned by Stambough (1982), industry portfolios also allow rejection of CAPM due to the presence of additional industry-related variables in the risk-return relation.

Table 36.1 indicates the number of securities in each portfolio, the SEC codes, and betas calculated from the market model and also from the MIMIC model. When the MIMIC model was used to estimate betas by the portfolio excess returns in subperiod 1, convergent problems were encountered during minimization so that the result is inappropriate. Consequently, raw returns in deviation form on the portfolios are used to estimate betas for subperiod 1. The betas estimated from the MIMIC model are very close to those from the market model in both periods. This evidence supports Stambough's discovery that inferences about CAPM are very insensitive to alternative market indexes.

Table 36.2 presents return-risk trade-off from the cross-sectional relationship in which the average monthly excess portfolio returns (monthly portfolio returns less monthly return on 3-month Treasury bills) is regressed on a beta estimated either from the MIMIC model or the market model from two different 120-month periods. Four different estimation procedures are used to estimate the intercepts and the market risk premia. The OLS method presents two sets of *t*-statistics shown in the parentheses under the same relevant regression coefficients. The first set (denoted *S*) assumes that the regression coefficients are constant or stationary over each 120-month period. The second set (denoted *NS*) of the OLS *t*-statistics allows the nonstationarities of the regression coefficients by computing the cross-sectional regression coefficients in each month and deriving the appropriate standard errors from the time series 120 estimates of the OLS regression coefficients. The GLS and MLE methods also permit the nonstationary coefficients. Thus, their *t*-statistics are derived from the same procedure in the OLS (NS) method. Although the OLS and the GLS estimators are biased and inconsistent due to measurement error in beta (see Litzenberger and Ramaswamy 1979), the maximum likelihood estimators are consistent simply because MLE takes care of measurement error in beta.

In Table 36.2, the coefficients in the *NS* regression of OLS, GLS, and MLE are obviously characterized by much larger standard errors so that they lose any significance shown in the stationary OLS regression.[5]

However, different estimated betas cause little changes in return-risk relationships. From the results of the OLS stationary method, there exists a significant return-risk relationship in subperiod 1, but not in subperiod 2 in both MIMIC and market models. The poor return-risk relationship in subperiod 2 may be due to the poor performance of CAPM in determining a pricing relation. In the next section, APT will be used to examine an alternative pricing relationship. Even though the null hypothesis of CAPM that $a_0 = 0$ cannot be rejected at the 5 % level in all four cases, all coefficients are positive. In addition, the intercept in subperiod 2, about

[5]The similar results were also found in the Friend and Westfield's (1980) study of co-skewness.

Table 36.1 Industry portfolio SEC codes, number of firms, and estimated betas

Portfolio description	SEC code	# of firms		Estimated betas			
				Period 1		Period 2	
		1972	1982	Market model	MIMC model	Market model	MIMC model
1. Mining	10–14	56	71	1.056	1.009	0.922	0.916
2. Food and beverages	20	75	51	0.894	0.841	0.803	0.803
3. Textile and apparel	22,23	58	45	1.264	1.220	1.081	1.082
4. Paper products	26	30	30	1.030	1.029	0.910	0.909
5. Chemical	28	87	83	0.949	0.947	0.847	0.846
6. Petroleum	29	28	22	0.706	0.792	0.745	0.740
7. Stone, clay, glass	32	43	31	1.050	1.106	1.045	1.045
8. Primary metals	33	56	49	1.136	1.193	0.932	0.930
9. Fabricated metals	34	45	46	1.145	1.155	1.102	1.102
10. Machinery	35	93	104	1.234	1.231	1.104	1.104
11. Appliance and elec. equip.	36	87	82	1.384	1.401	1.179	1.180
12. Transport. equip.	37	64	50	1.209	1.275	1.150	1.151
13. Misc. manufactrng.	38,39	64	59	1.375	1.314	1.197	1.198
14. Railroads	40	18	11	1.294	1.229	0.899	0.895
15. Other transport.	41, 42, 44 45, 47	34	35	1.335	1.447	1.203	1.203
16. Utilities	49	138	152	0.443	0.467	0.564	0.562
17. Department stores	53	35	28	1.149	1.104	1.125	1.125
18. Other retail trades	50–52, 54–59	103	97	1.144	1.088	1.123	1.124
19. Banking, finance, real estate	60–67	184	240	0.968	1.021	1.069	1.068

This table presents the number of firms and betas in each industry portfolio. The numbers of firms are obtained in the end of 1972 and the end of 1982. The betas shown in the first column are estimated from the market model, while those in the second column are estimated from the MIMIC model. Period 1 represents the sample period from 1963 to 1972 and period 2 represents the sample period from 1973 to 1982

4.3 % annual rate, is too high. This is consistent with prior tests of the traditional version of CAPM.

All nonstationary estimates of a_0 and a_1 in OLS, GLS, and MLE in Table 36.2 are insignificant. Because of the low test power for all nonstationary procedures (the standard errors are too high), in the following, only the magnitudes of the estimated coefficients are discussed. The MLE and GLS estimates of a_0 are much lower and much closer to zero than the corresponding OLS estimates in all four cases. In addition, the MLE estimates of a_1 is greater than the corresponding GLS estimates and is closer to the realized market risk premia in all four cases. The realized market risk premia are 0.751 % and 0.636 % monthly for periods 1 and 2, respectively.

Table 36.2 Return-risk cross-sectional relationships of CAPM: 1963–1982

	Panel A: MIMC model			Panel B: Market model		
Procedure	a_0	a_1	$\overline{R}^2$	a_0	a_1	$\overline{R}^2$
Period 1: 1963–1972						
OLS-S	0.159	0.539	0.258	0.187	0.516	0.243
	(0.71)	(2.70)*		(0.85)	(2.61)*	
OLS-NS	0.159	0.539	0.258	0.187	0.516	0.243
	(0.41)	(1.03)		(0.48)	(1.02)	
GLS	0.029	0.660		0.095	0.599	
	(0.06)	(1.29)		(0.19)	(1.22)	
MLE	−0.106	0.793		−0.082	0.770	
	(−0.24)	(1.29)		(−0.13)	(1.22)	
Period 2: 1973–1982						
OLS-S	0.370	0.266	−0.004	0.363	0.273	−0.001
	(1.33)	(0.97)		(1.29)	(0.99)	
OLS-NS	0.370	0.266	−0.004	0.363	0.273	−0.001
	(0.60)	(0.31)		(0.59)	(0.32)	
GLS	0.106	0.532		0.107	0.531	
	(0.24)	(0.71)		(0.24)	(0.71)	
MLE	0.081	0.556		0.081	0.556	
	(0.17)	(0.71)		(0.17)	(0.71)	

This table presents return-risk trade-off from the cross-sectional relationship in which the average monthly excess portfolio returns are regressed on a beta estimated either from the MIMC model or the market model from two different 120-month periods. For different estimation procedures, stationary OLS, nonstationary OLS, GLS, and MLE, are used to estimate the intercepts and the market risk premia. The stationary OLS version corresponds to the regression $\overline{R}_i - \overline{R}_f = a_0 + a_1\beta_i + \widetilde{e}_i$, ($i$=1,· · ·,N). The other versions correspond to the regression $\overline{R}_i - \overline{R}_f = a_0 + a_1 \beta_i + \widetilde{e}_i$, ($i = 1$,· · ·,N; and $t = 1$,· · ·T. The monthly returns are multiplied by 100 before regressing. The reported coefficients are arithmetic average of the time series, while t-statistics are in parentheses under each relevant coefficient. *represents significant at the 5 % level

This evidence proves that the MLE estimator in the return-risk cross-sectional regressions is more appropriate than OLS or GLS estimator in testing CAPM.

In sum, we have proposed an alternative estimator of betas by the MIMIC model in which measurement error in a market portfolio is allowed. Nevertheless, this reasonable alternative method does not gain much from the traditional OLS estimator. However, some interesting results have surfaced. This evidence supports Stambough's conclusion that the tests of CAPM are insensitive to different market indexes. In return-risk cross-sectional regressions, our evidence shows that the MLE estimator is more appropriate than the OLS or GLS estimator due to measurement error in beta. From these two interesting results, we conclude that measurement error on beta is more serious than measurement error on the market portfolio in testing CAPM.

36.4 Tests of APT by MIMIC Approach

This section tests APT using the MIMIC model demonstrated in Sect. 36.2. The objective is to investigate that (i) the proper number of factors is used to explain the data and (ii) the relationships between factors and indicators which are measured by macroeconomic variables. The same 19 industry portfolios described in previous section are used here. The macroeconomic variables are selected from those most likely related to common stock returns. In the following, the indicators selected in this study will be discussed.

36.4.1 Macroeconomic Variables as the Indicators

In early 1970s, several studies attempt to employ economic methods to investigate the relationship between money supply and aggregated common stock prices. Models developed by Keran (1971), Homa and Jaffee (1971), and Hamburger and Kochin (1972) appear to have met with considerable success in explaining the behavior of Standard and Poor's Composite Index. However, Pesando (1974) reexamines above models using different periods. He finds that the extraordinary success of these methods in tracking the behavior of stock prices during the sample period may be illusory. We believe that the above spurious regression phenomenon results from ignoring the autocorrelated errors in time series regression equations as pointed out by Granger and Newbold (1974). Gargett (1978) used a qualitative method to study the relationship between these two variables. He discovered that the Dow Jones Industrial Index follows changes in money supply with a lag of 3 months.

The relationship between stock returns and inflation has been extensively studied. In particular, Bodie (1976), Nelson (1976), and Fama and Schwert (1977) all present evidence that monthly returns to NYSE Composite Index are negatively related to the inflation rate as indicated by the consumer price index (CPI) since 1953. Cohn and Lessard (1981), Gultekin (1983), and Solnik (1983) also find that stock prices are negatively related to nominal interest rate and inflation in a number of countries. Fama (1981) suggests a reason why the stock market reaction to unexpected inflation is weak. He argues that unexpected inflation is contemporaneously correlated with unexpected movements in important "real" variables, such as capital expenditures or real GNP, so that the correlation between stock returns and unexpected inflation is spurious. After extensively reexamining the relationship between the stock returns and inflation, Geske and Roll (1983) conclude that only Nelson's (1979) and Fama's (1981) money demand explanation is logically consistent, but it seems unable to fully explain all of the empirical phenomena. Geske and Roll (1983) therefore propose the fiscal and monetary explanation. They argue that the basic underlying relation is between stock returns and changes in inflationary expectations.

In exploring the common stocks as hedges against the investment opportunity sets, Schipper and Thompson (1981) select two candidates for state variables.

They are price level as measured by consumer price index (CPI) and the real gross national product less corporate profit (GNP). They find that hedge portfolios offer meaningful hedging potential in portfolio-formation period. In addition, CAPM or the market model indicates that the return on a security or a portfolio most likely co-move with the return on the market portfolio.

In summary, the variables most likely correlated with a stock return would be classified as five categories: (1) money supply, (2) real production, (3) inflation, (4) interest rate, and (5) market return. Further, Brigham (1982) decomposes a risk premium into maturity risk premium and default risk premium. Thus, these two indicators are also included in our study. According to the above discussion, the following 11 variables are selected as the indicators.

1. Return on the market portfolio (RM): the return on NYSE common stock composite index.
2. Transaction volume (VL): the change rate in the transaction volume (shares) for all of the NYSE common stocks.
3. Real riskless rate (RF): the real interest rate on 3-month Treasury bills.
4. Maturity risk premium (MP): the difference between the real interest rates on long-term Treasury bonds (10 or more years) and on 3-month Treasury bills.
5. Default risk premium (DP): the difference between the real interest rates on new AA corporate bonds and on 3-month Treasury bills.
6. Consumer price index inflation rate (CPI): the change rate in urban consumer price index for all items.
7. Money supply (M2): the real change rate in money stock as measured by M2 (M1 + time deposits).
8. Velocity of money supply (PI/M2): the ratio of personal income to money supply M2. This is an alternative measure of money supply.
9. Real industrial production (IP): the change rate in real total industrial productions.
10. Real auto production (IPA): the change rate in real automotive products.
11. Real home production (IPH): the change rate in real home goods.

Since the automobile and housing industries generally lead the rest of the economy, the last two indicators, IPA and IPH, are used to catch up the first two biggest industries. The reason to select industrial production instead of GNP in this study is that all other indicators are published monthly while GNP is published quarterly. Since the industrial production is a very good proxy for GNP, we sacrifice GNP measure to gain the number of time periods.

36.4.2 Empirical Results

After carefully selecting the indicator candidates, the time lag or leading problem needs to be solved. The correlation coefficients between the market portfolio return and the other indicators with lags and leadings of zero to 5 months for both periods were examined. All indicators are contemporaneously correlated with the

Table 36.3 Eigenvalue as a percentage of the first eigenvalue for 19 industry portfolios: 1963–1982

Factor	PRC	ALP	SCF	ULS
Panel A: Period: 1963–1972				
1	100 %	100 %	100 %	100 %
2	3.8	4.4	3.9	2.9
3	3.1	1.5	2.6	1.7
4	2.1	1.0	2.3	0.8
5	1.9	0.7	2.0	0.7
6	1.6	0.6	1.9	0.6
Panel B: Period: 1973–1982				
1	100 %	100 %	100 %	100 %
2	7.9	6.9	7.1	7.2
3	2.1	2.8	2.9	1.8
4	1.3	1.0	1.9	0.9
5	1.0	0.6	1.6	0.5
6	1.0	0.4	1.4	0.4

This table shows the eigenvalues as a percentage of the first eigenvalue. Four methods are used to determine the number of factors in APT model. They are principal component analysis (PRC), alpha factor analysis (ALP), simple common factor analysis (SCF), and unweighted least squares method (ULS)

market portfolio except three real production indicators. The real production indicators follow the market portfolio return with a lag of 2 or 3 months. Therefore, in this study, all indicators are contemporaneous except three real production indicators which is a 2-month lag.

Before APT is directly tested by the MIMIC model, factor analysis is preliminarily used to determine the number of factors in both periods. Table 36.3 shows the eigenvalues as a percentage of the first eigenvalue. Clearly, it is only one factor in period 1, while it is perhaps two factors in period 2 by "scree" test described in Chapter 4 of Wei (1984).[6] Consequently, at most a two-factor model is enough to explain the historical data. Three alternative MIMIC models are proposed to test APT: (i) one-factor 11-indicator model, (ii) one-factor six-indicator model, and (iii) two-factor six-indicator model. When the two one-factor models are used to test APT in period 1, there is little difference between 11-indicator and six-indicator model. Thus, only six indicators are used in the two-factor model to save the computer time.[7]

The structural coefficients of APT in the MIMIC model for period 1 are reported in Table 36.4a. In both one-factor models, only the stock market-related variables,

[6] In his dissertation, Wei (1984) has shown that the "scree" test is a powerful test in identifying the number of relevant factors in the APT. By using simulation study, Wei has shown that Roll and Ross's (1980) ML method in estimating factors are inferior to methods listed in Table 36.3.

[7] It is very expensive to run LISREL program, especially for more than two factor models.

the market return (RM), and the transaction volume (VL) are significant at the 5 % level. From this evidence, if the pricing relation in period 1 is a one-factor APT, this common factor would be most likely related to only the stock market-related indicators, namely, the market portfolio return and the transaction volume. Other indicators may be correlated with this single common factor, but they are obviously not as important as the stock market-related indicators. Now, let us closely examine other indicators with an absolute *t*-value of greater than one for 11-indicator model. The real riskless interest rate (RF), CPI inflation rate (CPI), and the real auto production (IPA) are negatively correlated with this common factor, while the velocity of money supply (PI/M2) is positively related to this common factor. If we regard this common factor as a proxy of market portfolio because most of the weight is on the market portfolio, then, except for real auto production, this evidence supports previous studies done on the relationship between common stock returns and other indicators (Keran 1971; Homa and Jaffee 1971; Hamburger and Kochin 1972; Pesando 1974; Bodie 1976; Nelson 1976; Fama and Schwert 1977; and others). However, there is no previous study which examines the relationship between stock returns and the real auto production.

Some might argue that weak relationship between non-stock market indicators and the common factor is due to multicollinearities among the indicators. Therefore, six of the 11 indicators are selected to represent each category indicator. They are the market portfolio return (RM), the transaction volume (VL), real riskless interest rate (RF), CPI inflation rate (CPI), money supply (M2), and the real total industrial production (IP). This is the one-factor six-indicator model. The result of this model is shown in Table 36.4a column 2. The result of this one-factor six-indicator model is very close to that of the one-factor 11-indicator model. As mentioned before, only the stock market-related indicators are significantly correlated with the common factor. Real riskless interest rate, inflation, money supply, and real production are all negatively but insignificantly related to the common factor. For factor equation, the 11-indicator model has only a little high R-square than the six-indicator model. They are 0.5611 and 0.5412, respectively. Comparing the betas of these two one-factor models in Table 36.4a, they are very highly correlated with a correlation coefficient of about 1.000. In addition, the average R-square of each return equation in both one-factor models is the same with a value of 0.811. Up to this point, there is not much difference between the one-factor 11-indicator and the one-factor six-indicator models. Later on, the cross-sectional regression will be used to double check this result.

Even though we have already discussed that the appropriate model is a one-factor model for period 1 by the scree test of factor analysis, we want to use a two-factor model to double check whether the second factor is significant or not. A predetermined two-factor six-indicator model will be used to test APT. Factor analysis is employed to classify these six indicators into two groups. The first group includes the market portfolio and the transaction volume, while the second group includes other four indicators: real riskless interest rate (RF), inflation rate (CPI), money supply (M2), and real total industrial production (IP). The two-factor result shown in Table 36.4a columns 3 and 4 displays that only the first factor is significantly related to RM and VL, whereas the second factor is very insignificantly correlated with the second group indicators. This is also evident by examining from the second betas in the table. All of the second betas are insignificant. Furthermore, the first beta coefficient is very highly correlated with the betas in both one-factor models with both correlation coefficients of 0.996. This is further

Table 36.4 Structural coefficients of APT in the MIMIC model

Period 1: 1963–1972				
a's Coefficients				
	One-factor 11-indicator	One-factor 6-indicator	Two-factor 6-indicator	
Indicator	$f1$	$f1$	$f1$	$f2$
RM	0.908(8.08)*	0.925(7.55)*	0.923(8.85)*	
VL	0.040(2.55)*	0.043(2.60)*	0.040(2.28)*	
RF	−17.24(−1.2)	−3.451(−0.76)		−0.300(–)
MP	−15.84(−0.90)			
DP	1.693(0.12)			
CPI	−16.51(−1.2)	−4.596(−1.1)		−0.119(−0.12)
M2	0.039(0.03)	−0.344(−0.31)		0.458(0.58)
PI/M2	22.23(1.29)			
IP	0.141(0.20)	−0.469(−1.1)		−0.054(−0.06)
IPA	−0.078(−1.3)			
IPH	0.189(0.92)			
R-square	0.5611	0.5412	0.5912	0.0774
b's Coefficients				
	One-factor 11-indicator	One-factor 6-indicator	Two-factor 6-indicator	
Industry	$f1$	$f1$	$f1$	$f2$
1	1.000(–)	0.974(12.1)*	1.000(–)	0.410(0.06)
2	0.834(16.5)*	0.812(13.6)*	0.857(16.3)*	−0.555(−0.06)
3	1.210(16.2)*	1.178(13.4)*	1.211(16.1)*	0.390(−0.06)
4	1.020(14.7)*	0.993(12.5)*	1.059(14.4)*	−1.075(−0.06)
5	0.939(17.4)*	0.914(14.1)*	0.942(17.4)*	0.270(0.06)
6	0.785(10.5)*	0.764(9.62)*	0.831(10.3)*	−1.575(−0.06)
7	1.097(15.3)*	1.068(12.9)*	1.122(15.2)*	−0.555(−0.06)
8	1.183(15.8)*	1.151(13.2)*	1.190(15.7)*	0.120(0.05)
9	1.146(18.1)*	1.115(14.4)*	1.126(17.6)*	1.205(0.06)
10	1.221(18.2)*	1.188(14.5)*	1.192(17.4)*	1.670(0.06)
11	1.389(16.6)*	1.352(13.6)*	1.335(15.0)*	2.875(0.06)
12	1.264(17.7)*	1.231(14.2)*	1.236(17.0)*	1.675(0.06)
13	1.302(17.2)*	1.268(13.9)*	1.268(16.2)*	1.995(0.06)
14	1.218(13.2)*	1.186(11.5)*	1.215(13.0)*	0.485(0.06)
15	1.435(14.0)*	1.396(12.1)*	1.405(13.5)*	1.760(0.06)
16	0.463(7.66)*	0.450(7.30)*	0.527(7.33)*	−2.340(−0.06)
17	1.094(14.1)*	1.065(12.2)*	1.112(14.1)*	−0.265(−0.06)
18	1.078(16.5)*	1.050(13.6)*	1.094(16.4)*	−0.180(−0.06)
19	1.012(14.7)*	0.985(12.5)*	1.056(14.3)*	−1.260(−0.06)

(*continued*)

Table 36.4 (continued)

Period 2: 1973–1982			
a's Coefficients			
	One-factor 11-indicator	Two-factor 6- indicator	
Indicator	$f1$	$f1$	$f2$
RM	0.859(6.38)*	1.000(6.67)*	
VL	0.073(3.34)*	0.075(2.83)*	
RF	−5.586(−1.1)		−5.600(–)
MP	−2.254(−0.40)		
DP	−7.417(−0.70)		
CPI	−9.786(−1.6)*		1.566(0.32)
M2	−1.820(−1.2)		1.810(0.72)
PI/M2	28.60(1.16)		
IP	−0.290(−0.42)		
IPA	−0.142(−1.6)*		−0.368(−1.7)*
IPH	0.065(0.20)		
R-square	0.5811	0.5525	0.1836
b's Coefficients			
	One-factor 11-indicator	Two-factor 6-indicator	
Industry	$f1$	$f1$	$f2$
1	0.845(8.85)*	0.993(7.86)*	0.502(1.95)*
2	0.786(13.4)*	0.724(8.39)*	−0.184(−1.7)*
3	1.073(13.3)*	0.948(7.98)*	−0.379(−1.9)*
4	0.884(13.2)*	0.850(14.4)*	−0.081(−1.0)
5	0.829(13.5)*	0.804(8.81)*	−0.054(−0.77)
6	0.676(7.96)*	0.797(7.41)*	0.408(1.93)*
7	1.026(14.2)*	0.948(8.62)*	−0.218(−1.7)*
8	0.898(12.4)*	0.896(8.67)*	0.029(0.38)
9	1.083(14.8)*	1.026(8.97)*	−0.146(−1.4)
10	1.082(14.5)*	1.054(9.11)*	−0.053(−0.63)
11	1.171(14.7)*	1.098(8.87)*	−0.192(−1.6)
12	1.139(14.4)*	1.058(8.72)*	−0.222(−1.7)*
13	1.188(14.7)*	1.115(8.86)*	−0.208(−1.6)
14	0.850(11.0)*	0.869(8.33)*	0.089(0.99)
15	1.186(13.0)*	1.088(8.21)*	−0.288(−1.7)*
16	0.534(9.76)*	0.506(7.32)*	−0.076(−1.2)
17	1.122(12.8)*	0.961(7.55)*	−0.506(−2.0)*
18	1.113(14.3)*	1.001(8.40)*	−0.341(−1.9)*
19	1.042(13.1)*	0.959(8.29)*	−0.242(−1.7)*

This table presents the structural coefficients of the APT in the MIMIC model for period 1 and period 2. The structural model is written as

$$\tilde{r}_i = b_{i1}\tilde{f}_1 + b_{i2}\tilde{f}_2 + \tilde{u}_i, \qquad i = 1, \ldots, 19$$
$$\tilde{f}_j = a_{j1}(RM) + a_{j2}(VL) + a_{j3}(RF) + a_{j4}(MP) + a_{j5}(DP) + a_{j6}(CPI) + a_{j7}(M2) + a_{j8}(PI/M2) + a_{j9}(IP) + a_{j10}(IPA) + a_{j11}(IPH) + \tilde{e}_j, \qquad j = 1 \text{ or } 1, 2.$$

Variables used as indicators in the structural model are the market return (*RM*), transaction volume (*VL*), real risk interest rate (*RF*), maturity risk premium (*MP*), default risk premium (*DP*), CPI inflation rate (*CPI*), money supply (*M2*), velocity of money supply (*PI*/*M2*), real total industrial production (*IP*), real auto production (*IPA*), and real home production (*IPH*). *represents significant at the 5 % level supported by the R-square criteria. The average R-square of each return equation in the one-factor model is 0.811, while 0.844 in the two-factor model. If the adjusted R-square is used as the criteria, the increase in R-square will be insignificant. Thus, a one-factor APT is good enough to explain the first period data. We will reconfirm this argument by the cross-sectional regression.

Table 36.5 Panel A reports the return-risk cross-sectional relationship in period 1 for APT in the MIMIC model. Both one-factor models have a very similar result. Their adjusted R-squares are the same of 0.258. The intercepts are both insignificantly different from zero but positive (recall that the LHS variable is excess return). The factor risk premia for both one-factor models are positive and significant. This is exactly the result that should have been concluded from APT. On the other hand, the intercept in the two-factor model is much higher than those in the one-factor model, and the two factor risk premia are both insignificant.[8] The adjusted R-square is a little lower than those in the one-factor models. From the results of the MIMIC model and the cross-sectional regressions, it is probable that one-factor APT with the market portfolio and the transaction volume as the determinants of the single factor is the appropriate pricing model for period 1. Comparing the MIMIC CAPM and CAPM in Table 36.2 with this one-factor model, the MIMIC CAPM is closer to the one-factor model. In addition, the factor risk premium is closer to the realized market risk premium than that in the MIMIC CAPM model.

Now, let us examine APT in period 2. The structural coefficients of APT in the MIMIC model for period 2 are represented in Table 36.4b. Because six indicators do not make much difference from 11 indicators in the one-factor model in period 1, only the 11-indicator model is used to test APT for the one-factor model in period 2 in order to save the computer time. From the one-factor model, the result is similar to that in period 1. The stock market return and the transaction volume are the most significant indicators. However, inflation and the real auto production are also significant even just at the marginal level. This reinforces our suspicion that there may be more than one factor in period 2.

Let us closely examine the indicators with an absolute t-value greater than one. Real interest riskless rate, inflation, money supply, and real auto production are all negatively correlated with the common factor, while the velocity of money supply is positively correlated with this common factor. Overall, this result again supports previous studies. Now, let us turn to the two-factor six-indicator model. Follow the same procedures done for period 1. The result is displayed in Table 36.4b columns 2 and 3. The market portfolio (RM) and the transaction volume (LV) are significantly related to the first factor. However, the real auto production is significantly correlated with the second factor; this time even only at the marginal level. Comparing other indicators with those in period 1, real riskless interest rate is again negatively related to the second factor, but inflation rate and money supply are positively related to the second factor in the period. From the result of the relationship between the factors and the indicators, the second factor is still important even though less important than

[8]The loss of the significance of the first factor risk premium is due to the multicollinearity problem.

Table 36.5 Return-risk cross-sectional relationships of APT in the MIMIC model: 1963–1982

	$\hat{a}_0$	$\hat{a}_1$	$\hat{a}_2$	$\overline{R}^2$
Model	Panel A: Period 1, 1963–1972			
One-factor	0.159	0.543*		0.258
11-indicator	(0.710)	(2.693)		
One-factor	0.159	0.558*		0.258
6-indicator	(0.710)	(2.692)		
Two-factor	0.507	0.206	0.071	0.257
6-indicator	(1.219)	(0.529)	(1.221)	
	Panel B: Period 2, 1973–1982			
One-factor	0.370	0.285		−0.002
11-indicator	(1.328)	(0.967)		
Two-factor	0.040	0.684*	0.362*	0.215
6-indicator	(0.150)	(2.377)	(2.066)	

This table presents the return-risk cross-sectional relationship for the APT in the MIMIC model. The cross-sectional relationship between return and risk can be written as $\overline{R}_i = \hat{a}_0 + \hat{a}_1 b_1 + \hat{a}_2 b_2 + \widetilde{e}_i$, $i=1,\ldots,19$. Panel A reports the relationship in period 1 and Panel B reports the relationship in period 2. The average monthly returns are multiplied by 100 before regressing. *represents significant at the 5 % level

the first factor for period 2. This can also be checked by the significance of betas in Table 36.4b. Ten out of the 19 s factor betas are significant although the relationship is not as strong as those in the first factor betas. Because the first factor is correlated with stock market-related indicator, this factor can be regarded as a proxy of the market portfolio. The correlation coefficient between the beta in the first factor and the beta in the MIMIC CAPM is very high with a coefficient of 0.923. Further, the average R-square of each return equation in the one-factor model is 0.818, while 0.885 in the two-factor model. When the adjusted R-square is used as a criteria, the increase in R-square should not be trivial and will be significant. In addition, the R-square for the second factor equation is as high as 0.1836. All of these results indicate that the second factor should be important. We further check whether the second factor is important or not by the cross-sectional regression.

The risk-return cross-sectional relationship of APT in the MIMIC model for period 2 is shown in Table 36.5 Panel B. For the one-factor model, both the intercept and the factor risk premium are insignificant. The intercept is too high with an annual rate of 4.3 %, while the factor risk premium is far below the realized market risk premium with a monthly rate of 0.636 %. The adjusted R-square is negative. On the other hand, the intercept for the two-factor model is insignificant and very near to zero, while the two factor premia are significant with the first factor premium very close to the realized market risk premium. The adjusted R-square is as high as 0.215. The results of APT in the MIMIC model confirm the scree test of the factor analysis and the poor performance of CAPM in previous section for period 2. Furthermore, comparing the result in Table 36.5 Panel A with that in Table 36.2, we can conclude that the two-factor APT outperforms the one-factor APT, the MIMIC CAPM, and CAPM in period 2. But the one-factor APT does a better job than the two-factor APT in period 1, and there is not much difference among the one-factor APT, the MIMIC CAPM, and CAPM. This evidence supports Ross's argument that APT is more general than CAPM because APT allows more than one factor in the pricing relation.

It will be interesting to see what will happen if market variables (the market portfolio and transaction volume) are excluded from the model. Table 36.6 shows the structural coefficients of the one-factor APT without market variables.[9] The *b* coefficients (factor loadings) shown in Table 36.6 for both periods are very close to those estimated from the one-factor APT with market variables shown in Tables 36.4a and 36.4b, respectively.[10] However, the results from the factor equation have dramatically changed.

For period 1, the *R*-square is only 0.0611 which is much lower than 0.5611 from the one-factor APT with market variables. This result indicates that the market variables play the major role in the pricing behavior during period 1. When the market variables excluded from the model, among nine indicators, only the two money supply variables, M2 and PI/M2, are positively related to the unique common factor at the marginal level. For period 2, the *R*-square of the factor equation for the model without market variables is 0.3915 which is higher than the one in period 1. This evidence denotes that, in addition to the market variables, some other macroeconomic indicators play a relatively important role in the pricing behavior during period 2. Among the nine nonmarket variables, real risk-free rate (RF), default risk premium (DF), and inflation (CPI) are significantly negatively related to the common factor, while the velocity of money supply (PI/M2) and the real auto production (IPA) are significantly positively related to the common factor. All of the relationships are as we expect. The results from the model without market variables also confirm our previous evidence that the 1963–1972 data can be described by a one-factor APT with the market variables as its indicators, while the 1973–1982 data should be explained by more than one-factor APT.

36.5 Concluding Remarks

This paper bases on parts of Wei's (1984) dissertation using a MIMIC approach to test CAPM and APT. The results support the conclusion that APT outperforms CAPM, especially for the period from 1973 to 1982. The beta estimated from the MIMIC model by allowing measurement error on the market portfolio does not significantly improve the OLS beta. However, the MLE estimator does a better job than the OLS and GLS estimators in the cross-sectional regressions because the MLE estimator takes care of the measurement error in beta. Therefore, the measurement error problem

[9]Only the one-factor APT is used to investigate the difference between the models shown in Table 36.4 and in Table 36.6.

[10]If we normalize the one-factor 11-indicator model for period 2 shown in Table 36.4b by setting $b_1 = 1.00$, it is easily seen that the *b* coefficients of one-factor model shown in Tables 36.4b and 36.6 column 2 are very similar.

Table 36.6 Structural coefficients of the one-factor APT in the MIMIC model without market variables: 1963–1982

a's Coefficients		
Indicator	*1963–1972*	*1973–1982*
RF	−25.541(−1.26)	−18.666(−3.39)*
MP	−26.061(−1.18)	−1.067(−0.17)
DP	−4.350(−0.23)	−30.745(−2.68)*
CPI	−25.989(−1.32)	−23.146(−3.52)*
M2	3.652(1.85)*	−1.630(−0.98)
PI/M2	45.126(1.91)*	110.380(4.22)*
IP	−0.932(−0.97)	0.124(0.16)
I PA	0.046(0.57)	0.214(2.18)*
IPH	0.298(1.06)	−0.120(−0.33)
R-square	0.0611	0.3915
b's Coefficients		
Industry	*1963–1972*	*1973–1982*
1	1.000(–)	1.000(–)
2	0.833(16.45)*	0.930(10.10)*
3	1.209(16.20)*	1.270(10.03)*
4	1.019(14.70)*	1.045(9.99)*
5	0.939(17.43)*	0.981(10.11)*
6	0.785(10.53)*	0.800(9.64)*
7	1.097(15.33)*	1.214(10.41)*
8	1.183(15.79)*	1.063(9.64)*
9	1.146(18.11)*	1.283(10.63)*
10	1.221(18.25)*	1.282(10.53)*
11	1.389(16.58)*	1.386(10.60)*
12	1.265(17.74)*	1.349(10.50)*
13	1.302(17.16)*	1.407(10.63)*
14	1.218(13.19)*	1.006(8.96)*
15	1.436(14.05)*	1.404(9.92)*
16	0.463(7.66)*	0.632(8.23)*
17	1.094(14.13)*	1.329(9.84)*
18	1.078(16.46)*	1.318(10.47)*
19	1.012(14.70)*	1.233(9.95)*

This table shows estimated coefficients of the one-factor APT in the MIMIC model without market variables. The structural model is shown as follows. *represents significant at the 5 % level

$$\widetilde{r}_i = b_{i1}\widetilde{f}_1 + b_{i2}\widetilde{f}_2 + \widetilde{u}_i, \qquad i = 1, \ldots, 19$$

$$\begin{aligned}\widetilde{f}_j = {} & a_1(RF) + a_2(MP) + a_3(DP) \\ & + a_4(CPI) + a_5(M2) + a_6(PI/M2) + a_7(IP) \\ & + a_8(IPA) + a_9(IPH) + \widetilde{e}_j,\end{aligned}$$

on beta is more serious than that on the market portfolio. This evidence supports Stambough's (1982) argument that the inference about the tests of CAPM is insensitive to alternative market indexes. When the one-factor APT with market variables is compared with the model without market variables, we found that the market variables play a major role in pricing behavior. Therefore, we conclude that it is inappropriate for the study of the relationship between the common factors extracted from APT and the macroeconomic variables without including the market variables.

References

Attfield, C. (1983). Consistent estimation of certain parameters in the unobservable variable model when there is a specification error. *Review of Economics and Statistics, 65*, 164–167.

Bodie, Z. (1976). Common stocks as a hedge against inflation. *Journal of Finance, 31*, 459–740.

Brigham, S. (1983). *Fundamentals of financial management* (3rd ed.). Chicago: Dryden Press.

Brigham, E. F., & Houston, J. F. (2012). *Fundamentals of financial management* (13th ed.). Singapore: Cengage Learning.

Chang, C., Lee, A. C., & Lee, C. F. (2009). Determinants of capital structure choice: A structural equation modeling approach. *Quarterly Review of Economics and Finance, 49*, 197–213.

Chen, C. (1981). The EM approach to the multiple indicators and multiple causes model via the estimation of the latent variable. *Journal of the American Statistical Association, 76*, 702–708.

Chen, N. F., Roll, R., & Ross, S. A. (1983). *Economic forces and the stock market: Testing the APT and alternative asset pricing theories*. Working paper, University of Chicago.

Cohn, R., & Lessard, D. (1981). Are markets efficient? Tests of alternative hypothesis. On the effect of inflation on stock prices: International evidence. *Journal of Finance, 36*, 277–289.

Fama, E. (1981). Stock returns, real activity, inflation, and money. *The American Economic Review, 71*, 545–565.

Fama, E., & Schwert, G. (1977). Asset returns and inflation. *Journal of Financial Economics, 5*, 115–146.

Fogler, H. R., John, K., & Tipton, J. (1981). Three factor interest rate differentials and stock groups. *Journal of Finance, 36*, 323–336.

Friend, I., & Westerfield, R. (1980). Co-skewness and capital asset pricing. *Journal of Finance, 35*, 897–913.

Gargett, D. (1978). The link between stock prices and liquidity. *Financial Analysis Journal, 34*, 50–54.

Geske, R., & Roll, R. (1983). The fiscal and monetary linkage between stock returns and inflation. *Journal of Finance, 38*, 1–33.

Goldberger, A. (1972a). Maximum-likelihood estimation of regressions containing unobservable independent variables. *International Economic Review, 13*, 1–15.

Goldberger, A. (1972b). Structural equation methods in the social sciences. *Econometrica, 40*, 979–1001.

Granger, C., & Newbold, P. (1974). Spurious regression in econometrics. *Journal of Econometrics, 2*, 111–120.

Gultekin, N. (1983). Stock market returns and inflation: Evidence from other countries. *Journal of Finance, 38*, 49–65.

Hamburger, M., & Kochin, L. (1972). Money and stock prices: The channels of influence. *Journal of Finance, 27*, 231–249.

Homa, K., & Jaffee, D. (1971). The supply of money and common stock prices. *Journal of Finance, 26*, 1045–1066.

Joreskog, K., & Goldberger, A. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association, 70*, 631–639.

Joreskog, K., & Sorbom, D. (1981). *LISREL V: Analysis of linear structural relationships by the method of maximizing likelihood, user's guide*. Chicago: National Educational Resources.

Keran, M. (1971). Expectations, money, and stock market. *Review, Federal Reserve Bank of St. Louis, 53*, 16-31.

Lahiri, K. (1976). Inflationary expectations: Their formation and interest rate effects. *The American Economic Review, 66*, 124–131.

Lintner, J. (1965). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics, 47*, 13–37.
Litzenberger, R., & Ramaswamy, K. (1979). The effects of personal taxes and dividends on capital asset prices: Theory and empirical evidence. *Journal of Financial Economics, 7*, 163–195.
Maddala, G. S., & Nimalendran, M. (1996). Error-in-variables problems in financial models. In G. S. Maddala & C. R. Rao (Eds.), *Handbook of statistics*. New York: Elsevier Science.
Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica, 34*, 768–783.
Nelson, C. (1976). Inflation and rate of returns on common stocks. *Journal of Finance, 31*, 471–483.
Nelson, C. (1979). Recursive structure in U.S. income, prices and output. *Journal of Political Economy, 87*, 1307–1327.
Pesando, J. (1974). The supply of money and common stock prices: Further observations on econometric evidence. *Journal of Finance, 29*, 909–921.
Roll, R. (1977). A critique of the asset pricing theory's tests. *Journal of Financial Economics, 4*, 129–176.
Roll, R., & Ross, S. (1980). An empirical investigation of the arbitrage pricing theory. *Journal of Finance, 35*, 1073–1103.
Ross, S. (1976). The arbitrage theory of capital asset pricing. *Journal of Economic Theory, 13*, 341–360.
Ross, S. (1977). Return, risk, and arbitrage. In I. Friend & J. L. Bicksler (Eds.), *Risk and return in finance* (Vol. I). Ballinger Pub. Co.: Cambridge, MA.
Schipper, K., & Thompson, R. (1981). Common stocks as hedges against shifts in the consumption on investment opportunity set. *Journal of Business, 54*, 305–328.
Shanken, J. (1982). The arbitrage pricing theory: Is it testable? *Journal of Finance, 37*, 1129–1140.
Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance, 19*, 425–442.
Solnik, B. (1983). The relation between stock prices and inflationary expectations: The international evidence. *Journal of Finance, 38*, 35–48.
Stambough, R. (1982). On the exclusion of assets from tests of the two-parameter model: A sensitivity analysis. *Journal of Financial Economics, 10*, 237–268.
Titman, S., & Wessels, R. (1988). The determinants of capital structure choice. *Journal of Finance, 43*, 1–19.
Turnovsky, S. (1970). Empirical evidence on the formation of price expectations. *Journal of the American Statistical Association, 65*, 1441–1454.
Wei, K. (1984). *The arbitrage pricing theory versus the generalized intertemporal capital asset pricing model: Theory and empirical evidence*. Unpublished Ph.D. dissertation, University of Illinois.
Yang, C. C., Lee, C. F., Gu, Y. X., & Lee, Y. W. (2009). Co-determinantion of capital structure and stock returns – A LISREL approach: An empirical test of Taiwan stock markets. *Quarterly Review of Economics and Finance, 50*, 222–233.
Zellner, A. (1970). Estimation of regression relationships counting un-observable variables. *International Economic Review, 11*, 441–454.

Binomial OPM, Black–Scholes OPM, and Their Relationship: Decision Tree and Microsoft Excel Approach

37

John C. Lee

Contents

Abstract

This chapter will first demonstrate how Microsoft Excel can be used to create the decision trees for the binomial option pricing model. At the same time, this chapter will discuss the binomial option pricing model in a less mathematical fashion. All the mathematical calculations will be taken care by the Microsoft Excel program that is presented in this chapter. Finally, this chapter also uses the decision tree approach to demonstrate the relationship between the binomial option pricing model and the Black–Scholes option pricing model.

J.C. Lee
Center for PBBEF Research, North Brunswick, NJ, USA
e-mail: johnleejohnlee@yahoo.com; leeleeassociates@gmail.com

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_37,

Keywords
Binomial option pricing model • Decision trees • Black–Scholes option pricing model • Call option • Put option • Microsoft Excel • Visual Basic for Applications • VBA • Put-call parity • Sigma • Volatility • Recursive programming

37.1 Introduction

The binomial option pricing model derived by Rendleman and Barter (1979) and Cox et al. (1979) is one the most famous models used to price options. Only the Black–Scholes model (1973) is more famous. One problem with learning the binomial option pricing model is that it is computationally intensive. This results in a very complicated formula to price an option.

The complexity of the binomial option pricing model makes it a challenge to learn the model. Most books teach the binomial option model by describing the formula. This is not very effective because it usually requires the learner to mentally keep track of many details, many times to the point of information overload. There is a well-known principle in psychology that the average number of things that a person can remember at one time is seven.

This chapter will first demonstrate the power of Microsoft Excel. It will do this by demonstrating that it is possible to create large decision trees for the binomial pricing model using Microsoft Excel. A ten-period decision tree would require 2,047 call calculations and 2,047 put calculations. This chapter will also show the decision tree for the price of a stock and the price of a bond, each requiring 2,047 calculations. Therefore, there would be 2,047 * 4 = 8,188 calculations for a complete set of ten-period decision trees.

Secondly, this chapter will present the binomial option model in a less mathematical matter. It will try to make it so that the reader will not have to keep track of many things at one time. It will do this by using decision trees to price call and put options.

Finally, this chapter will show the relationship between the binomial option pricing model and the Black–Scholes option pricing model.

This chapter uses a Microsoft Excel workbook called *binomialBS_OPM.xls* that contains the VBA code to create the decision trees for the binomial option pricing model. The VBA code is published in Appendix 1. The password for the workbook is *bigsky* for those who want to study the VBA code. E-mail me at JohnLeeExcelVBA@gmail.com and indicate that the password is "bigsky" to get a copy of this Microsoft Excel workbook.

Section 37.2 discusses the basic concepts of call and put options. Section 37.3 demonstrates the one-period call and put option pricing models. Section 37.4 presents the two-period option pricing model. Section 37.5 demonstrates how to use the Microsoft Excel workbook *binomialBS_OPM.xls* to create the decision trees for an *n*-period binomial option pricing model. Section 37.6 demonstrates the use

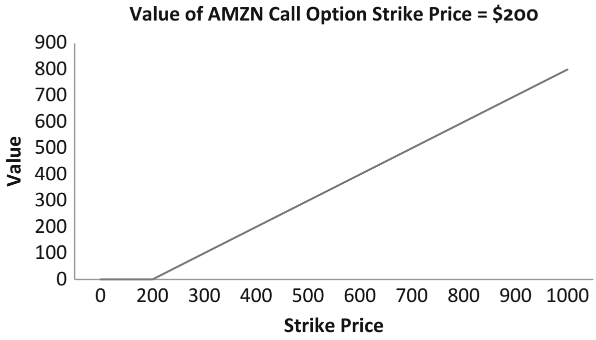


Fig. 37.1 Value of AMZN call option

of the Black–Scholes model. Section 37.7 shows the relationship between the binomial option pricing model and the Black–Scholes option pricing model. Section 37.8 demonstrates how to use the Microsoft Excel workbook *binomialBS_OPM.xls* to demonstrate the relationship between the binomial option pricing model and the Black–Scholes option pricing model.

37.2 Call and Put Options

A *call option* gives the owner the right but not the obligation to buy the underlying security at a specified price. The price in which the owner can buy the underlying price is called the *exercise price*. A call option becomes valuable when the exercise price is less than the current price of the underlying stock price.

For example, a call option on an AMZN stock with an exercise price of $200 when the stock price of an Amazon stock is $250 is worth $50. The reason it is worth $50 is because a holder of the call option can buy the AMZN stock at $200 and then sell the AMZN stock at the prevailing price of $250 for a profit of $50. Also, a call option on an AMZN stock with an exercise price of $300 when the stock price of an AMZN stock is $150 is worth $0.

A *put option* gives the owner the right but not the obligation to sell the underlying security at a specified price. A put option becomes valuable when the exercise price is more than the current price of the underlying stock price.

For example, a put option on an AMZN stock with an exercise price of $200 when the stock price of an AMZN stock is $150 is worth $50. The reason it is worth $50 is because a holder of the put option can buy the AMZN stock at the prevailing price of $150 and then sell the AMZN stock at the put price of $200 for a profit of $50. Also, a put option on an AMZN stock with an exercise price of $200 when the stock price of the AMZN stock is $250 is worth $0.

Figures 37.1 and 37.2 are charts showing the value of call and put options of the above GE stock at varying prices.

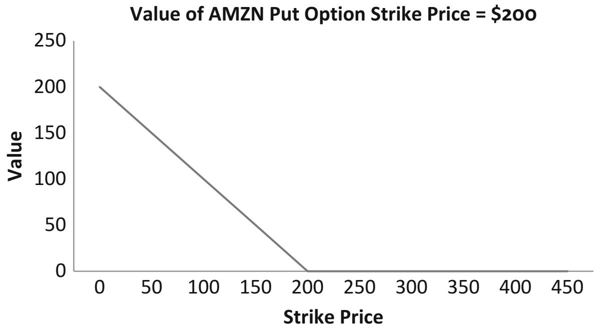


Fig. 37.2 Value of AMZN put option

37.3 One-Period Option Pricing Model

What should be the value of these options? Let's look at a case where we are only concerned with the value of options for one period. In the next period, a stock price can either go up or go down. Let's look at a case where we know for certain that an AMZN stock with a price of $200 will either go up 5 % or go down 5 % in the next period and the exercise after one period is $200. Figures 37.3, 37.4, and 37.5 show the decision tree for the AMZN stock price, the AMZN call option price, and the AMZN put option price, respectively.

Let's first consider the issue of pricing an AMZN call option. Using a one-period decision tree, we can illustrate the price of an AMZN stock if it goes up 5 % and the price of a stock AMZN if it goes down 5 %. Since we know the possible ending values of the AMZN stock, we can derive the possible ending values of a call option. If the stock price increases to $210, the price of the AMZN call option will then be $10 ($210 − $200). If the AMZ stock price decreases to $190, the value of the call option will worth $0 because it would be below the exercise price of $200. We have just discussed the possible ending value of an AMZN call option in period 1. But what we are really interested in is what the value is now of the AMZN call option knowing the two resulting values of the AMZN call option.

To help determine the value of a one-period AMZN call option, it's useful to know that it is possible to replicate the resulting two states of the value of the AMZN call option by buying a combination of stocks and bonds. Below is the formula to replicate the situation where the price increases to $210. We will assume that the interest rate for the bond is 3 %:

$$210S + 1.03B = 10$$
$$190S + 1.03B = 0$$

We can use simple algebra to solve for both S and B. The first thing that we need to do is to rearrange the second equation as follows:

$$1.03B = -190S$$

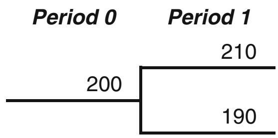


Fig. 37.3 AMZN stock price

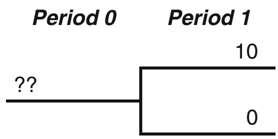


Fig. 37.4 AMZN call option price

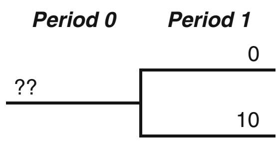


Fig. 37.5 AMZN put option price

With the above equation, we can rewrite the first equation as

$$\begin{aligned} 210\text{S} + (-190\text{S}) &= 10 \\ 20\text{S} &= 10 \\ \text{S} &= 0.5 \end{aligned}$$

We can solve for B by substituting the value 0.05 for S in the first equation:

$$\begin{aligned} 210(0.5) + 1.03\text{B} &= 10 \\ 105 + 1.03\text{B} &= 10 \\ 1.03\text{B} &= -95 \\ \text{B} &= -92.23 \end{aligned}$$

Therefore, from the above simple algebraic exercise, we should at period 0 buy 0.05 shares of AMZN stock and borrow 9.223 at 3 % to replicate the payoff of the AMZN call option. This means the value of an AMZN call option should be $0.5 * 200 - 92.23 = 7.77$.

If this were not the case, there would then be arbitrage profits. For example, if the call option were sold for \$30, there would be a profit of 22.23. This would result in the increase in the selling of the AMZN call option. The increase in the supply of AMZN call options would push the price down for the call options. If the call option were sold for \$5, there would be a saving of \$2.77. This saving would result in the increase demand for the AMZN call option. This increase demand would result in the price of the call option to increase. The equilibrium point would be \$7.77.

Using the above mentioned concept and procedure, Benninga (2000) has derived a one-period call option model as

$$C = q_u \mathrm{Max}[S(1+u)X, 0] + q_d \mathrm{Max}[S(1+d) - X, 0] \tag{37.1}$$

where

$$q_u = \frac{i-d}{(1+i)(u-d)}$$
$$q_d = \frac{u-i}{(1+i)(u-d)}$$
$$u = \text{increase factor}$$
$$d = \text{down factor}$$
$$i = \text{interest rate}$$

If we let $i = r$, $p = (r - d)/(u - d)$, $1 - p = (u - r)/(u - d)$, $R = 1/(1 + r)$, $C_u = \mathrm{Max}[S(1 + u) - X, 0]$, and $C_d = \mathrm{Max}[S(1 + d) - X, 0]$, then we have

$$C = [pC_u + (1-p)C_d]/R, \tag{37.2}$$

where
Cu = call option price after increase
Cd = call option price after decrease

Equation 37.2 is identical to Eq. 6B.6 in Lee et al. (2000, p. 234).[1]

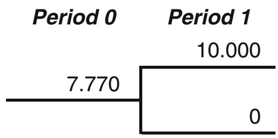
Below calculates the value of the above one-period call option where the strike price, X, is \$200 and the risk-free interest rate is 3 %. We will assume that the price of a stock for any given period will either increase or decrease by 5 %:

$$X = \$200$$
$$S = \$200$$
$$u = 1.05$$
$$d = 0.95$$
$$R = 1 + r = 1 + 0.03$$
$$p = (1.03 - 0.95)/(1.05 - 0.95)$$
$$C = [0.8(10) + 0.2(0)]/1.03 = \$7.77$$

Therefore, from the above calculations, the value of the call option is \$7.77. Figure 37.6 shows the resulting decision tree for the above call option.

[1]Please note that in Lee et al. (2000, p. 234) u = 1 + percentage of price increase and d = 1 – percentage of price increase.

Fig. 37.6 Call option price



Like the call option, it is possible to replicate the resulting two states of the value of the put option by buying a combination of stocks and bonds. Below is the formula to replicate the situation where the price decreases to \$190:

$$\begin{aligned} 210\mathrm{S} + 1.03\mathrm{B} &= 0 \\ 190\mathrm{S} + 1.03\mathrm{B} &= 10 \end{aligned}$$

We will use simple algebra to solve for both S and B. The first thing we will do is to rewrite the second equation as follows:

$$1.03\mathrm{B} = 10 - 190\mathrm{S}$$

The next thing to do is to substitute the above equation to the first put option equation. Doing this would result in the following:

$$210\mathrm{S} + 10 - 190\mathrm{S} = 0$$

The following solves for S:

$$\begin{aligned} 20\mathrm{S} &= -10 \\ \mathrm{S} &= -0.5 \end{aligned}$$

Now let us solve for B by putting the value of S into the first equation. This is shown below:

$$\begin{aligned} 210(-0.5) + 1.03\mathrm{B} &= 0 \\ 1.03\mathrm{B} &= 105 \\ \mathrm{B} &= 101.94 \end{aligned}$$

From the above simple algebra exercise, we have S = −0.5 and B = 101.94. This tells us that we should in period 0 lend \$101.94 at 3 % and sell 0.5 shares of stock to replicate the put option payoff for period 1. And the value of the AMZN put option should be 200(−0.5) + 101.94 = 1.94.

Using the same arbitrage argument that we used in the discussion of the call option, 0.194 has to be the equilibrium price of the put option.

As with the call option, Benninga (2000) has derived a one-period put option model as

$$P = q_u \mathrm{Max}[X - S(1+u), 0] + q_d \mathrm{Max}[X - S(1+d), 0] \tag{37.3}$$

where

$$q_u = \frac{i-d}{(1+i)(u-d)}$$
$$q_d = \frac{u-i}{(1+i)(u-d)}$$
$$u = \text{increase factor}$$
$$d = \text{down factor}$$
$$i = \text{interest rate}$$

If we let $i = r$, $p = (r - d)/(u - d)$, $1 - p = (u - r)/(u - d)$, $R = 1/(1 + r)$, $P_u = \text{Max}[X - S(1 + u), 0]$, and $P_d = \text{Max}[X - S(1 + d), 0]$, then we have

$$P = [pP_u + (1-p)P_d]/R, \quad (37.4)$$

where
P_u = put option price after increase
P_d = put option price after decrease

Below calculates the value of the above one-period put option where the strike price, X, is \$20 and the risk-free interest rate is 3 %:

$$P = [0.8(0) + 0.2(10)]/1.03 = \$1.94$$

From the above calculation, the put option pricing decision tree would look like the following.

Figure 37.7 shows the resulting decision tree for the above put option.

There is a relationship between the price of a put option and the price of all call option. This relationship is called the put-call parity. Equation 37.5 shows the relationship between the price of a put option and the price of a call option:

$$\mathrm{P} = \mathrm{C} + \mathrm{X}/\mathrm{R} - \mathrm{S} \quad (37.5)$$

where
C = call price
X = strike price
R = 1 + interest rate
S = stock price

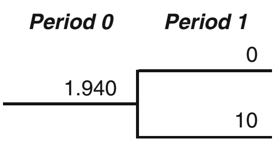


Fig. 37.7 AMZN put option price

The following uses the put-call parity to calculate the price of the AMZN put option:

$$\begin{aligned} P &= \$7.77 + \$200/(1.03) - \$200 \\ &= 7.77 + 194.17 - 200 \\ &= 1.94 \end{aligned}$$

37.4 Two-Period Option Pricing Model

We now will look at pricing options for two periods. Figure 37.8 shows the stock price decision tree based on the parameters indicated in the last section. This decision tree was created based on the assumption that a stock price will either increase by 5 % or decrease by 5 %.

How do we price the value of a call and put option for two periods?

The highest possible value for our stock based on our assumption is \$220.5. We get this value first by multiplying the stock price at period 0 by 105 % to get the resulting value of \$210 of period 1. We then again multiply the stock price in period 1 by 105 % to get the resulting value of \$220.5. In period two, the value of a call option when a stock price is \$220.5 is the stock price minus the exercise price, \$220.5 − \$200, or \$20.5. In period two, the value of a put option when a stock price is \$220.5 is the exercise price minus the stock price, \$200 − \$220.5, or −\$20.5. A negative value has no value to an investor so the value of the put option would be \$0.

The lowest possible value for our stock based on our assumptions is \$180.5. We get this value first by multiplying the stock price at period 0 by 95 % (decreasing the value of the stock by 5 %) to get the resulting value of \$190 of period 1. We then again multiply the stock price in period 1 by 95 % to get the resulting value of \$180.5. In period two, the value of a call option when a stock price is \$180.5 is the stock price minus the exercise price, \$180.5 − \$200, or −\$19.5. A negative value has no value to an investor so the value of a call option would be \$0. In period two, the value of a put option when a stock price is \$18.05 is the exercise price minus the stock price, \$200 − \$180.5, or \$19.5. We can derive the call and put option values for the other possible values of the stock in period 2 in the same fashion.

Figures 37.9 and 37.10 show the possible call and put option values for period 2.

We cannot calculate the value of the call and put options in period 1 the same way we did in period 2 because it's not the ending value of the stock. In period 1, there are two possible call values. One value is when the stock price increased, and one value is when the stock price decreased. The call option decision tree shown in Fig. 37.9 shows two possible values for a call option in period 1. If we just focus on the value of a call option when the stock price increases from period one, we will notice that it is like the decision tree for a call option for one period. This is shown in Fig. 37.11.

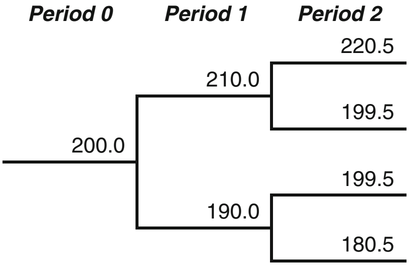


Fig. 37.8 AMZN stock price

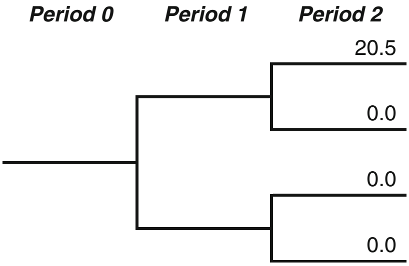


Fig. 37.9 AMZN call option

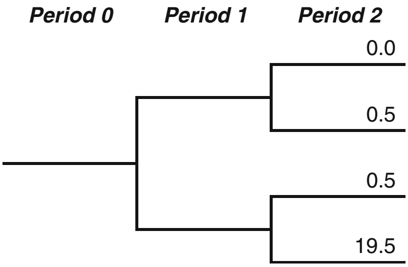


Fig. 37.10 AMZN put option

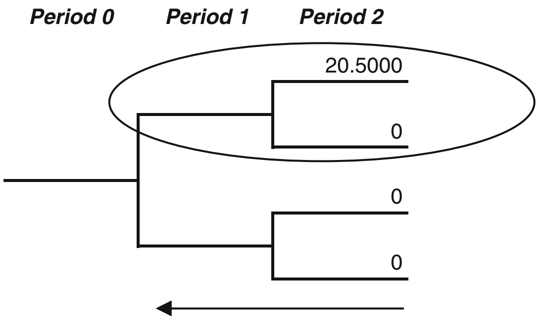


Fig. 37.11 AMZN call option

Using the same method for pricing a call option for one period, the price of a call option when stock price increases from period 0 will be \$15.922. The resulting decision tree is shown in Fig. 37.12.

In the same fashion, we can price the value of a call option when a stock price decreases. The price of a call option when a stock price decreases from period 0 is \$0. The resulting decision tree is shown in Fig. 37.13.

In the same fashion, we can price the value of a call option in period 0. The resulting decision tree is shown in Fig. 37.14.

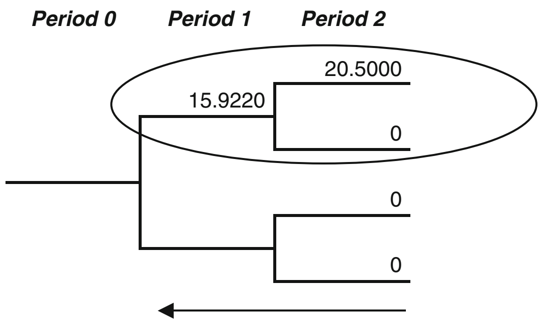


Fig. 37.12 AMZN call option

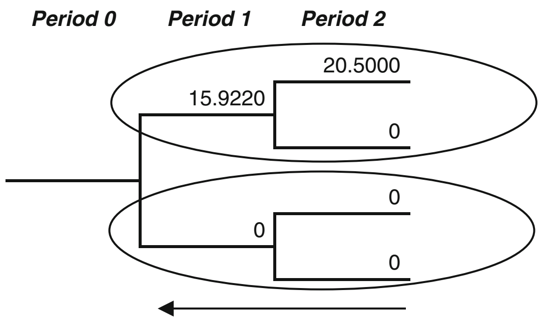


Fig. 37.13 AMZN call option

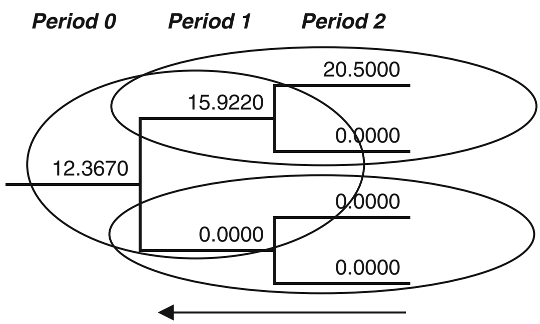


Fig. 37.14 AMZN call option

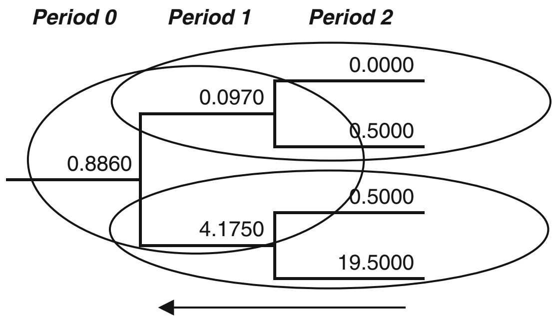


Fig. 37.15 AMZN put option

Table 37.1

Periods	Calculations
1	3
2	7
3	17
4	31
5	63
6	127
7	255
8	511
9	1,023
10	2,047
11	4,065
12	8,191

We can calculate the value of a put option in the same manner as we did in calculating the value of a call option. The decision tree for a put option is shown in Fig. 37.15.

37.5 Using Microsoft Excel to Create the Binomial Option Trees

In the previous section, we priced the value of a call and put option by pricing backwards, from the last period to the first period. This method of pricing call and put options will work for any n-period. To price the value of a call option for two periods required seven sets of calculations. The number of calculations increases dramatically as n increases. Table 37.1 lists the number of calculations for specific number of periods.

After two periods, it becomes very cumbersome to calculate and create the decision trees for a call and put option. In the previous section, we saw that calculations were very repetitive and mechanical. To solve this problem, this chapter will use Microsoft Excel to do the calculations and create the decision trees for the call and put options. We will also use Microsoft Excel to calculate and draw the related decision trees for the underlying stock and bond.

To solve this repetitive and mechanical calculation of the binomial option pricing model, we will look at a Microsoft Excel file called *binomialBS_OPM.xls*. We will use this Microsoft Excel workbook to produce four decision trees for the GE stock that was discussed in the previous sections. The four decision trees are:

1. Stock price
2. Call option price
3. Put option price
4. Bond price

This section will demonstrate how to use the *binomialBS_OPM.xls* Excel file to create the four decision trees. Figure 37.16 shows the Excel file *binomialBS_OPM.xls* after the file is opened. Pushing the button shown in Fig. 37.16 will get the dialog box shown in Fig. 37.17.

The dialog box shown in Fig. 37.17 shows the parameters for the *binomial option pricing model*. These parameters are changeable. The dialog box in Fig. 37.17 shows the default values.

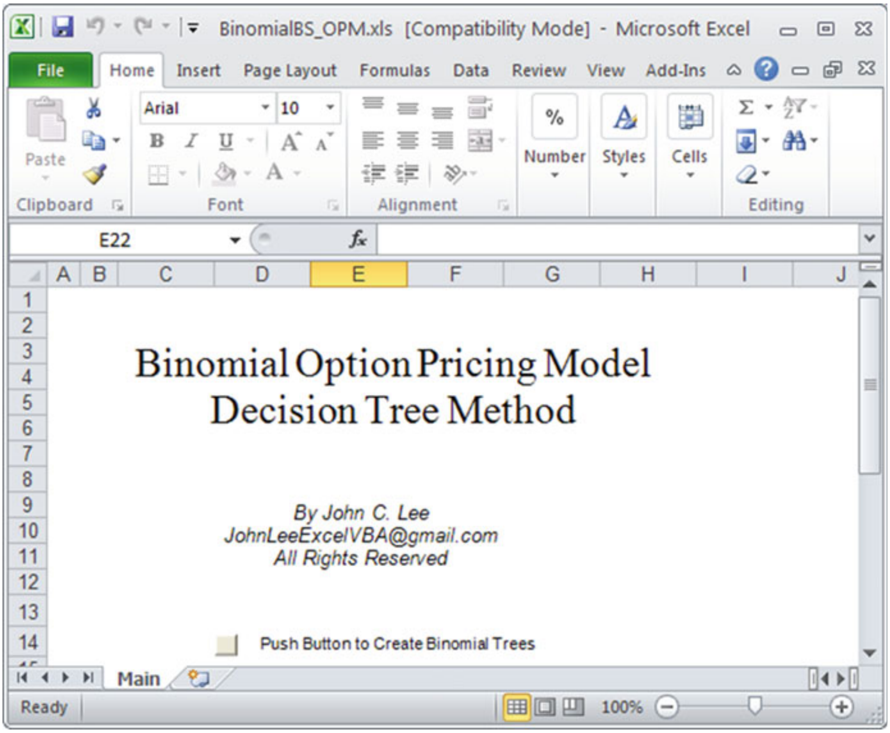


Fig. 37.16 Excel file BinomialBS_OPM.xls

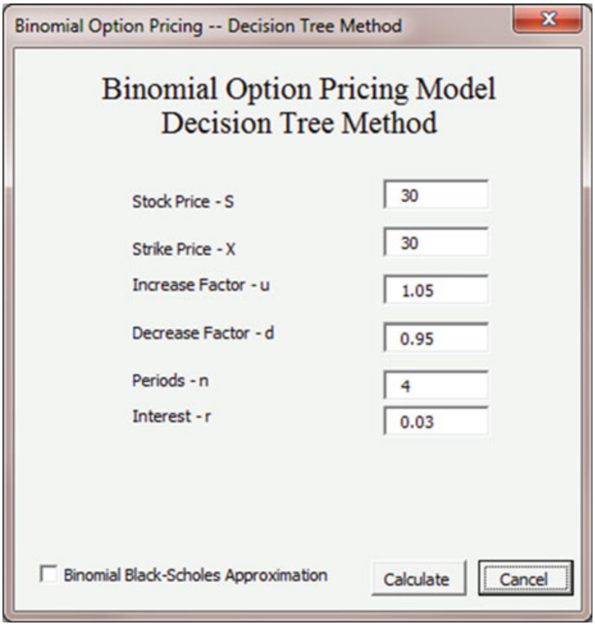


Fig. 37.17 Dialog box showing parameters for the binomial option pricing model

Pushing the *calculate* button shown in Fig. 37.17 will produce the four decision trees shown in Figs. 37.18, 37.19, 37.20, and 37.21.

The table at the beginning of this section indicated 31 calculations were required to create a decision tree that has four periods. This section showed four decision trees. Therefore, the Excel file did 31 * 4 = 121 calculations to create the four decision trees.

Benninga (2000, p. 260) has defined the price of a call option in a binomial option pricing model with n-periods as

$$C = \sum_{i=0}^{n} \binom{n}{i} q_u^i q_d^{n-i} \max\left[S(1+u)^i(1+d)^{n-i}, 0\right] \tag{37.6}$$

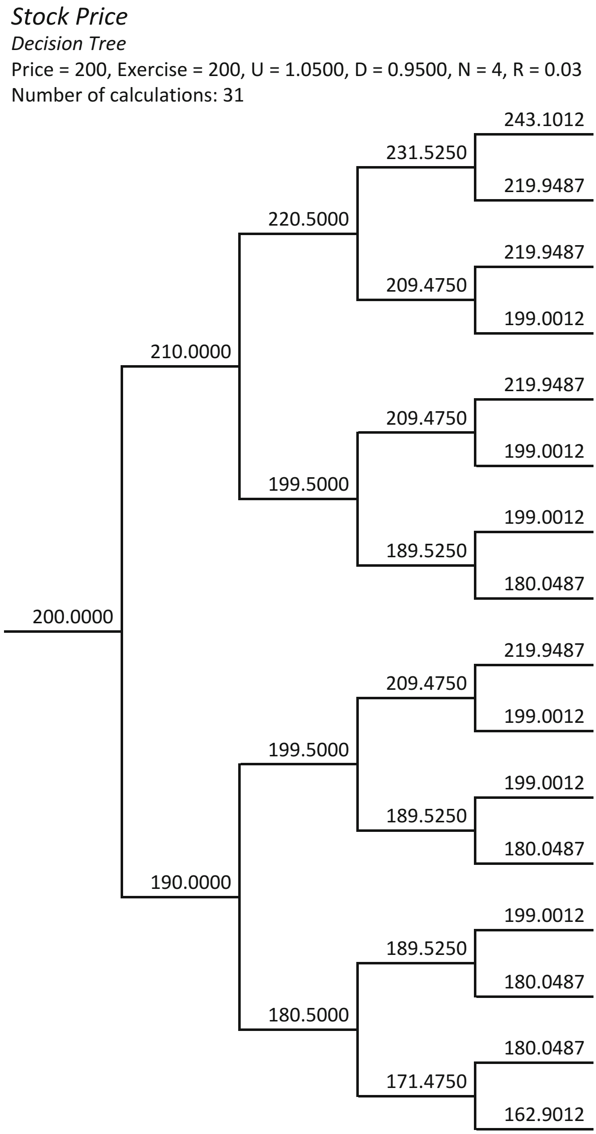
and the price of a put option in a binomial option pricing model with n-periods as

$$P = \sum_{i=0}^{n} \binom{n}{i} q_u^i q_d^{n-i} \max\left[X - S(1+u)^i(1+d)^{n-i}, 0\right] \tag{37.7}$$

Lee et al. (2000, p. 237) have defined the pricing of a call option in a binomial option pricing model with n-period as

$$C = \frac{1}{R^n} \sum_{k=0}^{n} \frac{n!}{k!(n-k!)} p^k (1-p)^{n-k} \max\left[0, (1+u)^k(1+d)^{n-k}, S - X\right] \tag{37.8}$$

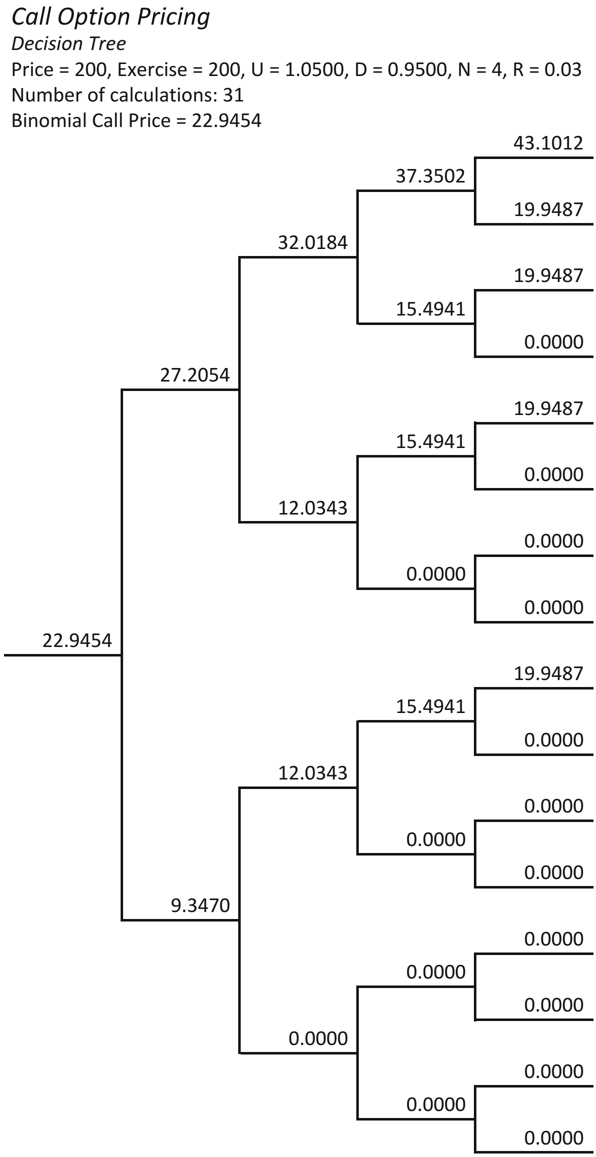
Fig. 37.18 Stock price decision tree



The definition of the pricing of a put option in a binomial option pricing model with n-period would then be defined as

$$P = \frac{1}{R^n} \sum_{k=0}^{n} \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} \max\left[0, X - (1+u)^k (1+d)^{n-k}, S\right] \quad (37.9)$$

Fig. 37.19 Call option pricing decision tree



37.6 Black–Scholes Option Pricing Model

The most famous option pricing model is the Black–Scholes option pricing model. In this section, we will demonstrate the usage of the Black–Scholes option pricing model. In latter sections, we will demonstrate the relationship between the binomial option pricing model and the Black–Scholes pricing model. The Black–Scholes

Fig. 37.20 Put option pricing decision tree

Put Option Pricing

Decision Tree

Price = 200, Exercise = 200, U = 1.0500, D = 0.9500, N = 4, R = 0.03

Number of calculations: 31

Binomial Put Price: 0.6428

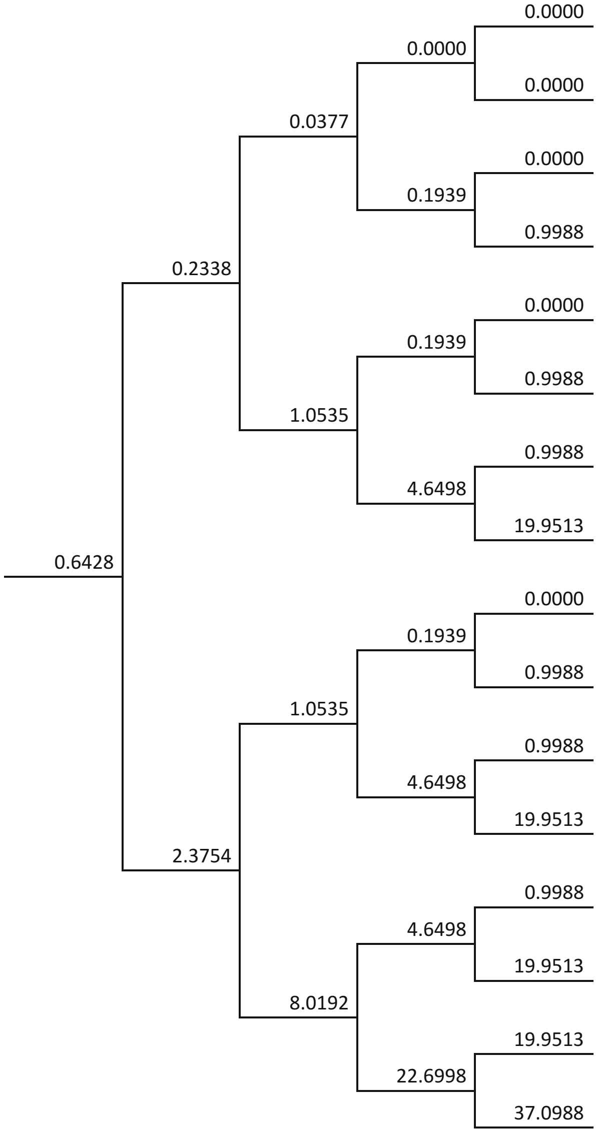
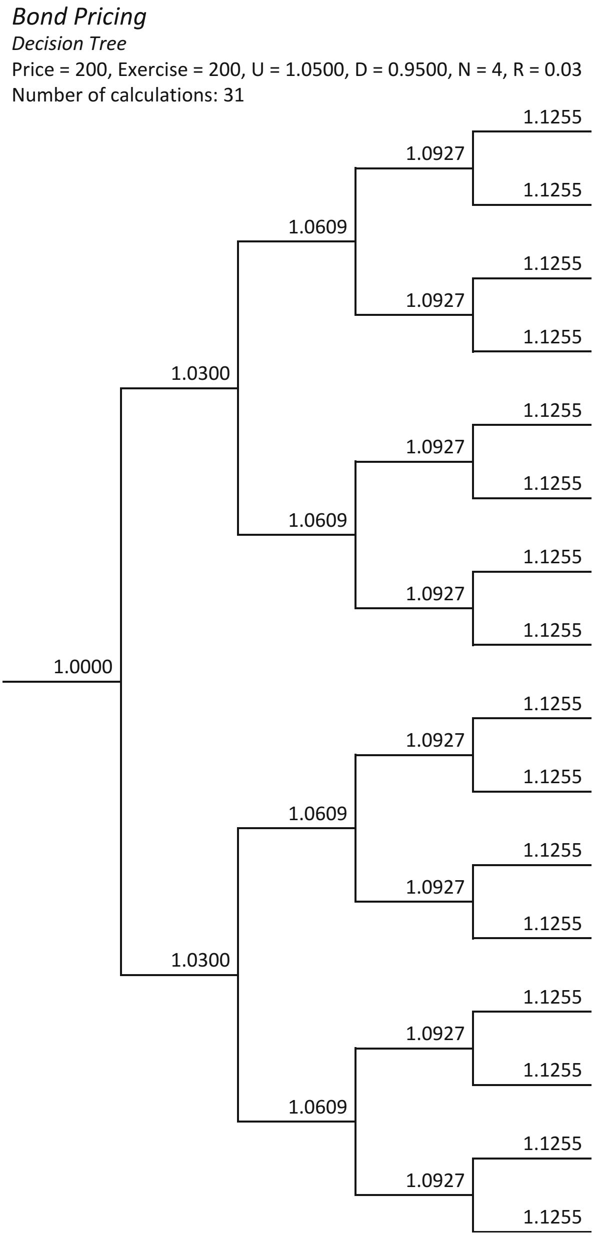


Fig. 37.21 Bond pricing decision tree



model prices European call and put options. The Black–Scholes model for a European call option is

$$C = SN(d1) - Xe^{-rT}N(d2) \tag{37.10}$$

where
C = call price
S = stock price
r = risk-free interest rate
T = time to maturity of option in years
N = standard normal distribution
σ = stock volatility

$$d1 = \frac{\ln(S/X) + \left(r + \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}}$$
$$d2 = d1 - \sigma\sqrt{T}$$

Let's manually calculate the price of a European call option in terms of Eq. 37.10 with the following parameter values, S = 200, X = 200, r = 3 %, T = 4, and σ = 20 %.

Solution

$$d1 = \frac{\ln(S/X) + \left(r + \frac{\sigma^2}{2}\right)T}{\sigma\sqrt{T}} = \frac{\ln(200/200) + \left(.03 + \frac{.2^2}{2}\right)(4)}{.2\sqrt{4}} = \frac{(.03 + .02) * 4}{.4} = \frac{.2}{.4} = .5,$$
$$d2 = .5 - .2\sqrt{4} = .1$$

$$\begin{aligned} N(d1) &= 0.69146, N(d2) = 0.5398, e^{-rT} = 0.8869 \\ C &= (200) * (0.69146) - (200) * (0.8869) * 0.5398 \\ &= 138.292 - 95.74972 = 42.5422 \end{aligned}$$

The Black–Scholes put-call parity equation is

$$P = C - S + Xe^{-rT}$$

The put option value for the stock would be

$$\begin{aligned} P &= 42.54 - 200 + 200(0.8869) \\ &= 42.54 - 200 + 177.38 = 19.92 \end{aligned}$$

37.7 Relationship Between the Binomial OPM and the Black–Scholes OPM

We can use either the binomial model or Black–Scholes to price an option. They both should result in similar numbers. If we look at the parameters in both models, we will notice that the binomial model has an *Increase Factor* (*U*), a *Decrease Factor* (*D*), and *n*-period parameters that the Black–Scholes model does not have. We also notice that the Black–Scholes model has the σ and T parameters that the binomial model does not have. Benninga (2008) suggests the following translation between the binomial and Black–Scholes parameters:

$$\Delta t = \mathrm{T}/n \quad \mathrm{R} = e^{r\Delta t} \quad \mathrm{U} = e^{\sigma\sqrt{\Delta t}}\ \mathrm{D} = e^{-\sigma\sqrt{\Delta t}}$$

In the Excel program, shown in Appendix 1, we use Benninga's (2008) *Increase Factor* and *Decrease Factor* definitions. They are defined as follows:

$$q_U = \frac{\mathrm{R} - \mathrm{D}}{\mathrm{R}(\mathrm{U} - \mathrm{D})},\quad q_D = \frac{\mathrm{U} - \mathrm{R}}{\mathrm{R}(\mathrm{U} - \mathrm{D})}$$

where
$\mathrm{U} = 1$ + percentage of price increase
$\mathrm{D} = 1 -$ percentage of price increase
$\mathrm{R} = 1$ + interest rate

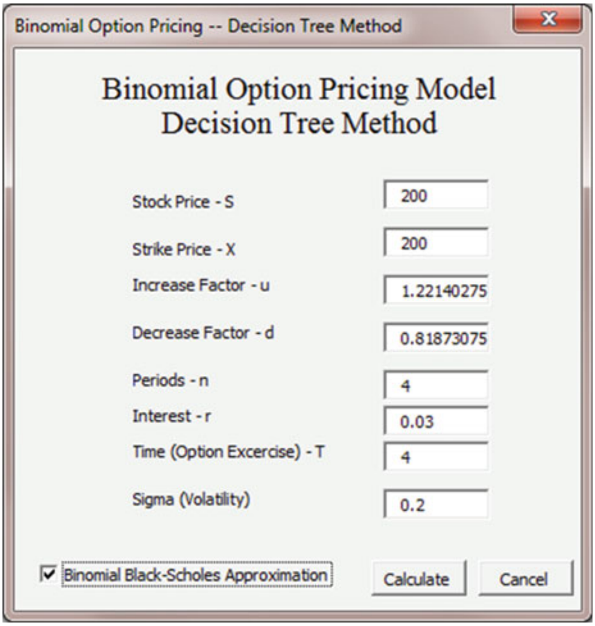


Fig. 37.22 Dialog box showing parameters for the binomial option pricing model

Call Option Pricing

Decision Tree

Price = 200, Exercise = 200, U = 1.2214, D = 0.8187, N = 4, R = 0.03

Number of calculations: 31

Binomial Call Price = 40.6705

Black-Scholes Call Price = 42.5356, d1 = 0.5000, d2 = 0.1000, N(d1) = 0.6915, N(d2) = 0.5398

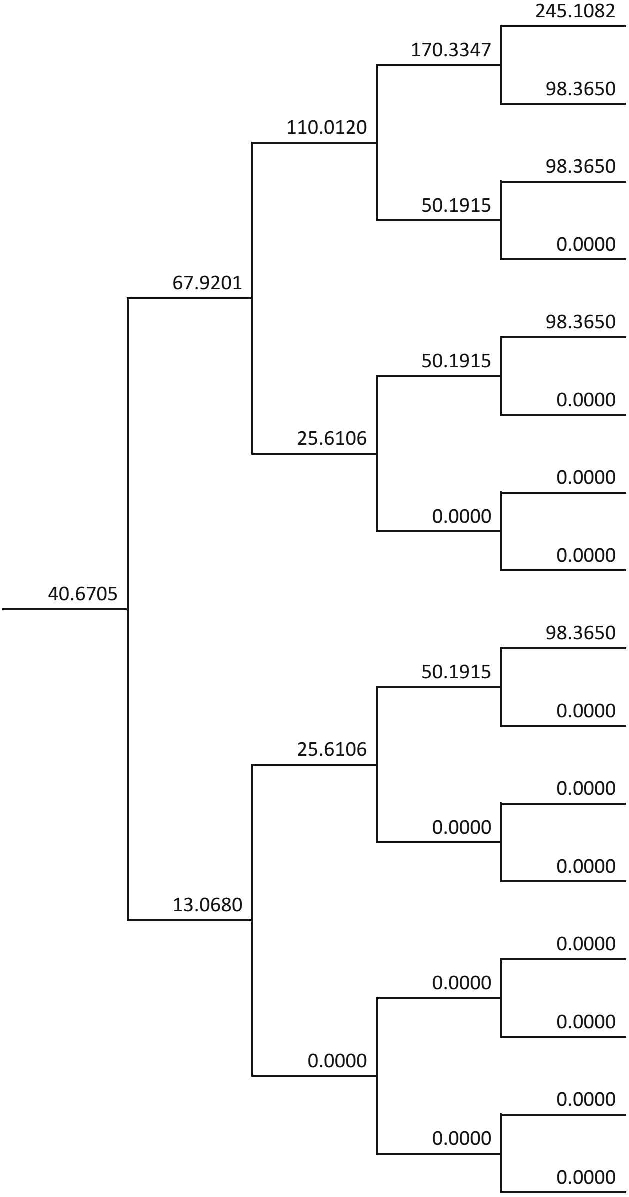


Fig. 37.23 Decision tree approximation of Black–Scholes call pricing

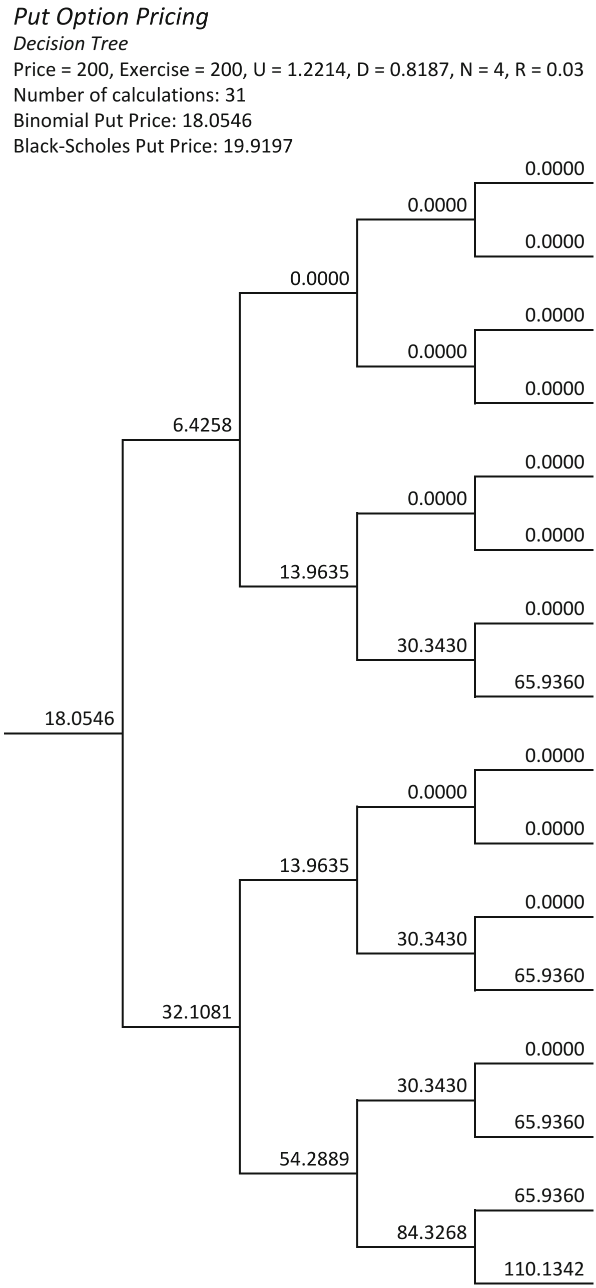


Fig. 37.24 Decision tree approximation of Black–Scholes put pricing

37.8 Decision Tree Black–Scholes Calculation

We will now use the *BinomialBS_OPM.xls* Excel file to calculate the binomial and Black–Scholes call and put values illustrated in Sect. 37.5. Notice that in Fig. 37.22 the *Binomial Black–Scholes Approximation* check box is checked. Checking this box will cause T and Sigma parameters to appear and will adjust the *Increase Factor – u and Decrease Factor – d* parameters. The adjustment was done as indicated in Sect. 37.7.

Notice in Figs. 37.23 and 37.24 that the binomial option pricing model value does not agree with the Black–Scholes option pricing model. The binomial OPM value will get very close to the Black–Scholes OPM value once the binomial parameter *n* gets very large. Benninga (2008) demonstrated that the binomial value will be close to the Black–Scholes when the binomial *n* parameter gets larger than 500.

37.9 Summary

This chapter demonstrated, with the aid of Microsoft Excel and decision trees, the binomial option model in a less mathematical fashion. This chapter allowed the reader to focus more on the concepts by studying the associated decision trees, which were created by Microsoft Excel. This chapter also demonstrates that using Microsoft Excel releases the reader from the computation burden of the binomial option model.

This chapter also published the Microsoft Excel Visual Basic for Application (VBA) code that created the binomial option decision trees. This allows for those who are interested in studying the many advance Microsoft Excel VBA programming concepts that were used to create the decision trees. One major computer science programming concept used by the Excel VBA program in this chapter is recursive programming. Recursive programming is the ideal of a procedure calling itself many times. Inside the procedure, there are statements to decide when not to call itself.

This chapter also used decision trees to demonstrate the relationship between the binomial option pricing model and the Black–Scholes option pricing model.

Appendix 1: Excel VBA Code: Binomial Option Pricing Model

It is important to note that the thing that makes Microsoft Excel powerful is that it offers a powerful professional programming language called Visual Basic for Applications (VBA). This section shows the VBA code that generated the decision trees for the binomial option pricing model. This code is in the form *frmBinomiaOption*. The procedure *cmdCalculate_Click* is the first procedure to run.

```
'/************************************************
'/ Relationship Between the Binomial OPM
'/ and Black-Scholes OPM:
'/ Decision Tree and Microsoft Excel Approach
'/
'/ by John Lee
'/ JohnLeeExcelVBA@gmail.com
'/ All Rights Reserved
'/************************************************
Option Explicit
Dim mwbTreeWorkbook As Workbook
Dim mwsTreeWorksheet As Worksheet
Dim mwsCallTree As Worksheet
Dim mwsPutTree As Worksheet
Dim mwsBondTree As Worksheet
Dim mdblPFactor As Double
Dim mBinomialCalc As Long
Dim mCallPrice As Double 'jcl 12/8/2008
Dim mPutPrice As Double 'jcl 12/8/2008
'/************************************************
'/Purpose: Keep track the numbers of binomial calc
'/************************************************
Property Let BinomialCalc(l As Long)
mBinomialCalc = l
End Property
Property Get BinomialCalc() As Long
BinomialCalc = mBinomialCalc
End Property
Property Set TreeWorkbook(wb As Workbook)
Set mwbTreeWorkbook = wb
End Property
Property Get TreeWorkbook() As Workbook
Set TreeWorkbook = mwbTreeWorkbook
End Property
Property Set TreeWorksheet(ws As Worksheet)
Set mwsTreeWorksheet = ws
End Property
Property Get TreeWorksheet() As Worksheet
Set TreeWorksheet = mwsTreeWorksheet
End Property
```

```
Property Set CallTree(ws As Worksheet)
Set mwsCallTree = ws
End Property
Property Get CallTree() As Worksheet
Set CallTree = mwsCallTree
End Property
Property Set PutTree(ws As Worksheet)
Set mwsPutTree = ws
End Property
Property Get PutTree() As Worksheet
Set PutTree = mwsPutTree
End Property
Property Set BondTree(ws As Worksheet)
Set mwsBondTree = ws
End Property
Property Get BondTree() As Worksheet
Set BondTree = mwsBondTree
End Property
Property Let CallPrice(dCallPrice As Double)
'12/8/2008
mCallPrice = dCallPrice
End Property
Property Get CallPrice() As Double
Let CallPrice = mCallPrice
End Property
Property Let PutPrice(dPutPrice As Double)
'12/10/2008
mPutPrice = dPutPrice
End Property
Property Get PutPrice() As Double
'12/10/2008
Let PutPrice = mPutPrice
End Property
Property Let PFactor(r As Double)
Dim dRate As Double
dRate = ((1 + r) - Me.txtBinomialD) / (Me.txtBinomialU -
Me.txtBinomialD)
Let mdblPFactor = dRate
End Property
Property Get PFactor() As Double
Let PFactor = mdblPFactor
End Property
Property Get qU() As Double
Dim dblDeltaT As Double
```

```
Dim dblDown As Double
Dim dblUp As Double
Dim dblR As Double
dblDeltaT = Me.txtTimeT / Me.txtBinomialN
dblR = Exp(Me.txtBinomialr * dblDeltaT)
dblUp = Exp(Me.txtSigma * VBA.Sqr(dblDeltaT))
dblDown = Exp(-Me.txtSigma * VBA.Sqr(dblDeltaT))
qU = (dblR - dblDown) / (dblR * (dblUp - dblDown))
End Property
Property Get qD() As Double
Dim dblDeltaT As Double
Dim dblDown As Double
Dim dblUp As Double
Dim dblR As Double
dblDeltaT = Me.txtTimeT / Me.txtBinomialN
dblR = Exp(Me.txtBinomialr * dblDeltaT)
dblUp = Exp(Me.txtSigma * VBA.Sqr(dblDeltaT))
dblDown = Exp(-Me.txtSigma * VBA.Sqr(dblDeltaT))
qD = (dblUp - dblR) / (dblR * (dblUp - dblDown))
End Property
Private Sub chkBinomialBSApproximation_Click()
On Error Resume Next
'Time and Sigma only BlackScholes parameter
Me.txtTimeT.Visible = Me.chkBinomialBSApproximation
Me.lblTimeT.Visible = Me.chkBinomialBSApproximation
Me.txtSigma.Visible = Me.chkBinomialBSApproximation
Me.lblSigma.Visible = Me.chkBinomialBSApproximation
txtTimeT_Change
End Sub
Private Sub cmdCalculate_Click()
Me.Hide
BinomialOption
Unload Me
End Sub
Private Sub cmdCancel_Click()
Unload Me
End Sub
Private Sub txtBinomialN_Change()
'jcl 12/8/2008
On Error Resume Next
If Me.chkBinomialBSApproximation Then
Me.txtBinomialU = Exp(Me.txtSigma * Sqr(Me.txtTimeT /
Me.txtBinomialN))
Me.txtBinomialD = Exp(-Me.txtSigma * Sqr(Me.txtTimeT /
Me.txtBinomialN))
```

```
End If
End Sub
Private Sub txtTimeT_Change()
'jcl 12/8/2008
On Error Resume Next
If Me.chkBinomialBSApproximation Then
Me.txtBinomialU = Exp(Me.txtSigma * Sqr(Me.txtTimeT /
Me.txtBinomialN))
Me.txtBinomialD = Exp(-Me.txtSigma * Sqr(Me.txtTimeT /
Me.txtBinomialN))
End If
End Sub
Private Sub UserForm_Initialize()
With Me
.txtBinomialS = 20
.txtBinomialX = 20
.txtBinomialD = 0.95
.txtBinomialU = 1.05
.txtBinomialN = 4
.txtBinomialr = 0.03
.txtSigma = 0.2
.txtTimeT = 4
Me.chkBinomialBSApproximation = False
End With
chkBinomialBSApproximation_Click
Me.Hide
End Sub
Sub BinomialOption()
Dim wbTree As Workbook
Dim wsTree As Worksheet
Dim rColumn As Range
Dim ws As Worksheet
Set Me.TreeWorkbook = Workbooks.Add
Set Me.BondTree = Me.TreeWorkbook.Worksheets.Add
Set Me.PutTree = Me.TreeWorkbook.Worksheets.Add
Set Me.CallTree = Me.TreeWorkbook.Worksheets.Add
Set Me.TreeWorksheet = Me.TreeWorkbook.Worksheets.Add
Set rColumn = Me.TreeWorksheet.Range("a1")
With Me
.BinomialCalc = 0
.PFactor = Me.txtBinomialr
.CallTree.Name = "Call Option Price"
.PutTree.Name = "Put Option Price"
.TreeWorksheet.Name = "Stock Price"
.BondTree.Name = "Bond"
```

```
  End With
  DecisionTree rCell:=rColumn, nPeriod:=Me.txtBinomialN
+ 1, _
  dblPrice:=Me.txtBinomialS, sngU:=Me.txtBinomialU, _
  sngD:=Me.txtBinomialD
  DecitionTreeFormat
  TreeTitle  wsTree:=Me.TreeWorksheet,  sTitle:="Stock
Price "
  TreeTitle  wsTree:=Me.CallTree,  sTitle:="Call  Option
Pricing"
  TreeTitle  wsTree:=Me.PutTree,  sTitle:="Put  Option
Pricing"
  TreeTitle wsTree:=Me.BondTree, sTitle:="Bond Pricing"
  Application.DisplayAlerts = False
  For Each ws In Me.TreeWorkbook.Worksheets
  If Left(ws.Name, 5) = "Sheet" Then
  ws.Delete
  Else
  ws.Activate
  ActiveWindow.DisplayGridlines = False
  ws.UsedRange.NumberFormat = "#,##0.0000_);(#,##0.0000)"
  End If
  Next
  Application.DisplayAlerts = True
  Me.TreeWorksheet.Activate
  End Sub
  Sub TreeTitle(wsTree As Worksheet, sTitle As String)
  wsTree.Range("A1:A5").EntireRow.Insert (xlShiftDown)
  With wsTree
  With.Cells(1)
  .Value = sTitle
  .Font.Size = 20
  .Font.Italic = True
  End With
  With.Cells(2, 1)
  .Value = "Decision Tree"
  .Font.Size = 16
  .Font.Italic = True
  End With
  With.Cells(3, 1)
  .Value = "Price = " & Me.txtBinomialS & _
  ",Exercise = " & Me.txtBinomialX & _
  ",U = " & Format(Me.txtBinomialU, "#,##0.0000") & _
  ",D = " & Format(Me.txtBinomialD, "#,##0.0000") & _
```

```
  ",N = " & Me.txtBinomialN & _
  ",R = " & Me.txtBinomialr
  .Font.Size = 14
  End With
  With.Cells(4, 1)
  .Value = "Number of calculations: " & Me.BinomialCalc
  .Font.Size = 14
  End With
  If wsTree Is Me.CallTree Then
  With.Cells(5, 1)
  .Value = "Binomial Call Price= " & Format(Me.CallPrice,
"#,##0.0000")
  .Font.Size = 14
  End With
  If Me.chkBinomialBSApproximation Then
  wsTree.Range("A6:A7").EntireRow.Insert (xlShiftDown)
  With.Cells(6, 1)
  .Value = "Black-Scholes Call Price= " & Format(Me.
BS_Call, "#,##0.0000") _
  & ",d1=" & Format(Me.BS_D1, "#,##0.0000") _
  & ",d2=" & Format(Me.BS_D2, "#,##0.0000") _
  & ",N(d1)=" & Format(WorksheetFunction.NormSDist
(BS_D1), "#,##0.0000") _
  & ",N(d2)=" & Format(WorksheetFunction.NormSDist
(BS_D2), "#,##0.0000")
  .Font.Size = 14
  End With
  End If
  ElseIf wsTree Is Me.PutTree Then
  With.Cells(5, 1)
  .Value = "Binomial Put Price: " & Format(Me.PutPrice,
"#,##0.0000")
  .Font.Size = 14
  End With
  If Me.chkBinomialBSApproximation Then
  wsTree.Range("A6:A7").EntireRow.Insert (xlShiftDown)
  With.Cells(6, 1)
  .Value = "Black-Scholes Put Price: " & Format(Me.BS_PUT,
"#,##0.0000")
  .Font.Size = 14
  End With
  End If
  End If
  End With
```

```
End Sub
Sub BondDecisionTree(rPrice As Range, arCell As Variant,
iCount As Long)
Dim rBond As Range
Dim rPup As Range
Dim rPDown As Range
Set rBond = Me.BondTree.Cells(rPrice.Row, rPrice.
Column)
Set rPup = Me.BondTree.Cells(arCell(iCount - 1).Row,
arCell(iCount - 1).Column)
Set rPDown = Me.BondTree.Cells(arCell(iCount).Row,
arCell(iCount).Column)
If rPup.Column = Me.TreeWorksheet.UsedRange.Columns.
Count Then
rPup.Value = (1 + Me.txtBinomialr) ^ (rPup.Column - 1)
rPDown.Value = rPup.Value
End If
With rBond
.Value = (1 + Me.txtBinomialr) ^ (rBond.Column - 1)
.Borders(xlBottom).LineStyle = xlContinuous
End With
rPDown.Borders(xlBottom).LineStyle = xlContinuous
With rPup
.Borders(xlBottom).LineStyle = xlContinuous
.Offset(1, 0).Resize((rPDown.Row - rPup.Row), 1). _
Borders(xlEdgeLeft).LineStyle = xlContinuous
End With
End Sub
Sub PutDecisionTree(rPrice As Range, arCell As Variant,
iCount As Long)
Dim rCall As Range
Dim rPup As Range
Dim rPDown As Range
Set rCall = Me.PutTree.Cells(rPrice.Row, rPrice.Column)
Set rPup = Me.PutTree.Cells(arCell(iCount - 1).Row,
arCell(iCount - 1).Column)
Set rPDown = Me.PutTree.Cells(arCell(iCount).Row,
arCell(iCount).Column)
If rPup.Column = Me.TreeWorksheet.UsedRange.Columns.
Count Then
rPup.Value = WorksheetFunction.Max(Me.txtBinomialX -
arCell(iCount - 1), 0)
rPDown.Value = WorksheetFunction.Max(Me.txtBinomialX -
arCell(iCount), 0)
```

```
  End If
  With rCall
  '12/10/2008
  If Not Me.chkBinomialBSApproximation Then
  .Value = (Me.PFactor * rPup + (1 - Me.PFactor) * rPDown) /
(1 + Me.txtBinomialr)
  Else
  .Value = (Me.qU * rPup) + (Me.qD * rPDown)
  End If
  Me.PutPrice =.Value '12/8/2008
  .Borders(xlBottom).LineStyle = xlContinuous
  End With
  rPDown.Borders(xlBottom).LineStyle = xlContinuous
  With rPup
  .Borders(xlBottom).LineStyle = xlContinuous
  .Offset(1, 0).Resize((rPDown.Row - rPup.Row), 1). _
  Borders(xlEdgeLeft).LineStyle = xlContinuous
  End With
  End Sub
  Sub CallDecisionTree(rPrice As Range, arCell As Variant,
iCount As Long)
  Dim rCall As Range
  Dim rCup As Range
  Dim rCDown As Range
  Set rCall = Me.CallTree.Cells(rPrice.Row, rPrice.
Column)
  Set rCup = Me.CallTree.Cells(arCell(iCount - 1).Row,
arCell(iCount - 1).Column)
  Set rCDown = Me.CallTree.Cells(arCell(iCount).Row,
arCell(iCount).Column)
  If rCup.Column = Me.TreeWorksheet.UsedRange.Columns.
Count Then
  With rCup
  .Value = WorksheetFunction.Max(arCell(iCount - 1) - Me.
txtBinomialX, 0)
  .Borders(xlBottom).LineStyle = xlContinuous
  End With
  With rCDown
  .Value = WorksheetFunction.Max(arCell(iCount) - Me.
txtBinomialX, 0)
  .Borders(xlBottom).LineStyle = xlContinuous
  End With
  End If
  With rCall
```

```
If Not Me.chkBinomialBSApproximation Then
.Value = (Me.PFactor * rCup + (1 - Me.PFactor) * rCDown) /
(1 + Me.txtBinomialr)
Else
.Value = (Me.qU * rCup) + (Me.qD * rCDown)
End If
Me.CallPrice =.Value '12/8/2008
.Borders(xlBottom).LineStyle = xlContinuous
End With
rCup.Offset(1, 0).Resize((rCDown.Row - rCup.Row), 1). _
Borders(xlEdgeLeft).LineStyle = xlContinuous
End Sub
Sub DecitionTreeFormat()
Dim rTree As Range
Dim nColumns As Integer
Dim rLast As Range
Dim rCell As Range
Dim lCount As Long
Dim lCellSize As Long
Dim vntColumn As Variant
Dim iCount As Long
Dim lTimes As Long
Dim arCell() As Range
Dim sFormatColumn As String
Dim rPrice As Range
Application.StatusBar = "Formatting Tree.. "
Set rTree = Me.TreeWorksheet.UsedRange
nColumns = rTree.Columns.Count
Set rLast = rTree.Columns(nColumns).EntireColumn.
SpecialCells(xlCellTypeConstants, 23)
lCellSize = rLast.Cells.Count
For lCount = nColumns To 2 Step -1
sFormatColumn = rLast.Parent.Columns(lCount).
EntireColumn.Address
Application.StatusBar = "Formatting column " &
sFormatColumn
ReDim vntColumn(1 To (rLast.Cells.Count / 2), 1)
Application.StatusBar = "Assigning values to array for
column " & _
rLast.Parent.Columns(lCount).EntireColumn.Address
vntColumn = rLast.Offset(0, -1).EntireColumn.Cells(1).
Resize(rLast.Cells.Count / 2, 1)
rLast.Offset(0, -1).EntireColumn.ClearContents
ReDim arCell(1 To rLast.Cells.Count)
lTimes = 1
```

```
Application.StatusBar = "Assigning cells to arrays.
Total number of cells: " & lCellSize
For Each rCell In rLast.Cells
Application.StatusBar = "Array to column " &
sFormatColumn & " Cells " & rCell.Row
Set arCell(lTimes) = rCell
lTimes = lTimes + 1
Next
lTimes = 1
Application.StatusBar = "Formatting leaves for column "
& sFormatColumn
For iCount = 2 To lCellSize Step 2
Application.StatusBar = "Formatting leaves for cell "
& arCell(iCount).Address
If rLast.Cells.Count <> 2 Then
Set rPrice = arCell(iCount).Offset(-1 * ((arCell
(iCount).Row - arCell(iCount -1).Row) / 2), -1)
rPrice.Value = vntColumn(lTimes, 1)
Else
Set rPrice = arCell(iCount).Offset(-1 * ((arCell
(iCount).Row - arCell(iCount -1).Row) / 2), -1)
rPrice.Value = vntColumn
End If
arCell(iCount).Borders(xlBottom).
LineStyle = xlContinuous
With arCell(iCount - 1)
.Borders(xlBottom).LineStyle = xlContinuous
.Offset(1, 0).Resize((arCell(iCount).Row - arCell
(iCount - 1).Row), 1). _
Borders(xlEdgeLeft).LineStyle = xlContinuous
End With
lTimes = 1 + lTimes
CallDecisionTree rPrice:=rPrice, arCell:=arCell,
iCount:=iCount
PutDecisionTree rPrice:=rPrice, arCell:=arCell,
iCount:=iCount
BondDecisionTree rPrice:=rPrice, arCell:=arCell,
iCount:=iCount
Next
Set rLast = rTree.Columns(lCount - 1).EntireColumn.
SpecialCells(xlCellTypeConstants, 23)
lCellSize = rLast.Cells.Count
Next ' / outer next
rLast.Borders(xlBottom).LineStyle = xlContinuous
Application.StatusBar = False
```

```
End Sub
'/**************************************************
'/Purpse: To calculate the price value of every state
of the binomial
'/ decision tree
'/**************************************************
Sub DecisionTree(rCell As Range, nPeriod As Integer, _
dblPrice As Double, sngU As Single, sngD As Single)
Dim lIteminColumn As Long
If Not nPeriod = 1 Then
'Do Up
DecisionTree rCell:=rCell.Offset(0, 1), nPeriod:=
nPeriod - 1, _
dblPrice:=dblPrice * sngU, sngU:=sngU, _
sngD:=sngD
'Do Down
DecisionTree rCell:=rCell.Offset(0, 1), nPeriod:=
nPeriod - 1, _
dblPrice:=dblPrice * sngD, sngU:=sngU, _
sngD:=sngD
End If
lIteminColumn = WorksheetFunction.CountA(rCell.
EntireColumn)
If lIteminColumn = 0 Then
rCell = dblPrice
Else
If nPeriod <> 1 Then
rCell.EntireColumn.Cells(lIteminColumn + 1) = dblPrice
Else
rCell.EntireColumn.Cells(((lIteminColumn + 1) * 2) -
1) = dblPrice
Application.StatusBar = "The number of binomial calcs are
: " & Me.BinomialCalc _ & " at cell " & rCell.EntireColumn.
Cells(((lIteminColumn + 1) * 2) - 1).Address
End If
End If
Me.BinomialCalc = Me.BinomialCalc + 1
End Sub
Function BS_D1() As Double
Dim dblNumerator As Double
Dim dblDenominator As Double
On Error Resume Next
dblNumerator = VBA.Log(Me.txtBinomialS / Me.
txtBinomialX) + _
```

```
((Me.txtBinomialr + Me.txtSigma ^ 2 / 2) * Me.txtTimeT)
dblDenominator = Me.txtSigma * Sqr(Me.txtTimeT)
BS_D1 = dblNumerator / dblDenominator
End Function
Function BS_D2() As Double
On Error Resume Next
BS_D2 = BS_D1 - (Me.txtSigma * VBA.Sqr(Me.txtTimeT))
End Function
Function BS_Call() As Double
BS_Call = (Me.txtBinomialS * WorksheetFunction.
NormSDist(BS_D1)) _
- Me.txtBinomialX * Exp(-Me.txtBinomialr * Me.txtTimeT) * _
WorksheetFunction.NormSDist(BS_D2)
End Function
'Used put-call parity theorem to price put option
Function BS_PUT() As Double
BS_PUT = BS_Call - Me.txtBinomialS + _
(Me.txtBinomialX * Exp(-Me.txtBinomialr * Me.txtTimeT))
End Function
```

References

Benninga, S. (2000). *Financial modeling*. Cambridge: MIT Press.

Benninga, S. (2008). *Financial modeling*. Cambridge: MIT Press.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy, 31*, 637–659.

Cox, J., Ross, S. A., & Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of Financial Economics, 7*, 229–263.

Daigler, R. T. (1994). *Financial futures and options markets concepts and strategies*. New York: Harper Collins.

Jarrow, R., & TurnBull, S. (1996). *Derivative securities*. Cincinnati: South-Western College Publishing.

Lee, J. C. (2001). Using Microsoft excel and decision trees to demonstrate the binomial option pricing model. *Advances in Investment Analysis and Portfolio Management, 8*, 303–329.

Lee, C. F. (2009). *Handbook of quantitative finance*. New York: Springer.

Lee, C. F., & Lee, A. C. (2006). *Encyclopedia of finance*. New York: Springer.

Lee, C. F., Lee, J. C., & Lee, A. C. (2000). *Statistics for business and financial economics*. New Jersey: World Scientific.

Lee, J. C., Lee, C. F., Wang, R. S., & Lin, T. I. (2004). On the limit properties of binomial and multinomial option pricing models: Review and integration. In C. F. Lee (Ed.), *Advances in quantitative analysis of finance and accounting new series* (Vol. 1). Singapore: World Scientific.

Rendleman, R. J., Jr., & Barter, B. J. (1979). Two-state option pricing. *Journal of Finance, 34*(5), 1093–1110.

Walkenbach, J. (2010). *Excel 2010 power programming with VBA*. Indianapolis: Wiley.

Wells, E., & Harshbarger, S. (1997). *Microsoft excel 97 developer's handbook*. Redmond: Microsoft Press.

Dividend Payments and Share Repurchases of US Firms: An Econometric Approach

38

Alok Bhargava

Contents

A. Bhargava
School of Public Policy, University of Maryland, College Park, MD, USA
e-mail: bhargava@umd.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_38,

Abstract

The analyses of dividends paid by firms and decisions to repurchase their own shares require an econometric approach because of the complex dynamic interrelationships. This chapter begins by, first, highlighting the importance of developing comprehensive econometric models for these interrelationships. It is common in finance research to spell out "specific hypotheses" and conduct empirical research to investigate validity of the hypotheses. However, such an approach can be misleading in situations where variables are simultaneously determined as is often the case in financial applications. Second, financial and accounting databases such as Compustat are complex and contain useful longitudinal information on variables that display considerable heterogeneity across the firms. Empirical analyses of financial databases demand the use of econometric and computational methods in order to draw robust inferences. For example, using longitudinal data on the same US firms, it was found that dividends were neither "disappearing" nor "reappearing" but were relatively stable in the period 1992–2007 Bhargava (*Journal of the Royal Statistical Society A, 173*, 631–656, 2010). Third, the econometric methodology tackled the dynamics of relationships and investigated endogeneity of certain explanatory variables. Identification of the model parameters is achieved in such models by exploiting the cross-equations restrictions on the coefficients in different time periods. Moreover, the estimation entails using nonlinear optimization methods to compute the maximum likelihood estimates of the dynamic random effects models and for testing statistical hypotheses using likelihood ratio tests. For example, share repurchases were treated as endogenous explanatory variable in the models for dividend payments, and dividends were treated as endogenous variables in the models for share repurchases. The empirical results showed that dividends are decided quarterly at the first stage, and higher dividends payments lowered share repurchases by firms that are made at longer intervals. These findings cast some doubt on evidence for the simple "substitution" hypothesis between dividends and share repurchases. The appendix outlines some of the econometric estimation techniques and tests that are useful for research in finance.

Keywords

Compustat database • Corporate policies • Dividends • Dynamic random effects models • Econometric methodology • Endogeneity • Maximum likelihood • Intangible assets • Model formulation • Nonlinear optimization • Panel data • Share repurchases

38.1 Introduction and Background

The proximate determinants of dividends payments by firms are of interest to researchers in economics and finance. From a historical perspective, Tinbergen (1939) proposed a model for predicting stock prices based on firms' earnings and dividends per share. Later, Lintner (1956) proposed a partial adjustment scheme for dividends and represented it by a first-order autoregressive ("dynamic") model where dividends per share in the current time period were a function of the previous levels. Such models were also employed in the economics literature to reflect "habit persistence" in the consumption function (Duesenberry 1949), i.e., current consumption depends on the past consumption. The work by Lintner (1956) was extended to data on US firms by Fama and Babiak (1968). However, the methods used in the early research for estimation of model parameters were rather elementary.

While data on macroeconomic variables such as consumption and income typically span for long periods and facilitate the estimation of econometric models (Koopmans 1950), the estimation of models using firm level data presents additional complications. For example, the large amount of heterogeneity across firms needs to be taken into account. Such issues are important for finance research because the neglect of unobserved between-firm differences can vitiate the consistency properties of the estimated model parameters. Moreover, financial databases such as (Compustat 2008) contain elaborate information at the firm level and several variables affecting dividend payments by firms need to be accounted for in the econometric models.

Further, shares repurchase, where firms purchase their own shares from stock holders and place them in "Treasury stock," have become popular since the 1980s in the United States and in the European Union (von Eije and Megginson 2008). Share repurchases reduce the numbers of shares held by investors ("common shares outstanding"), thereby increasing share values and hence are a form of payment to stock holders. It has also become popular for executives and employees of firms to receive part of their remuneration as options to purchase company stock at the current price in the future. The longitudinal information in financial data sets such as Compustat can facilitate the investigation of substantive issues such as the proximate determinants of dividend payments and the interrelationships between dividends and share repurchases by firms (e.g., Jagannathan et al. 2000; Fama and French 2001; Grullon and Michaely 2002; De Angelo et al. 2006; Skinner 2008; Bhargava 2010; Jones and Wu 2010).

Further, while analyses of the Compustat data can provide useful insights, methodological aspects are critical. For example, most previous analyses have not exploited the longitudinal ("panel") nature of the Compustat data. Instead, cross-sectional regressions have been estimated in different time periods for modeling firms' chances of paying dividends, i.e., dependent variables take the values 0 or 1 (Jagannathan et al. 2000; Fama and French 2001; De Angelo et al. 2006). In applications where longitudinal analyses were conducted for dividends payments (Skinner 2008), the model parameters were likely to be inconsistently estimated because the appropriate estimators for dynamic models were not employed. This chapter discusses several difficulties in the previous literature and presents findings on the interrelationships between dividends and share repurchases by over 2,000 US industrial firms (see Bhargava 2010).

First, it is essential to model the time structure of the interrelationships between dividends and share repurchases using longitudinal data on firms. While Lintner (1956) argued that firms may be reluctant to alter dividend rates, firms often adjust dividends per share to the market circumstances. Thus, classification of firms between dividend "payers" and "non-payers" is not a concrete one since a majority of firms may pay dividends during some periods. For example, using eight 2-yearly averages of dividends paid by a panel of 3,154 US industrial firms in the period 1992–2007 (i.e., for 1992–1993, 1994–1995, 1996–1997, 1998–1999, 2000–2001, 2002–2003, 2004–2005, 2006–2007), Bhargava (2010) found that 42 % of the firms never paid dividends, 33 % paid dividends in all 8 periods, and 25 % paid dividends in 1–7 periods. Moreover, dividends per share ranged from \$0.00 to \$67.61 so that it is appropriate to model the payments via continuous random variables (Fama and Babiak 1968). Similarly, firms' share repurchase are continuous variables and can be defined using the variables "Treasury stock," "purchase of common stock," and "sale of common stock." Using longitudinal data on 2,907 firms in 8 periods, 20 % of firms made no repurchases, 5 % repurchased shares in all 8 periods, and 75 % made 1–7 repurchases (Bhargava 2010). Thus, models with continuous dependent variables for dividends and share repurchases are suitable for analyses of payout decisions. Moreover, it is essential to use longitudinal data on the *same* firms for analyzing the interrelationships between dividends and share repurchases.

Second, from a methodological standpoint, the estimation of dynamic models from longitudinal data covering large number of firms observed for a few time periods requires the treatment of lagged dependent variables as "endogenous" variables that are correlated with firm-specific random effects (Anderson and Hsiao 1981; Bhargava and Sargan 1983); this issue will be discussed below and in the Appendix. Despite the popularity of Lintner (1956) type models in finance literature, such issues have seldom been addressed by researchers and critically affect consistency properties of the estimated model parameters. Moreover, it is informative to model the actual magnitudes of dividends and share repurchases, since these are likely to be related to firm characteristics such as size, earnings, and debt. It would also be useful to consider their joint determination; models for continuous dependent variables using longitudinal data can handle situations where some explanatory variables are correlated with errors affecting the models (Bhargava and Sargan 1983; Bhargava 1991).

Third, there has been a discussion in the finance literature regarding possible "substitution" between dividends payments and share repurchases by firms. While some researchers have argued that these payout methods may not be substitutes (e.g., John and Williams 1985), substitution of dividends by share repurchases (or vice versa) are consistent with certain economic formulations (Miller and Modigliani (1961). For example, Grullon and Michaely (2002) interpreted negative correlations between forecasting errors in changes in dividends and share repurchases as evidence of substitution. However, these issues demand an econometric approach for addressing the possible asymmetries and endogeneity in the relationships. For example, using the Compustat data on 2,800 industrial firms in the eight time periods noted above, bivariate correlations between dividends per share and repurchases were 0.11, 0.11, 0.08, 0.07, 0.11, 0.03, 0.05, and 0.06, respectively. These *positive* correlations were statistically significant; similar results were obtained for correlations between total dividend payments and share repurchases. Thus, investigation of possible substitution between dividends and share repurchases demands a comprehensive econometric framework (see below).

This chapter outlines a comprehensive analysis of the interrelationships between dividends and share repurchases using longitudinal Compustat data at 2-year intervals on over 2,000 US firms for the period 1992–2007 (Bhargava 2010). A brief discussion of issues arising in econometric methodology that is important for formulation of models and testing hypotheses in finance research is presented in Sect. 38.2. The analytical framework for dividends and share repurchase interrelationships is developed in Sect. 38.3, and econometric models embodying various hypotheses are outlined. In Sect. 38.4, Compustat variables are described and construction of the longitudinal data set is explained. For example, 2-yearly averages for variables are well suited for the analyses because share repurchases occur roughly at 2-year intervals (e.g., Stephens and Weisbach 1998); averaging also reduces the impact of missing values. The comprehensive econometric models for dividends and share repurchase relationships are outlined in Sect. 38.5. The econometric framework is outlined in Sect. 38.6 and it yields consistent and efficient estimates of model parameters; diagnostic tests for model adequacy and certain steps in the applications of the methods are outlined in the Appendix. The descriptive results from the Compustat data are in Sect. 38.7; results from estimating simple dynamic models for dividends and share repurchases are in Sect. 38.8. In Sect. 38.9, the results from estimating comprehensive dynamic econometric models for dividends and share repurchases are discussed. It is emphasized that one can draw robust conclusions from the estimated parameters of models that capture salient aspects of the various relationships. The conclusions are summarized in Sect. 38.10.

38.2 Specific Hypotheses, Financial Databases, and the Formulation of Econometric Models

It would be useful to discuss the role of conceptual aspects in the formulation of econometric models for finance relationships especially in the context of interrelationships between dividends and share repurchases. First, note that in biomedical sciences, randomized controlled trials are designed to investigate "specific hypotheses." The design of experiments is influenced by previous findings that are treated as assumptions in the trial (Cox 1958) and enable the investigation of specific hypotheses under consideration (Bhargava 2008). The validity of specific hypotheses is investigated using the emerging data from control and treatment groups of the trial. In contrast, finance researchers often spell out hypotheses that are influenced by economic postulates and investigate the support for the hypotheses using existing databases such as Compustat for the United States. However, the hypotheses driving the economic and finance investigation are often based on untested assumptions and/or "stylized facts" (Granger 1992). By contrast, assumptions invoked in biomedical sciences reflect the knowledge accumulated from previous experimental studies, and hypotheses are tested using the new data.

Further, in social sciences research, most phenomena being explained result from the interactions between economic and social factors. Because different dimensions of the relationships merit different emphasis depending on the context, social science researchers are forced to address several hypotheses simultaneously. Typically, this is done via formulation of comprehensive models where, for example, the systematic part of the relationships incorporates the relevant variables. The stochastic properties of the dependent variable and error terms are also tackled in the formulation of comprehensive econometric models.

The analyses of databases such as Compustat entail the modeling of accounting, finance, and economic variables so that financial research often demands econometric modeling of variables that are jointly determined. Moreover, economic theory is helpful in identifying variables and often implies restrictions on certain model parameters. Thus, investigation of hypotheses in finance applications entails the development of comprehensive econometric models within which several hypotheses can be embedded and tested using the estimated model parameters. This is in contrast with the approach in biomedical sciences where trials are designed to investigate specific hypotheses, and researchers are not concerned with interdependence between the variables. Thus, the currently popular approach in finance of spelling out a few specific hypotheses needs to be augmented by development of comprehensive econometric models reflecting the hypotheses. It is only *after* the estimation of comprehensive econometric models from databases such as Compustat that investigators can satisfactorily test the validity of specific hypotheses using statistical procedures.

As an illustration of the importance of methodological aspects for drawing inferences, it would be helpful to reappraise the work by Grullon and Michaely (2002) claiming empirical support for the "substitution" hypothesis between dividend payments and share repurchases. First, as noted in the Introduction, the correlations between dividend payments and share repurchases estimated from Compustat data were *positive* so the interrelationships are likely to be complex. Second, if dividends and share repurchases are substitutes, then they are likely to be simultaneously determined; modeling the interrelationships would require the use of econometric methods for handling endogenous explanatory variables in longitudinal analyses. Also, longitudinal data need to cover the *same* firms over time in order to shed light on the interrelationships. While such issues were not considered by Grullon and Michaely (2002), the authors explained forecasting errors in changes in dividends based on share repurchases and interpreted negative correlations as evidence for substitution hypothesis. However, dividends are announced quarterly, whereas share repurchases are made at 2–3-yearly intervals. Thus, the stochastic properties of variables cast doubt on adequacy of the models that explain a higher frequency variable such as dividends by the slowly changing share repurchases. In fact, it is likely that dividend payments that are made on a quarterly basis affect share repurchases but not vice versa. As shown below, firms' share repurchases are constrained by dividend payments and repurchases are relatively flexible (Bhargava 2010). Overall, from a methodological standpoint, it is important in financial analyses to develop comprehensive models *prior* to the testing of specific hypotheses.

38.3 Analytical Framework for Modeling the Dividends and Share Repurchases Interrelationships

38.3.1 Background Issues

The earliest dynamic model for dividends is given by Lintner (1956):

$$D_{i\,t} = a_{i\,t} + b\,P_{i\,t} + d\,D_{i\,t-1} + u_{i\,t} \quad (i = 1, 2, \ldots, H; t = 2, 3, \ldots, T) \tag{38.1}$$

For firm i in time period t, $D_{i\,t}$ is the dividend payment, $P_{i\,t}$ is profit or earnings, and $u_{i\,t}$ is a random error term; it is assumed that there are H firms in the sample that are observed in T time periods. The subsequent work by Fama and Babiak (1968) expressed dividends and earnings in per share terms. An important feature of these models is that the short-run effect of a unit increase in profit is given by b, while the long-run equilibrium impact is $[b/(1 - d)]$ with $|d| < 1$.

Explanatory variables such as firms' debt, assets, and investments can be included in Eq. 38.1. However, the estimation methods employed in previous research (Lintner 1956; Fama and Babiak 1968; Lee et al. 1987) were appropriate for time series data on a single firm rather than for a panel of firms. Because firms in Compustat and other databases are heterogeneous, it is important to include firm-specific "random" or "fixed" effects in the model. For example, treating the unobserved between-firm differences as randomly distributed variables, error terms $u_{i\,t}$ can be decomposed as:

$$u_{i\,t} = \delta_i + v_{i\,t} \tag{38.2}$$

where, δ_i are firm-specific random effects that follow some distribution (e.g., normal) with zero mean and constant variance, and $v_{i\,t}$ are distributed with zero mean and constant variance (Anderson and Hsiao 1981; Bhargava and Sargan 1983). Moreover, one can invoke the general assumption that $u_{i\,t}$'s are drawings from a multivariate distribution with a symmetric and positive definite dispersion matrix.

38.3.2 Some Conceptual Aspects of Dividends and Share Repurchase Interrelationships

From a conceptual standpoint, one would expect comprehensive models for dividends or dividends per share to resemble the dynamic model in Eq. 38.1, with additional explanatory variables such as firms' assets, debt, and investments (Bhargava 2010). Similar models can be developed for share repurchases that have become popular following the 1986 Tax Reform Act in the United States. However, there are likely to be asymmetries in the interrelationships between dividend payments and share repurchases. For example, firms regularly paying dividends may be reluctant to lower dividends in order to increase share repurchases since that might send ambiguous signals to investors. Thus, higher dividend payments are likely to lower firms' ability to make repurchases, i.e., in a model for repurchases, dividend payments are likely to be estimated with a negative coefficient. By contrast, decisions to make repurchases may be influenced by firms' unexpectedly large cash holdings (Guay and Harford 2000; Brav et al. 2005; Chen and Wang 2012) so that in a model for dividends, coefficient of share repurchases need not be statistically significant. These conceptual aspects can be embedded in a simultaneous equation model taking into account endogeneity of dividend payments and share repurchases (Sect. 38.5, below).

38.4 Processing the Compustat Database from the United States

Databases such as Compustat compile detailed information on accounting and finance variables for large numbers of firms listed on stock exchanges. However, many firms merge or exit after some years and are removed from the database. Cross-sectional studies analyzing data on firms for different years are therefore likely to include many firms that will be dropped in subsequent periods. From the standpoint of modeling the proximate determinants of dividends and share repurchases, this raises several issues. First, firms staying on the stock market for only a few time periods may typically pay small or no dividends. Also, the interrelationships between dividend payments and share repurchases require a longer time frame since dividends per share are announced on a quarterly basis, while share repurchases are made roughly at 2-year intervals. Thus, while cross-sectional analyses can provide insights, the interrelationships between dividend payments and share repurchases require that the *same* firms are observed for a certain number of years.

Second, there are trade-offs between analyzing longitudinal data that span long versus short time periods. If, for example, the data cover long periods of (say) over 30 years, then the number of firms in the sample is likely to be small and mainly the well-established firms will be included. By contrast, if one analyzes longitudinal data for only 5 years, then many firms included in the sample will be dropped from the stock exchange in later periods. In view of the fact that share repurchase decisions are made roughly at 2-year intervals, a reasonable approach would be to analyze data covering around 15 years. Moreover, one can create 2-yearly averages so that if observations were missing for a firm in one of the years, the firm could still be included in the sample. Since the data on executive compensation are available from 1992 (ExecuComp 2008), the sample period 1992–2007 is appealing (Bhargava 2011).

Third, even when the same firms are retained in the sample, there is large intra- and interfirm variation in financial variables. In the context of dividends, some problems can be tackled by expressing variables in per share terms (Fama and Babiak 1968). However, as recognized in the early statistical literature by Pearson (1897) and Neyman (1952), transforming variables into ratios can induce spurious correlations. Such practices were criticized by Kronmal (1993), Bhargava (1994) developed a likelihood ratio statistic for testing the restrictions on coefficients that enable expressing variables as ratios in empirical models. While research in finance often expresses variables in ratio forms, such transformation can affect the magnitudes and signs of estimated regression coefficients. It is important to conduct robustness check to ensure that transformation of variables do alter the conclusions.

Finally, there is often high inter- and intra-firm variation in variables such as share repurchases, assets, and debt that are measured in millions of dollars in the Compustat database. Transformations to natural logarithms can reduce internal variation in the data. For example, share repurchase can be expressed in dollars and then transformed into natural logarithms with zero dollar values set equal to one. This procedure facilitates the estimation of model parameters using numerical optimization techniques and obviates the need for arbitrarily truncating large values of variables (Bhargava 2010).

38.5 Comprehensive Empirical Models for Dividends and Share Repurchase Interrelationships

A dynamic random effects model for dividends per share adjusted for "stock splits" using cumulative adjustment factor (#27), with Compustat item numbers next to the variables, can be written as (Bhargava 2010):

$$\begin{aligned}
&(\text{Dividends per share; \#26, \#27})_{i\,t} = a_0 + a_1(\text{Earnings; \#18, \#17})_{i\,t} + a_2 \ln(\text{Total assets; \#6})_{i\,t} \\
&\quad + a_3(\text{Market-to-book value; \#199, \#25, \#60})_{i\,t} + a_4 \ln(\text{Long} - \text{term debt; \#9})_{i\,t} \\
&\quad + a_5 \ln(\text{Short-term investments; \#193})_{i\,t} + a_6(\text{Time dummy period 3})_{i\,t} \\
&\quad + a_7(\text{Time dummy period 4})_{i\,t} + a_8(\text{Time dummy period 5})_{i\,t} \\
&\quad + a_9(\text{Time dummy period 6})_{i\,t} + a_{10}(\text{Time dummy period 7})_{i\,t} \\
&\quad + a_{11}(\text{Time dummy period 8})_{i\,t} + a_{12}(\text{Dividends per share; \#26, \#27})_{i\,t-1} \\
&\quad + u_{i\,t}\ (i = 1, 2, \ldots, H;\, t = 2, 3, \ldots, 8)
\end{aligned} \tag{38.3}$$

In the empirical model in Eq. 38.3, it is recognized that firms sometimes split their shares especially if share prices are high. Earnings are defined as "income before extraordinary items" (#18) minus 0.6 times "special items" (#17), measured in million of dollars (Skinner 2008). Income before extraordinary items is the firms' incomes taking into account all expenses, while special items reflect adjustments for debts and losses. “Total assets” (#6), “long-term debt” (#9), and “short-term investments” (#193) were expressed in dollars and then converted into natural logarithms. The “market-to-book value” variable was constructed by multiplying “share price” (#199) by number of “common shares outstanding” (#25) and dividing this product by “common equity total” (#60). “Dividends per share” in period $(t-1)$ will be referred to as “lagged dependent variable.” Models of the type in Eq. 38.3 were also estimated for dividends expressed as ratios to firms’ book values, while dropping the explanatory variable market-to-book value from Eq. 38.3.

It is assumed in Eq. 38.3 that 8-time observations at 2-yearly intervals {1992–1993, 1994–1995, 1996–1997, 1998–1999, 2000–2001, 2002–2003, 2004–2005, 2006–2007} were available on H firms. Moreover, as explained in Sect. 38.6 and the Appendix, initial values of the dependent variable in time period 1 were modeled using a “reduced form” equation that includes a separate coefficient, c_0, for the constant term. Thus, the model in Eq. 38.3 includes separate coefficients of the constant term for each of the 8-time periods. This formulation allows the variables in the model to have different means in the 8-time periods and is useful if there are trends in the dependent variables in the observation period.

Further, $u_{i\,t}$’s are error terms that can be decomposed in the simple random effect fashion as in Eq. 38.2. The variance of $v_{i\,t}$ is the “within” (or intra-) firm variance, and [var (δ_i)/var (v_{it})] is the “between-to-within” variance ratio that can be estimated when errors are decomposed as in Eq. 38.2. In another version of the model in Eq. 38.3, firms’ intangible assets (Compustat item #33) were included in the model, though greater numbers of observations were missing for this variable. Also, firms’ share repurchases were introduced as potentially endogenous explanatory variables to test if they influenced dividend payments. Last, static versions of the model in Eq. 38.3 that excluded the lagged dividends were estimated to assess robustness of the results.

The model for share repurchases was similar to that in Eq. 38.3 except that dividends per share were included as a potentially endogenous explanatory variable:

$$\begin{aligned}
&\ln(\text{Share repurchases};\#226,\#115,\#108)_{i\,t} = b_0 + b_1(\text{Earnings};\#18,\#17)_{i\,t}\\
&\quad + b_2 \ln(\text{Total assets};\#6)_{i\,t} + b_3\,(\text{Market-to-book value};\#199,\#25,\#60)_{i\,t}\\
&\quad + b_4 \ln(\text{Long-term debt};\#9)_{i\,t} + b_5 \ln(\text{Short-term investments};\#193)_{i\,t}\\
&\quad + b_6\,(\text{Time dummy period 3})_{i\,t} + b_7\,(\text{Time dummy period 4})_{i\,t} + b_8(\text{Time dummy period 5})_{i\,t}\\
&\quad + b_9\,(\text{Time dummy period 6})_{i\,t} + b_{10}\,(\text{Time dummy period 7})_{i\,t} + b_{11}(\text{Time dummy period 8})_{i\,t}\\
&\quad + b_{12}\,(\text{Dividends per share};\#26,\#27)_{i\,t} + b_{13} \ln(\text{Share repurchases};\#226,\#115,\#108)_{i\,t-1}\\
&\quad + u_{2\,i\,t} \quad (i = 1,2,\ldots,H;\, t = 2,3,\ldots,8)
\end{aligned} \tag{38.4}$$

Note that share repurchases were defined as the change in “Treasury stock” (#226), and if this was less than or equal to zero for the firm, the nonnegative difference between “purchase of common stock” (#115) and “sale of common stock” (#108) was used (Fama and French 2001). This definition covered the situation where firms "retire" shares from Treasury stock so that when the change in Treasury stock was less than or equal to zero, nonnegative difference between purchase of common stock and sale of common stock was used as the measure of repurchases. Banyi et al. (2008) have advocated the use of purchase of common stock as a measure of repurchases, and this variable was also modeled to assess robustness of the results. Finally, alternative models for share repurchases were estimated with intangible assets included as an explanatory variable, and also where dividends per share was replaced by dividends expressed as ratios to firms' book values.

38.6 The Econometric Framework for Addressing Simultaneity and Between-Firm Heterogeneity

The methodology used for estimation of dynamic and static random effect models, where some explanatory variables are endogenous, was developed in Bhargava and Sargan (1983) and Bhargava (1991). Let the dynamic model be given by:

$$y_{i\,t} = \sum_{j=1}^{m} z_{i\,j}\gamma_j + \sum_{j=1}^{n_1} x_{1\,i\,j\,t}\beta_j + \sum_{j=n_1+1}^{n} x_{2\,i\,j\,t}\beta_j + \alpha\, y_{i\,t-1} + u_{i\,t} \quad (i = 1, 2, \ldots, H;\, t = 2, 3, \ldots, T) \tag{38.5}$$

where z's are time-invariant variables, x_1 and x_2 are, respectively, n_1 exogenous and n_2 endogenous time-varying variables ($n_1 + n_2 = n$). In the model for share repurchases, for example, dividends per share is potentially an endogenous time-varying variable; unobserved factors affecting share repurchases, reflected in firm-specific random effects (δ_i) in Eq. 38.2, can influence dividend payments. For exposition purposes, first assuming that all time-varying variables are exogenous so that $n_2 = 0$ and subscripts on the x variables can be dropped, the dynamic model can be written in a simultaneous equations framework by defining a "reduced form" equation for initial observations that do not include any endogenous explanatory variables, and a "triangular" system of (T-1) "structural" equations for the remaining periods (Bhargava and Sargan 1983):

$$y_{i\,1} = \sum_{j=1}^{m} z_{i\,j}\zeta_j + \sum_{j=1}^{n}\sum_{k=1}^{T} \upsilon_{j\,k} x_{i\,j\,k} + u_{i\,1} \quad (i = 1, \ldots, H) \tag{38.6}$$

And

$$\underset{(T-1)\mathrm{x}T}{B}\ \underset{T\mathrm{x}H}{Y'} + \underset{(T-1)\mathrm{x}\,m}{C_z}\ \underset{m\mathrm{x}H}{Z'} + \underset{(T-1)\mathrm{x}\,nT}{C_x}\ \underset{nT\mathrm{x}H}{X'} = \underset{(T-1)\mathrm{x}H}{U'} \tag{38.7}$$

Here, Y, Z, and X are, respectively, matrices containing observations on the dependent, time-invariant, and time-varying explanatory variables; dimensions of the matrices are written below the respective symbols. B is a (T−1) x T lower triangular matrix of coefficients:

$$B_{i\,i} = \alpha, B_{i,i+1} = -1, B_{i\,j} = 0 \text{ otherwise} \quad (i = 1, \ldots, T-1; j = 1, \ldots, T) \tag{38.8}$$

where α is the coefficient of lagged dependent variable. Matrices C_z and C_x contain coefficients of time-invariant and time-varying regressors, respectively; U contains the error terms.

Note that in the reduced form Eq. 38.6, all T realizations of the n exogenous time-varying variables appear as explanatory variables for modeling the systematic or predicted part of $y_{i\ 1}$. Because maximum likelihood or instrumental variable methods treat $\{y_{i\ 1}, y_{i\ 2}, \ldots, y_{i\ T}\}$ as jointly determined variables, it is essential to explain the reduced form Eq. 38.6 using all exogenous variables in the model. The estimation of model parameters and certain econometric tests are described in the Appendix.

38.7 Descriptive Statistics from the Annual Compustat Database

The Compustat annual data on firms for the period 1992–2007 were processed for the analyses (Bhargava 2010). As is customary in the dividends literature, financial firms with Standard Industry Classification (SIC) codes between 6,000 and 6,999 and utility firms with SIC codes between 4,900 and 4,949 were dropped since there are regulations on dividend payments by such firms (Fama and French 2001). Moreover, observations were retained on firms that were in the Compustat database for at least 14 years, and 2-yearly averages were created at 8-time points, i.e., for {1992–1993, 1994–1995, 1996–1997, 1998–1999, 2000–2001, 2002–2003, 2004–2005, 2006–2007}. An alternative data set created 3-yearly averages at 5-time points, i.e.,{1992–1994, 1995–1997, 1998–2000, 2001–2003, 2004–2006}, though the analyses of the 2-yearly averages led to more robust parameter estimation. Averaging over 2 years led to a sample of 3,290 industrial firms observed in 8-time periods. Also, estimation of dynamic models requires "balanced" panels, i.e., where firms are included at all eight time points. While it is possible to analyze "unbalanced" panels using static models in software packages such as Stata (2008), it is difficult to address endogeneity issues in the analyses.

The sample means and standard deviations of 2-year averages of variables are reported in Table 38.1 for some of the years with Compustat annual item numbers noted next to the variables. The mean dividends per share were approximately \$0.32 for the period 1992–1999, and declined slightly to \$0.27 for 2000–2003, and increased to \$0.39 in 2006–2007. While a simple paired *t*-test on the set of firms' observations showed significant ($P < 0.05$) increase in dividends per share between

Table 38.1 Sample means and standard deviations of 2-yearly averages of selected variables from the Compustat database for up to 3,290 US industrial firms in 1992–2007

	1992–1993		1998–1999		2002–2003		2006–2007	
Variable	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Income before extraordinary items (#18), $ million	53.99	317.55	138.05	702.98	124.09	1,015.57	309.43	1,818.94
Special items (#17), $ million	−15.99	215.30	−1.56	372.44	−55.10	552.09	−28.20	731.89
Treasury common stock (#226), $ million	37.08	398.01	108.79	852.50	194.16	1,461.74	406.76	2,994.53
Purchase of common stock (#115), $ million	10.39	66.72	54.54	282.34	51.85	349.32	221.96	1,538.69
Sale of common stock (#108), $ million	21.66	112.12	31.07	160.90	25.69	104.93	72.96	1,022.77
Share repurchases (#226, #115, #108), $ million	9.95	111.13	44.44	316.76	49.49	467.39	174.77	1,564.85
Dividends (#21), $ million	32.40	166.22	63.22	291.33	63.09	376.21	113.18	629.87
Dividends per share (#26), $	0.304	0.902	0.298	0.882	0.281	1.272	0.390	1.095
Total assets (#6), $ million	2,292.84	13,195.74	3,845.47	23,638.40	5,491.57	33,186.31	8,083.80	61,202.42
Cumulative adjustment factor (#27)	2.48	4.48	1.41	1.24	1.18	0.72	1.03	0.89
Share price (#199), $	17.55	19.61	19.40	27.03	17.21	25.29	23.55	42.62
Long-term debt (#9), $ million	429.26	2,417.02	695.61	3,664.86	1,070.41	6,085.70	1,519.05	11,825.75
Short-term investments (#193), $ million	88.31	1,181.56	173.91	2,767.14	238.97	3,735.60	368.76	8,385.86
Intangible assets (#33), $ million	99.91	746.35	279.68	1,279.77	651.12	4,629.00	980.21	5,507.11
Common shares outstanding (#25), million	39.53	133.66	84.22	279.68	159.55	1,682.32	177.41	1,684.17
Common equity total (#60), $ million	541.70	2,214.41	958.11	3,917.14	1,351.86	6,265.71	1,985.57	8,140.76
Percentages of firms paying dividends[a]	36.50	–	40.79	–	39.89	–	44.00	–
Percentages of firms making repurchases[a]	26.29	–	49.82	–	37.93	–	43.46	–

Means and standard deviations of variables using 2-yearly averages at four time points for up to 3,290 US industrial firms observed for at least 14 years in 1992–2007 (Source: Bhargava 2010); specific number of firms included depends on the extent of unavailable information for each variable in each period; Compustat items and units of measurements are listed next to variables

[a]Percentages are based on firms' with nonzero payments

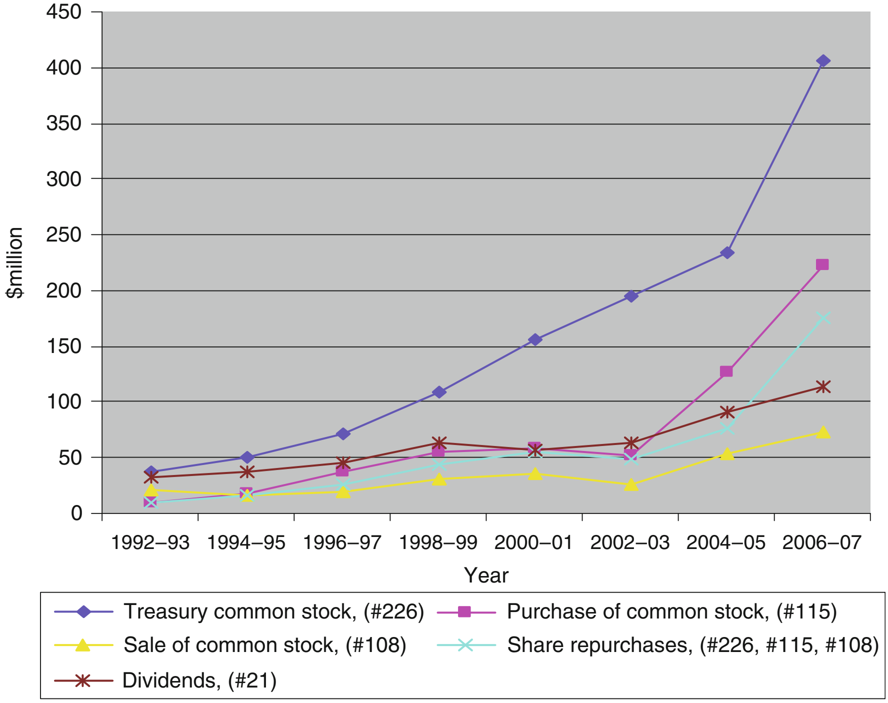


Fig. 38.1 Sample means of dividends and components of share repurchases by US industrial firms for 1992–2007. Share repurchases were calculated as the change in Treasury stock (#226), and if this was less than or equal to zero, the nonzero difference between purchase of common stock (#115) and sale of common stock (#108) was used (Source: Bhargava 2010)

1992–1993 and 2006–2007, the pattern over time was more complex partly because of tax changes in 2003. Mean share repurchases increased from \$10 million in 1992–1993 to \$175 million in 2006–2007.

Figure 38.1 plots the sample means (in \$ millions) of Treasury stock, purchase of common stock, sale of common stock, dividends, and share repurchases. There were upward trends in Treasury stock, share repurchases, and purchases of common stock. There was an increase in dividends in the period from 2002–2003 to 2006–2007, where dividends increased from \$63 million to \$113 million. This could be due to reductions in 2003 in tax rates on dividends (Julio and Ikenberry 2004).

Further, focusing on 2,880 firms with non-missing observations on dividends and share repurchases, percentages of firms paying dividends and making share repurchases in the 8-time periods, were {36.5, 39.9, 40.6, 40.8, 39.6, 39.9, 43.7, 44.0} and {26.3, 30.7, 38.2, 49.8, 48.1, 37.9, 38.1, 43.5}, respectively. While percentages of firms making share repurchases increased from 26.3 in 1992–1993 to 43.5 in 2006–2007, the peak of 49.8 % was reached in the 1998–1999 period. The mean share repurchases in 1998–1999 were 44.44 million, while they amounted to \$174.77 million in 2006–2007 so that share repurchases were significantly higher in 2006–2007 after adjusting for the price level. The other salient feature of the Compustat data was that the mean intangible assets increased from \$99.9 million in 1992–1993 to \$980.2 million in 2006–2007 constituting a tenfold increase.

38.8 Results from Simple Dynamic Models for Dividends and Share Repurchase Controlling for Between-Firm Heterogeneity

The results from estimating simple dynamic random effects models for dividends per share are in Table 38.2 for the sample of 3,113 industrial firms (Bhargava 2010). The results are presented for the cases where firms paid nonzero dividends at least in one of the eight time periods and where firms paid dividends in all 8-time periods. The results for ratios of dividends to book value of firms are presented in the last column of Table 38.2. A set of six dummy variables for time periods 3–8 was included to account for differences in dividend payments in the 8-time periods. The models were estimated by maximum likelihood and provide consistent and efficient estimates of the parameters.

For the pooled sample, coefficients of the dummy variables in the model for dividends per share were positive and significant ($P < 0.05$) in time periods 7 and 8 corresponding to 2004–2005 and 2006–2007, respectively. The positive coefficients were consistent with reported increases in dividends following reduction in tax rates in 2003. Coefficients of the dummy variables were qualitatively similar for the reduced sample of 1,827 firms that paid dividends at least in one of the eight periods and in the case where 1,035 firms paid dividends in all eight periods. Moreover, coefficients of the dummy variables for time periods 7 and 8 corresponding to 2004–2005 and 2006–2007, respectively, were approximately twice as large for firms paying dividends than for firms that never paid dividends. Thus, reduction in tax rates on dividends in 2003 appeared to have increased dividends per share. These results also show the robustness of dynamic models in situations where firms decide against paying dividends in some time periods, i.e., to zero values of the dependent variables. This is not surprising since one is modeling the deviations of dividends per share from an overall mean.

The coefficients of the lagged dependent variables for the first three cases presented in Table 38.2 were close to 0.28. Moreover, the ratios of between-to-within variances for the three cases, i.e., pooled sample, firms paying nonzero dividends at least once, and firms paying dividends in all 8 periods were, 0.38, 0.33, and 0.45, respectively. These estimates were close despite reduction in sample size from 3,113 firms in the pooled sample to 1,035 in the case of firms paying nonzero dividends in all 8 periods. The within firm variances for the three cases were 0.52, 0.88, and 1.07, respectively. Finally, in the last column, the results for the ratio of dividends to firms' book values were similar to those for the pooled sample for dividends per share. While the estimated coefficient of lagged dependent

Table 38.2 Maximum likelihood estimates of simple dynamic models for 2-yearly averages of dividend paid by US industrial firms in 1992–2007

	Dependent variable: dividends per share (Compustat items #26/#27)						Dependent variable:dividends/ book value (#21,#60)	
	All firms		Firms paying nonzero dividends at least once		Firms paying nonzero dividends in all 8 periods		All firms	
Explanatory variables	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Constant	0.172	0.015	0.294	0.018	0.421	0.031	0.018	0.004
Time period 3 dummy variable	0.017	0.020	0.029	0.026	0.029	0.046	0.001	0.006
Time period 4 dummy variable	0.033	0.020	0.058*	0.026	0.081	0.046	0.003	0.006
Time period 5 dummy variable	−0.002	0.019	−0.003	0.026	0.010	0.046	−0.0002	0.006
Time period 6 dummy variable	0.017	0.019	0.030	0.026	0.082	0.046	−0.00002	0.006
Time period 7 dummy variable	0.087*	0.020	0.148*	0.027	0.130*	0.046	0.005	0.006
Time period 8 dummy variable	0.133*	0.019	0.228*	0.027	0.235*	0.046	0.025*	0.006
Lagged dependent variable	0.285*	0.009	0.283*	0.010	0.258*	0.013	0.290*	0.018
(between-to-within) variance ratio	0.375*	0.017	0.328*	0.019	0.451*	0.034	0.016*	0.007
Within variance	0.518	–	0.879	–	1.073	–	0.040	–
2 x (maximized log-likelihood function)	11,552.55	–	−741.58	–	−2,301.34		58,929.64	
Number of firms	3,113		1,827		1,035		2,315	

Dependent variable is dividends per share (#26) divided by cumulative adjustment factor (#27) (Source: Bhargava 2010); 8-time observations at 2-yearly intervals on firms were used; slope coefficients and standard errors are reported

*$P < 0.05$

Table 38.3 Maximum likelihood estimates of simple dynamic models for 2-yearly averages of share repurchases of US industrial firms in 1992–2007

	Dependent variable: ln (share repurchases $) (using Compustat items #226, #115, #108)					
	All firms		Firms making nonzero repurchases at least once		Firms making nonzero repurchases in all 8 periods	
Explanatory variables	Coefficient	SE	Coefficient	SE	Coefficient	SE
Constant	3.216	0.061	4.057	0.153	4.376	0.587
Time period 3 dummy variable	1.008*	0.122	1.262*	0.205	−0.054	0.192
Time period 4 dummy variable	2.428*	0.117	3.048*	0.200	0.067	0.193
Time period 5 dummy variable	1.426*	0.126	1.818*	0.208	−0.128	0.197
Time period 6 dummy variable	−0.095	0.117	−0.083	0.203	−0.540*	0.183
Time period 7 dummy variable	0.738*	0.122	0.938*	0.208	−0.017	0.192
Time period 8 dummy variable	1.730*	0.123	2.180*	0.205	0.171	0.194
Lagged dependent variable	0.362*	0.008	0.354*	0.009	0.754*	0.038
(between-to-within) variance ratio	0.153*	0.010	0.066*	0.008	0.184*	0.092
Within variance	39.176	–	48.341	–	1.841	–
2 x (maximized log-likelihood function)	−88,260.15	–	−73,671.73	–	−886.97	
Number of firms	2,907		2,329		127	

Dependent variable share repurchases were calculated from Compustat items #226, #115, and #108 and expressed in dollars and transformed into natural logarithms with zero value assigned to zero purchases (source: Bhargava 2010); 8-time observations at 2-yearly intervals on the firms in 1992–2007 were used; slope coefficients and standard errors are reported

*$P < 0.05$

variable was very close (0.29), the estimated between-to-within variance ratio was 0.016 that was lower than the corresponding estimate (0.38) in the model for dividends per share.

The results from simple dynamic models for logarithms of US firms' share repurchases are in Table 38.3 for the three cases, i.e., pooled sample, firms making at least one repurchase, and firms making repurchase in all 8-time periods. For the pooled sample of 2,907 firms, coefficients of dummy variables for five of the time periods were positive and statistically significant; coefficient of the dummy variable for time period six corresponding to 2002–2003 was not significant. The results were very similar for the case where 2,329 firms made repurchases at least once. The dummy variables were generally insignificant in the third case for firms making repurchases in all 8 periods; there were 127 firms in this group which was a small sample size for estimating dynamic models.

The coefficients of lagged dependent variables in the three cases were 0.36, 0.35, and 0.75, respectively; these estimates were significantly less than unity thereby indicating that share repurchase series were stationary. The between-to-within variance ratios for the three cases were 0.15, 0.07, and 0.18, respectively; within firm variances were quite large in these models. The higher coefficient of lagged dependent variable in the model where 127 firms made repurchases in all 8 periods may be due to the small sample size. Moreover, such firms may have been less heterogeneous in some respects and had a smooth pattern of share repurchases. Such issues can be systematically investigated by controlling for firm characteristics and are addressed below.

38.9 Results from Comprehensive Dynamic Models for Dividends and Share Repurchases

The results from empirical models for dividends per share and share repurchases, outlined in Eqs. 38.3 and 38.4, are presented in this section. Dynamic and static random effects models were estimated and the results are initially presented for the pooled samples and for the case where firms paid dividends or made repurchases at least once during 8-time periods.

38.9.1 Results from Dynamic and Static Models for Firms' Dividends per Share

Table 38.4 presents the results from dynamic and static models for dividends per share paid by US industrial firms; a set of six dummy variables for time periods was included in the models though the coefficients are not reported (Bhargava 2010). The firms' total assets, long-term debt, and short-term investments were expressed in dollars and converted to natural logarithms. Because firms' earnings were defined as the difference between income before extraordinary items and special items, this variable sometimes assumed negative values and was not transformed into logarithms.

For the dynamic model estimated using the pooled sample, coefficient of earnings was estimated with a positive coefficient that was statistically significant. The coefficients of earnings and total assets were also significant in the static model for the pooled sample and in dynamic and static models for the subsample where firms paid dividends at least once during the 8-time periods. These results provide evidence that earnings and total assets of US firms were positively and significantly associated with dividends per share. By contrast, Skinner (2008, Table 8) found these variables to be generally insignificant predictors of the ratio of dividends to total payout (dividends plus repurchases) using cross-sectional regressions and data on 345 US firms. When some of the coefficients were statistically significant, they were positive for some years and negative for others. The contradictory findings were likely to be due to combining dividends and share repurchases into a single variable and using estimation methods that are appropriate for cross-sectional analyses. The results in Table 38.4, however, were consistent with evidence from the European Union (von Eije and Megginson 2008), where static random effects models were estimated using unbalanced panel data on approximately 3,000 firms.

Table 38.4 Maximum likelihood estimates from comprehensive dynamic and static random effects models for 2-yearly averages of dividends per share paid by US industrial firms in 1992–2007

	Dependent variable: dividends per share (adjusted for stock splits; Compustat items # 26/#27)							
	Dynamic model: all firms		Static model: all firms		Dynamic model: firms paying at least once		Static model: firms paying at least once	
Explanatory variables:	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Constant	−9.383	2.664	−9.169	0.464	−9.357	1.306	−11.091	0.817
Earnings (#18, #17), $million	0.0001*	0.00001	0.0001*	0.00001	0.0002*	0.00001	0.0001*	0.00001
ln (total assets) (#6), $	0.058*	0.019	0.060*	0.003	0.054*	0.010	0.072*	0.006
Market-to-book value (#199, #25, #60)	0.0004	0.004	0.0004	0.0004	0.0005	0.0008	0.001	0.001
ln (long-term debt) (#9), $	0.0007	0.009	−0.0008	0.0006	0.001	0.003	−0.002	0.001
ln (short-term investments) (#193), $	0.002	0.007	0.0006	0.0005	0.005*	0.002	0.002*	0.001
Lagged dependent variable	0.201*	0.021	–		0.193*	0.014	–	
(between-to-within) variance ratio	0.126*	0.014	–		0.114*	0.014	–	
Within variance	0.611	–	–		0.980	–	–	
2x (maximized log-likelihood function)	5,878.43		–		−630.16	–	–	
Number of firms	1,854		1,854		1,144		1,144	

Dependent variable is dividends per share (#26) divided by cumulative adjustment factor (#27) (Source: Bhargava 2010); 8-time observations at 2-yearly intervals on firms in 1992–2007 were used; slope coefficients and standard errors are reported; time dummies for 6 periods were included though their coefficients are not reported in the tables

*$P < 0.05$

The variable market-to-book value was not statistically significant in the dynamic and static models estimated from the pooled or disaggregated samples. Because market-to-book value is a composite variable based on Compustat items price, common shares outstanding, and common equity total, it may be difficult to unscramble the effects on dividend payments. Coefficients of long-term debt were not statistically significant in the four cases reported in Table 38.4. Because firms may be reluctant to frequently alter dividend policy, it is likely that long-term debt might influence share repurchases rather than dividends. Coefficients of short-term investments were positively and significantly associated with dividends per share in the dynamic and static models estimated using the subsample of firms that paid dividends at least once.

The coefficients of lagged dependent variables from the pooled sample and the subsample of firms paying dividends at least once were 0.20 and 0.19, respectively. These coefficients implied that the long-run effects of explanatory variables were 1.25 times the respective short-run effects reported in Table 38.4. Moreover, the ratios of between-to-within variance in the two dynamic models were 0.13 and 0.11, respectively. While these ratios were smaller than those reported in Table 38.2 for simple dynamic models for dividends per share (0.38 and 0.33, respectively), the estimates suggest that some heterogeneity across firms could not be accounted for by explanatory variables in the models. The within firm variances in the two dynamic models were 0.61 and 0.98. Overall, the results in Table 38.4 supported the view that larger firms paid higher dividends per share.

38.9.2 Results from Dynamic and Static Models for Firms' Share Repurchases

The results for logarithms of share repurchases by US industrial firms are in Table 38.5 for the pooled sample and for firms making repurchases at least once in the sample period. As in the results from the model for dividends per share, firms' earnings and total assets were estimated with significant positive coefficients in dynamic and static models for the pooled sample and for the subsample of firms making repurchases at least once. Coefficients of the logarithm of total assets in dynamic models were the short-run "elasticities" of share repurchases with respect to total assets, i.e., 0.61 and 0.78, respectively, for the pooled sample and subsample of firms making repurchases. Because coefficients of lagged dependent variables were approximately 0.40 in this model, the long-run effects were about 1.67 times the respective short-run effects. For example, the long-run elasticity of share repurchases with respect to total assets was 1.02 and this was larger than the point estimate (0.70) from the static model for the pooled sample. Similarly, short- and long-run elasticities of share repurchases with respect to total assets for firms making repurchases at least once were 0.78 and 1.33, respectively. Thus, the size of the firm was an important determinant of magnitudes of share repurchases. However, there were no nonlinearities apparent with respect to firms' earnings and total assets, i.e., squared terms of these variables were not significant predictors of share repurchases.

Table 38.5 Maximum likelihood estimates from comprehensive dynamic and static random effects models for 2-yearly averages of share repurchases by US industrial firms in 1992–2007

	Dependent variable: ln (share repurchases $) (using Compustat items #226, #115, #108)							
	Dynamic model: all firms		Static model: all firms		Dynamic model: firms repurchasing least once		Static model: firms repurchasing at least once	
Explanatory variables	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Constant	−99.008	6.428	−114.142	7.960	−116.688	7.620	−147.576	9.851
Earnings (#18, #17), $million	0.0004*	0.0001	0.0008*	0.0001	0.0004*	0.0001	0.0007*	0.0001
ln (total assets) (#6), $	0.613*	0.041	0.703*	0.051	0.776*	0.052	0.961*	0.066
Market-to-book value (#199, #25, #60)	−0.010*	0.004	−0.010*	0.003	−0.011*	0.005	−0.012*	0.005
ln (long-term debt) (#9), $	−0.021*	0.008	−0.020*	0.010	−0.054*	0.012	−0.053*	0.017
ln (short-term investments) (#193), $	0.001	0.006	0.003	0.008	−0.020	0.008	−0.019	0.012
Dividends per share (#26/#27), $	−0.201*	0.057	−0.139*	0.059	−0.179*	0.078	−0.153*	0.077
Lagged dependent variable	0.404*	0.011	–		0.393*	0.011	–	
(between-to-within) variance ratio	0.168*	0.012	–		0.014*	0.005	–	
Within variance	29.619	–	–		46.627	–	–	
2 x (maximized log-likelihood function)	−52,088.12		–		−35,401.14		–	
Chi-square test exogeneity of dividends (8df)	10.15		11.53		12.14		11.26	
Number of firms	1,845		1,845		1,141		1,141	

Dependent variable share repurchases were based on items #226, #115, and #108 (Source: Bhargava 2010); 8-time observations at 2-yearly intervals on the firms in 1992–2007 were used; slope coefficients and standard errors are reported; time dummies for 6 periods were included though their coefficients are not reported in the tables

*$P < 0.05$

In all the models in Table 38.5, firms' market-to-book value ratio was estimated with negative and significant coefficients in the dynamic and static models using the pooled and disaggregated samples. Thus, firms with higher market-to-book value ratio made lower share repurchases. Also, firms' long-term debt was estimated with negative and significant coefficients in dynamic models from the pooled sample and from the subsample of firms making repurchases at least once. The short-run elasticity from the pooled sample was −0.02, while the long-run elasticity was −0.03. Doubling of firms' long-term debt predicted a 3 % decline in long-run share repurchases. The coefficients of short-term investments were not significant in any of the models in Table 38.5.

An important aspect of the results in Table 38.5 was that the estimated coefficients of dividends per share were negative and statistically significant in the dynamic and static models for the pooled sample and for the subsample of firms that made at least one repurchase. Thus, controlling for factors such as earnings and size of the firm, firms paying higher dividends per share made significantly lower share repurchases. The short-run effects of dividends per share on share repurchases were −0.20 and −0.18, respectively, for the pooled sample and subsample of firms making repurchases; the respective long-run effects were −0.33 and −0.30. The relatively large magnitudes of these effects indicated that firms with higher dividend payments were likely to make smaller share repurchases.

Last, exogeneity hypotheses for dividends per share were accepted at the 5 % level using test statistics that were distributed as Chi-square variables with 8° of freedom (5 % critical limit of Chi-square (8) = 15.5). The values of Chi-square statistics around 12 in Table 38.5 indicated that there might be some dependence in the unobserved components of firms' decisions to pay dividends and make share repurchases though it was not significant at the 5 % level. Such hypotheses will be tested below for the effects of share repurchases on firms' dividend payments.

38.9.3 Results from Dynamic and Static Models for Dividends per Share with Share Repurchases and Intangible Assets Included as Explanatory Variables

The results from models for dividends per share and ratios of dividends to firms' book value, with share repurchases included as a potentially endogenous explanatory variable, are in Table 38.6; firms' intangible assets were included in these models. Because short-term investments were not significant predictors of share repurchases in Table 38.5, this variable was dropped from the models for dividends and share repurchases in Tables 38.6 and 38.7, respectively. There were greater number of missing observations on intangible assets and sample sizes in Table 38.6 were lower than sample sizes in Table 38.4.

The noteworthy feature of the results for dividends per share in Table 38.6 was that coefficients of share repurchases was estimated with a small negative coefficient (−0.004) in the dynamic model that reached statistical significance at the 5 % level. However, this coefficient from the static model was 0.0002 and was not

Table 38.6 Maximum likelihood estimates from comprehensive dynamic and static random effects models for 2-yearly averages of dividends per share paid by US industrial firms in 1992–2007 with intangible assets and share repurchases included as explanatory variables

	Dependent variable: dividends per share (#26/#27)				Dependent variable: (dividends/book value) (#21/#60)			
	Dynamic model		Static model		Dynamic model		Static model	
Explanatory variables	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Constant	−11.348	0.844	−10.855	0.612	−0.077	0.027	−0.068	0.010
Earnings (#18, #17), $million	0.0001*	0.00001	0.0001*	0.00001	0.00001*	0.000003	0.00001*	0.000001
ln (total assets) (#6), $	0.075*	0.006	0.067*	0.004	0.005*	0.002	0.005*	0.001
Market-to-book value (#199, #25, #60)	0.001	0.001	0.0004	0.0004	–		–	
ln (long-term debt) (#9), $	0.0004	0.002	−0.0003	0.001	−0.00003	0.0005	−0.0004*	0.0001
ln (intangible assets) (#33), $	−0.004*	0.001	−0.003*	0.001	−0.0002	0.0004	−0.0001	0.0001
ln (share repurchases) (#226, #115, #108),$	−0.004*	0.002	0.0002	0.0005	−0.0003	0.0004	0.0002	0.0001
Lagged dependent variable	0.191*	0.012	–		0.255*	0.023	–	
(between-to-within) variance ratio	0.122*	0.013	–		0.012*	0.006	–	
Within variance	0.761		–		0.063	–	–	
2x (maximized log-likelihood function)	2,063.30		–		31,108.69	–	–	
Chi-square test exogeneity repurchases (8df)	6.235		6.74		13.26		11.23	
Number of firms	1,429		1,429		1,418		1,418	

Dependent variables are dividends per share (#26) divided by cumulative adjustment factor (#27) and the ratio of dividends to book value (#21/#60) (Source: Bhargava 2010); 8-time observations at 2-yearly intervals on firms in 1992–2007 were used; slope coefficients and standard errors are reported; time dummies for 6 periods were included though their coefficients are not reported

*$P < 0.05$

Table 38.7 Maximum likelihood estimates from dynamic and static random effects models for 2-yearly averages of share repurchases by US industrial firms in 1992–2007 with intangible assets included as an explanatory variable

	Dependent variable: ln (Share repurchases $) (Using Compustat items #226, #115, #108)							
	Specification 1:				Specification 2:			
	Dynamic model		Static model		Dynamic model		Static model	
Explanatory variables:	Coefficient	SE	Coefficient	SE	Coefficient	SE	Coefficient	SE
Constant	−89.993	5.547	−103.897	8.953	−81.508	4.659	−94.684	8.243
Earnings (#18, #17), $million	0.0004*	0.0001	0.001*	0.0001	0.0004*	0.0001	0.001*	0.0001
ln (total assets) (#6), $	0.524*	0.039	0.602*	0.058	0.517*	0.036	0.599*	0.058
Market-to-book value (#199, #25, #60)	−0.012*	0.0004	−0.011*	0.003	–		–	
ln (long-term debt) (#9), $	−0.027*	0.009	−0.019	0.013	−0.028*	0.009	−0.021*	0.013
ln (intangible assets) (#33), $	0.044*	0.003	0.044*	0.011	0.046*	0.004	0.046*	0.011
Dividend per share (#26/#27), $	−0.217*	0.054	−0.183*	0.058	–		–	
Dividends/Book value (#21/#60)	–		–		−0.583*	0.248	−0.537*	0.207
Lagged dependent variable	0.412*	0.012	–		0.412*	0.011	–	
(between-to-within) variance ratio	0.171*	0.014	–		0.166*	0.014	–	
Within variance	28.123	–	–		28.310	–	–	
2x (maximized log-likelihood function)	−39,778.71		–		−39,523.88		–	
Chi-square test exogeneity of dividends (8df)	9.62		11.30		10.34		15.40*	
Number of firms	1,429		1,429		1,418		1,418	

Dependent variable share repurchases were based on items #226, #115, and #108 (Source: Bhargava 2010); Specification 1 included dividends per share as explanatory variable, while Specification 2 included dividends expressed in terms of book values; 8-time observations at 2-yearly intervals on firms in 1992–2007 were used; slope coefficients and standard errors are reported; time dummies for 6 periods were included though their coefficients are not reported

*$P < 0.05$

significant. Also, coefficients of share repurchases in the model for the ratio of dividends to book value were statistically not different from zero. Exogeneity hypotheses for share repurchases could not be rejected using the Chi-square statistics in dynamic and static models for dividends per share and for the ratio of dividends to firms' book value. Because share repurchases in the comprehensive models generally did not affect dividend payments, it seems likely that firms decided their dividends payments at an earlier stage than share repurchases. By contrast, dividends per share were significant predictors of share repurchases in Table 38.5.

Another important finding in Table 38.6 was that intangible assets were estimated with negative and significant coefficients in dynamic and static models for dividends per share. The results for other explanatory variables were similar to those presented in Table 38.4, thereby showing robustness of the estimates to changes in model specification and to reductions in sample sizes due to greater number of missing observations on intangible assets.

38.9.4 Results from Dynamic and Static Models for Share Repurchases with Intangible Assets Included as Explanatory Variables

Table 38.7 presents the results for logarithms of share repurchases with intangible assets included as explanatory variables in the models; in Specification 2, dividends per share were replaced by the ratio of dividends to firms' book values. The results in Table 38.7 showed that intangible assets were positively and significantly associated with share repurchases. This was in contrast with the results in Table 38.6 where intangible assets were negatively associated with dividends per share. Thus, firms possessing higher intangible assets paid lower dividends per share and made greater share repurchases; these results indicate the inappropriateness of modeling the ratio of dividend payments to total payouts (Skinner 2008). Furthermore, while intangible assets have received attention in the literature (Lev 2001), their effects on dividends and share repurchases have not been rigorously investigated. For example, it has been suggested that information technology firms may pay low (or zero) dividends and make frequent repurchases in part because they have higher investments in research and development. A dummy variable was created using SIC codes for information technology firms in manufacturing, transportation, and communications sectors. However, the estimated coefficient was not significantly different from zero. Thus, the reported coefficients in Table 38.7 appear not to suffer from biases due to omission of such variables.

Further, dividends per share and the ratio of dividends to book value were negatively and significantly associated with share repurchases, with large magnitudes of the estimated coefficients. Thus, the models with intangible assets as an explanatory variable again showed that firms paying higher dividends made smaller share repurchases. Also, exogeneity hypothesis for dividends-to-book value ratio was close to rejection using the Chi-square statistic at 5 % level in the static model.

The rejection of the null indicated that there was some dependence in the unobserved factors affecting the firms' decisions to pay dividends and make share repurchases. For example, if firms made "unexpectedly" large share repurchases, then such decisions in turn can affect dividend payments.

Finally, in view of the literature on choice of measures for share repurchases (Banyi et al. 2008), the variable "purchase of common stock" was used as a proxy for share repurchases, and dynamic and static models were estimated. The results from modeling purchase of common stock were similar to those reported in Tables 38.5 and 38.7 using the more complex definition of share repurchases. However, greater numbers of observations were missing on purchase of common stock variable and these were reduced in the more complex definition because changes in Treasury stock were often positive so that the purchase of common stock variable was not utilized. Overall, the results indicated that these two measures for share repurchases yield similar results, perhaps because the models captured many salient features of the interrelationships.

38.10 Conclusion

This chapter presented a detailed analysis of the interrelationships between dividend payments and share repurchases by US industrial firms and modeled the proximate determinants of the payout methods (Bhargava 2010). The usefulness of analyzing longitudinal Compustat data on large number of firms and of appropriate econometric estimators was underscored. The empirical results provided several insights that can be summarized as follows. First, the magnitudes of dividends per share and dividend payments were broadly stable over the period of 1992–2007 using the data on the *same* firms. Moreover, simple dynamic models showed that dividends were higher in the period 2003–2007, following reduction in tax rates in 2003. These results were in some contrast with previous cross-sectional findings where entry and exit by firms in the Compustat database affected the numbers of firms paying dividends and magnitudes of dividend payments. For example, dividends were not seen to be "disappearing" (Fama and French 2001) and, apart from the effects of the tax cut in 2003, dividends were not "reappearing" (Julio and Ikenberry 2004). The simple dynamic models for share repurchases indicated steady increases in the period 1992–2007.

Second, the analytical and econometric frameworks led to comprehensive dynamic models for dividends per share and repurchases. The empirical results showed that variables such as firms' earnings, total assets, and investments were positively and significantly associated with dividends per share. Similar results were found for share repurchases, though firms' long-term debt was negatively and significantly associated with repurchases but was significantly associated with dividends only in a few models. Moreover, while dividends per share were negatively and significantly associated with share repurchases, coefficients of share repurchases were generally insignificant in the models for dividends. Exogeneity hypotheses for dividends per share were accepted in most models for repurchases, and exogeneity of share repurchases was also accepted in the models for dividends. These results suggest that firms' decisions to pay dividends at particular rates are taken prior to the decisions to repurchase shares that are more strongly influenced by the firms' current financial situations. For example, while firms can adjust repurchases given their debt levels, it may be more difficult to alter dividend payments. Thus, interpreting negative correlations between dividend forecasting errors and share repurchases as evidence of "substitution" between payment forms (Grullon and Michaely 2002) seems inconsistent with the elaborate two stage decision process documented by econometric analyses of the longitudinal Compustat database.

Third, the effects of firms' intangible assets on dividends per share were negative and significant, while these coefficients were positive and significant in the models for share repurchases. Such findings indicate that it is appropriate to model the two payout methods via separate models rather than combining them into a single variable. These results support the insights of Pearson (1897) and Neyman (1952) cautioning against employing ratios of variables in applied work. Such problems are exacerbated in longitudinal data analyses because stochastic properties of variables combined can evolve differently over time. It might have been useful to disaggregate intangible assets into components such as goodwill since they may differentially affect dividends and repurchase decisions. Because of missing observations on such variables in the Compustat database, further analyses were not pursued.

Finally, while econometric methods were useful for modeling the dynamic interactions between dividend payments and share repurchases, it is important in future research to investigate the role played by executive remuneration or compensation in making such allocations. Options to purchase company stock at some point in the future are regularly granted to the top executives and other employees of US firms and these may have potentially increased firms' share repurchases. If such decisions in turn reduce funds available for investment and research and development, then tax policies should discourage excessive repurchase activity. The ExecuComp (2008) database contains information on salary, bonus, and stock options granted to top executives of approximately 1,500 US firms from 1992. A recent analysis of the merged ExecuComp and Compustat databases in fact showed that beyond certain thresholds, share repurchases lower firms' expenditures on research and development and investments and should be discouraged (Bhargava 2012).

Appendix

The concentrated or "profile" log-likelihood functions of the model in Eq. 38.3 was computed by a FORTRAN program and was optimized using the numerical scheme

E04 JBF from Numerical Algorithm Group (1991). Note that the data were read firm-by-firm and several second moment matrices were created. Thus, there was no limit on the number of firms in the sample. The profile likelihood function, which depends only on the structural form parameters of interest, was computed using the second moment matrices. FORTRAN programs for the computations of the likelihood functions are available on request from the author.

The likelihood functions were separately computed for the case where the variance-covariance matrix of the errors was unrestricted and where the errors were decomposed in a random effect fashion, as in Eq. 38.2. Likelihood ratio tests were applied to discriminate between these two alternative formulations. Assuming that the number of firms (H) is large but the number of time observations is fixed, asymptotic standard errors of the parameters were obtained by approximating second derivatives of the functions at the maximum.

Further, assuming that n_1 and n_2 time-varying variables are exogenous and endogenous, respectively, as in Eq. 38.5, one can tackle the correlation between the errors u's and x_2's. It is reasonable to assume in short panels that a variable such as dividends per share in Eq. 38.5 may be correlated only with firm-specific random effects δ_i. Thus, the correlation pattern can be decomposed as:

$$x_{2ijt} = \lambda_j \delta_i + x^*_{2ijt} \tag{38.9}$$

where x^*_{2ijt} are uncorrelated with δ_i, and δ_i are randomly distributed variables with zero mean and finite variance as in Eq. 38.2. This correlation pattern was invoked by Bhargava and Sargan (1983) and has the advantage that deviations of x_{2ijt}'s from their time means:

$$x^+_{2ijt} = x_{2ijt} - x^-_{2ij} \quad (t = 2, \ldots, T; j = n_1 + 1, \ldots, n; i = 1, \ldots, H) \tag{38.10}$$

Where

$$x^-_{2ij} = \sum_{t=1}^{T} x_{2ijt}/T \quad (j = n_1 + 1, \ldots n; i = 1, \ldots, H) \tag{38.11}$$

can be used as additional $[(T-1)\,n_2]$ instrumental variables to facilitate identification and estimation of parameters. Because the T deviations from means (x^+_{2ijt}) in Eq. 38.10 will sum to zero for every firm, one time observation per variable needs to be omitted so that the (T−1) instrumental variables are linearly independent. The reduced form Eq. 38.6 for y_1 is modified in this case as:

$$y_{i1} = \sum_{j=1}^{m} z_{ij}\zeta_j + \sum_{j=1}^{n_1}\sum_{k=1}^{T} \upsilon_{jk}\, x_{1ijk} + \sum_{j=1}^{n_2}\sum_{k=2}^{T} \mu_{jk} x^+_{2ijk} + u_{i1}\,(i = 1, \ldots, H) \tag{38.12}$$

Further, likelihood ratio tests can be applied to test for exogeneity of time means of the n_2 endogenous variables (x_2) in Eq. 38.11 by testing if correlations between the T errors ($u_{i\ t}$) affecting the dependent variables and n_2 time means in Eq. 38.11 are zero. For example, given 8-time observations ($T = 8$), likelihood ratio statistic for testing zero correlation between errors affecting share repurchases in Eq. 38.4 ($n_2 = 1$) and time means of dividends per share is distributed for large H as a Chi-square variable with 8 degrees of freedom.

The identification of parameters is achieved via the n_1 time-varying exogenous variables in the model (e.g., Sargan 1958). For example, in the system representing ($T-1$) time periods in Eq. 38.7, only the exogenous explanatory variables in time period t explain the dependent variable in period t. The remaining ($T-1$) n_2 variables are excluded from the t th equation and are used in the set of instrumental variables for identifying the coefficients of endogenous variables. Sufficient conditions for identification, exploiting the time structure of longitudinal models, were developed in Bhargava and Sargan (1983). For example, each exogenous time-varying variable in the model for share repurchases in Eq. 38.4 effectively provides eight exogenous variables of which seven are excluded from the equations. The null hypothesis that dividends per share do not affect share repurchases can be tested using the estimated coefficient b_{12} of dividends per share in Eq. 38.4. Also, one can test if share repurchases affect dividends per share by including repurchases as a potentially endogenous variable in Eq. 38.3.

For static version of the models not containing lagged dependent variables, one can use stepwise estimation procedures with equality restrictions on coefficients across time periods (Bhargava 1991). Efficient instrumental variables estimators were used to estimate model parameters of Eq. 38.4, assuming the correlation patterns for $x_{2\ i\ j\ t}$ as in Eq. 38.9 and without restricting variance-covariance matrices of the errors, i.e., the T x T dispersion matrix was assumed to be symmetric and positive definite but not of the simple random effects form as in Eq. 38.2. This formulation has the advantage that the errors $v_{i\ t}$ in Eq. 38.2 may be serially correlated which was likely to be the case in static models since the lagged dependent variables were omitted. Exogeneity hypotheses can be tested in the stepwise estimation procedures via Chi-square tests that are asymptotically equivalent to likelihood ratio tests.

Finally, while one can use “fixed” effects estimators (with dummy variables for each firm) to circumvent certain endogeneity problems, increase in the number of parameters with sample size leads to the problem of “incidental parameters” (Neyman and Scott 1948). For example, coefficient of the lagged dependent variable cannot be consistently estimated due to the incidental parameters in fixed effects models unless the number of time observation is large. Moreover, from a modeling standpoint, the use of random effects models obviates the need for estimating the coefficients of large numbers of dummy variables for firms thereby enhancing the efficiency of estimates. Exogeneity hypotheses for variables such as dividends per share can be tested in the model for share repurchases. If the null were rejected, then the models were estimated under the appropriate assumption that dividends per share were correlated with errors affecting the models for share repurchases.

References

Anderson, T. W., & Hsiao, C. (1981). Estimation of dynamic models with error components. *Journal of the American Statistical Association, 76*, 598–606.

Banyi, M., Dyl, E., & Kahle, K. (2008). Errors in estimating share repurchases. *Journal of Corporate Finance, 14*, 460–474.

Bhargava, A. (1991). Identification and panel data models with endogenous regressors. *Reviews of Economic Studies, 58*, 129–140.

Bhargava, A. (1994). Modeling the health of Filipino children. *Journal of the Royal Statistical Society A, 157*, 417–432.

Bhargava, A. (2008). Randomized controlled experiments in health and social sciences. *Economics and Human Biology, 6*, 293–298.

Bhargava, A. (2010). An econometric analysis of the dividends and share repurchases by U.S. firms. *Journal of the Royal Statistical Society A, 173*, 631–656.

Bhargava, A. (2012, in press). Executive compensation, share repurchases and investment expenditures: Econometric evidence from U.S. firms. *Review of Quantitative Finance and Accounting*.

Bhargava, A., & Sargan, J. D. (1983). Estimating dynamic random effects models from panel data covering short time periods. *Econometrica, 51*, 1635–1660.

Brav, A., Graham, J. R., Harvey, C. R., & Michaely, R. (2005). Payout policy in the 21st century. *Journal of Financial Economics, 77*, 483–527.

Chen, S. -S., & Wang, Y. (2012, in press). Financial constraints and share repurchases. *Journal of Financial Economics*.

Compustat. (2008). *Standard & Poor's Compustat*. http://www.compustat.com/

Cox, D. (1958). *The planning of experiments*. New York: Wiley.

De Angelo, H., De Angelo, L., & Stultz, R. (2006). Dividend policy and the earned/contributed capital mix: A test of the life cycle theory. *Journal of Financial Economics, 81*, 227–254.

Duesenberry, J. (1949). *Income, saving, and the theory of consumer behavior*. Cambridge: Harvard University Press.

ExecuComp. (2008). *Standard & Poor's ExecuComp*. http://www.compustat.com/products.aspx?id=2147492873

Fama, E., & Babiak, H. (1968). Dividend policy: An empirical analysis. *Journal of the American Statistical Association, 63*, 1132–1161.

Fama, E., & French, K. (2001). Disappearing dividends: Changing firm characteristics or lower propensity to pay? *Journal of Financial Economics, 60*, 3–43.

Fenn, G. W., & Liang, N. (2001). Corporate payout policy and managerial stock incentives. *Journal of Financial Economics, 60*, 45–72.

Granger, C. (1992). Evaluating economic theory. *Journal of Econometrics, 51*, 3–5.

Grullon, G., & Michaely, R. (2002). Dividends, share repurchases and the substitution hypothesis. *Journal of Finance, 57*, 1649–1684.

Gu, F., & Wang, W. (2005). Intangible assets, information complexity, and analysts' earnings forecasts. *Journal of Business Finance and Accounting, 32*, 1673–1702.

Guay, W., & Harford, J. (2000). The cash-flow permanence and information content of dividend increases versus repurchases. *Journal of Financial Economics, 57*, 385–415.

Jagannathan, M., Stephens, C., & Weisbach, M. (2000). Financial flexibility and the choice between dividends and repurchases. *Journal of Financial Economics, 57*, 355–384.

John, K., & Williams, J. (1985). Dividends, dilution, and taxes: A signaling equilibrium. *Journal of Finance, 40*, 1053–1070.

Jones, R., & Wu, Y. (2010). Executive compensation, earnings management and share holder litigation. *Review of Quantitative Finance and Accounting, 35*, 1–20.

Julio, B., & Ikenberry, D. (2004). Reappearing dividends. *Journal of Applied Corporate Finance, 16*, 89–100.

Kahle, K. (2002). When a buyback isn't a buyback: Open market repurchases and employee options. *Journal of Financial Economics, 63*, 235–261.
Koopmans, T. C. (Ed.). (1950). *Statistical inference in dynamic economic models*. New York: Wiley.
Kronmal, R. (1993). Spurious correlation and the fallacy of the ratio standard revisited. *Journal of the Royal Statistical Society A, 156*, 379–392.
Lee, C., Wu, C., & Djarraya, M. (1987). A further empirical investigation of the dividend adjustment process. *Journal of Econometrics, 35*, 267–285.
Lev, B. (2001). *Intangibles*. Washington, DC: Brookings Institution Press.
Lintner, J. (1956). Distributions of incomes of corporations among dividends, retained earnings, and taxes. *American Economic Review, 46*, 97–113.
Miller, M., & Modigliani, F. (1961). Dividend policy, growth, and the valuation of shares. *Journal of Business, 34*, 411–433.
Neyman, J. (1952). *Lectures and conferences on mathematical statistics and probability* (2nd ed., pp. 143–154). Washington, DC: U.S. Department of Agriculture.
Neyman, J., & Scott, E. (1948). Consistent estimates based on partially consistent observations. *Econometrica, 16*, 1–32.
Numerical Algorithm Group. (1990). *Numerical Algorithm Group library. Mark 13*. Oxford: Oxford University.
Pearson, K. (1897). Mathematical contributions to the theory of evolution- on a form of spurious correlation which may arise when indices are used in the measurement of organs. *Proceedings of the Royal Society of London, 60*, 489–497.
Sargan, J. D. (1958). The estimation of economic relationships using instrumental variables. *Econometrica, 26*, 393–415.
Skinner, D. J. (2008). The evolving relation between earnings, dividends, and stock repurchases. *Journal of Financial Economics, 87*, 582–609.
Stata. (2008). *Stata version 10*. College Station: StataCorp.
Stephens, C., & Weisbach, M. (1998). Actual share reacquisition in open-market repurchase programs. *Journal of Finance, 53*, 313–333.
Tinbergen, J. (1939). The dynamics of share-price formation. *Review of Economics and Statistics, 21*, 153–160.
Von Eije, H., & Megginson, W. (2008). Dividends and share repurchases in the European Union. *Journal of Financial Economics, 89*, 347–374.

Term Structure Modeling and Forecasting Using the Nelson-Siegel Model 39

Jian Hua

Contents

Abstract

In this chapter, we illustrate some recent developments in the yield curve modeling by introducing a latent factor model called the dynamic Nelson-Siegel model. This model not only provides good in-sample fit, but also produces superior out-of-sample performance. Beyond Treasury yield curve, the model can also be useful for other assets such as corporate bond and volatility. Moreover, the model also suggests generalized duration components corresponding to the level, slope, and curvature risk factors.

The dynamic Nelson-Siegel model can be estimated via a one-step procedure, like the Kalman filter, which can also easily accommodate other variables of interests. Alternatively, we could estimate the model through a two-step process by fixing one parameter and estimating with ordinary least squares. The model is flexible and capable of replicating a variety of yield curve shapes: upward

J. Hua
Baruch College (CUNY), New York, NY, USA
e-mail: jian.hua@baruch.cuny.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_39,

sloping, downward sloping, humped, and inverted humped. Forecasting the yield curve is achieved through forecasting the factors and we can impose either a univariate autoregressive structure or a vector autoregressive structure on the factors.

Keywords
Term structure • Yield curve • Factor model • Nelson-Siegel curve • State-space model

39.1 Introduction

There have been major advances in theoretical term structure models, as well as their econometric estimation. Two popular approaches to term structure modeling are no-arbitrage models and equilibrium models. The no-arbitrage tradition focuses on eliminating arbitrage opportunities by perfectly fitting the term structure at a point in time, which is important for pricing derivatives; see Hull and White (1990) and Health et al. (1992), among others. The equilibrium tradition models the dynamics of the instantaneous rate typically through affine models so that yields at other maturities can be derived under various assumptions about the risk premium (e.g., Vasicek 1977; Cox et al. 1985; Duffie and Kan 1996).

For many finance questions, such as bond portfolio management, derivatives pricing, and risk management, it is crucial to both produce accurate estimates of the current term structure as well as forecast future interest rate dynamics. One class of models that has one potential satisfactory answer to these questions is that of the Nelson-Siegel class of models (see, Nelson and Siegel 1987). Here, we survey some recent developments of this model. This model not only provides good in-sample fit, but also produces superior out-of-sample performance. Moreover, the model also suggests generalized duration components corresponding to the level, slope, and curvature risk factors.

39.2 Modeling the Term Structure

Let $P_t(\tau)$ be the date t price of a zero-coupon riskless bond that pays \$1 in τ periods. Then,

$$P_t(\tau) = \exp(-\tau y_t(\tau)). \tag{39.1}$$

In practice, yield curves, discount curves, and forward curves are not observed. Instead, they must be estimated from observed bond prices. A popular approach is to estimate forward rates at the observed maturities, and then, construct unsmoothed yields by averaging appropriate estimated forward rates. These yields exactly price the included bonds (see Fama and Bliss 1987).

The original Nelson and Siegel (1987) framework is a convenient and parsimonious three-component exponential approximation. They work with the forward

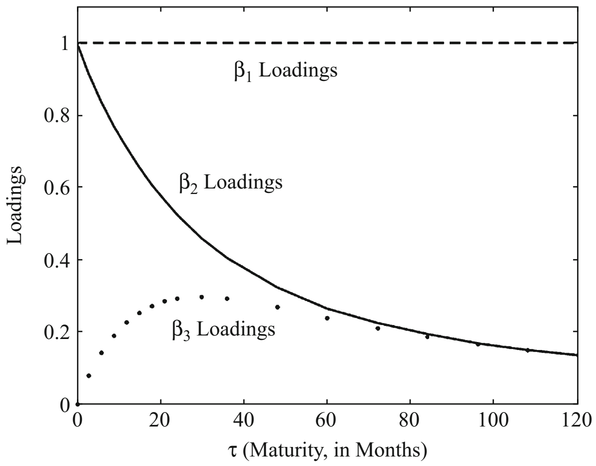


Fig. 39.1 Factor loadings. This figure plots the factor loadings when λ is prefixed at 0.0609

rate, which can be viewed as a constant plus a Laguerre function. Diebold and Li (2006) made it dynamic, and thus, the corresponding yield curve is

$$y_t(\tau) = \beta_{1t} + \beta_{2t}\left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau}\right) + \beta_{3t}\left(\frac{1 - e^{-\lambda\tau}}{\lambda\tau} - e^{-\lambda\tau}\right) + \varepsilon_t, \tag{39.2}$$

and call it DNS hereafter. β_1 changes all yields uniformly, and it can be called the level factor (L_t). β_2 loads the short rate more heavily, its loading decays to zero as maturity lengthens, and it can be called the slope factor (S_t). β_3 loads the medium term more heavily, its loading starts at zero and decays back to zero as maturity increases, and it can be called the curvature factor (C_t). In Fig. 39.1, we plot the factor loadings for a specific value of λ. λ determines the maturity at which the medium-term (or the curvature factor) loading achieves its maximum, which effectively controls the location of the hump of the curve. The DNS model is different from the factor analysis, in which one estimates both the unobserved factors and factor loadings. Here, the framework imposes structure on the factor loadings.

39.3 Fitting the Yield Curve

39.3.1 Data Construction

The standard way of measuring the term structure is by means of the spot rate curve (or equivalently the yield-to-maturity) on zero-coupon bonds.[1] The problem with

[1]Zero coupon bonds are chosen to limit the "coupon effect," which implies that two bonds that are identical in every respect except for bearing different coupon rates can have a different yield-to-maturity.

zero-coupon yields is that these are not usually directly observed. For example, the US Treasury Bills do not bear coupon, but they are only available for maturities of 1-year or less, and the longer maturity zero-coupon yields need to derived from coupon bearing Treasury Notes and Bonds. A popular approach to resolve such an issue is the bootstrapping procedure by Fama and Bliss (1987), which sequentially extracts forward rates from bond prices with successively longer maturities and then takes advantages of the interchangeable nature of the spot rate curve, discount curve, and forward rate curve.

39.3.2 Estimation Procedure

Because of the structure the model imposes, estimation can be achieved with high precision. There are two popular approaches: the one-step and two-step procedures. The one-step procedure can be achieved in two ways. One method is simply to estimate the model by nonlinear least squares for each month t. The other way is to transform the system into a state-space representation and estimate λ and the factors via a Kalman filter. Alternatively, for a two-step procedure, Diebold and Li (2006) advocate prefixing the λ, and estimating the factors via ordinary least squares (OLS). The appendix provides more discussion of these two approaches.

As we can see in Fig. 39.2, the model is flexible and is capable of replicating a variety of yield curve shapes: upward sloping, downward sloping, humped, and inverted humped. Figure 39.3 displays the autocorrelations of the estimated factors and residuals. The level factor displays high persistence and is, of course, positive. In contrast, the slope and curvature factors are less persistent and assume both positive and negative values.

39.4 Forecasting the Yield Curve

Forecasting the yield curve can be achieved through forecasting the factors. We can impose either a univariate autoregressive structure or a vector autoregressive structure on the factors.

The yield forecasts based on underlying univariate AR(1) factor specifications are

$$\hat{y}_{t+h/t}(\tau) = \beta_{1,t+h/t} + \beta_{2,t+h/t}\left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau}\right) + \beta_{3,t+h/t}\left(\frac{1-e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau}\right), \quad (39.3)$$

where β forecasts are made by the following:

$$\begin{aligned} \hat{\beta}_{1,t+h/t} &= \hat{c}_1 + \hat{\gamma}_1 \hat{\beta}_{1t} \\ \hat{\beta}_{2,t+h/t} &= \hat{c}_2 + \hat{\gamma}_2 \hat{\beta}_{2t} \\ \hat{\beta}_{3,t+h/t} &= \hat{c}_3 + \hat{\gamma}_3 \hat{\beta}_{3t} \end{aligned} \quad (39.4)$$

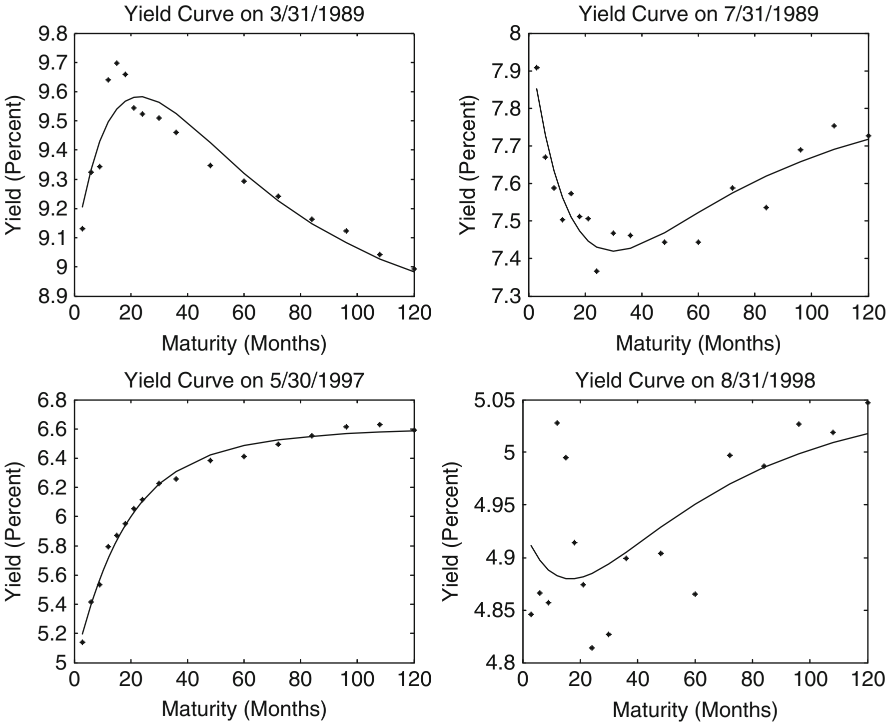


Fig. 39.2 Selected fitted yield curves. This figure presents fitted yield curve for selected dates, together with actual yields (Source: Diebold and Li 2006)

The yield forecasts based on an underlying multivariate autoregressive specification (Vector AR(1)) are

$$\hat{y}_{t+h/t}(\tau) = \beta_{1,\,t+h/t} + \beta_{2,\,t+h/t}\left(\frac{1 - e^{-\lambda_t \tau}}{\lambda_t \tau}\right) + \beta_{3,\,t+h/t}\left(\frac{1 - e^{-\lambda_t \tau}}{\lambda_t \tau} - e^{-\lambda_t \tau}\right), \quad (39.5)$$

where

$$\hat{\beta}_{t+h/t} = \hat{c} + \hat{\gamma}\hat{\beta}_t, \quad (39.6)$$

and $\beta = \{\beta_1, \beta_2, \beta_{3,}\}'$.

The random walk model,

$$\hat{y}_{t+h/t}(\tau) = y_t(\tau), \quad (39.7)$$

is a no-change forecast model, which is used as the benchmark.[2]

[2]Diebold and Li (2006) conduct a more extensive comparisons with competing models.

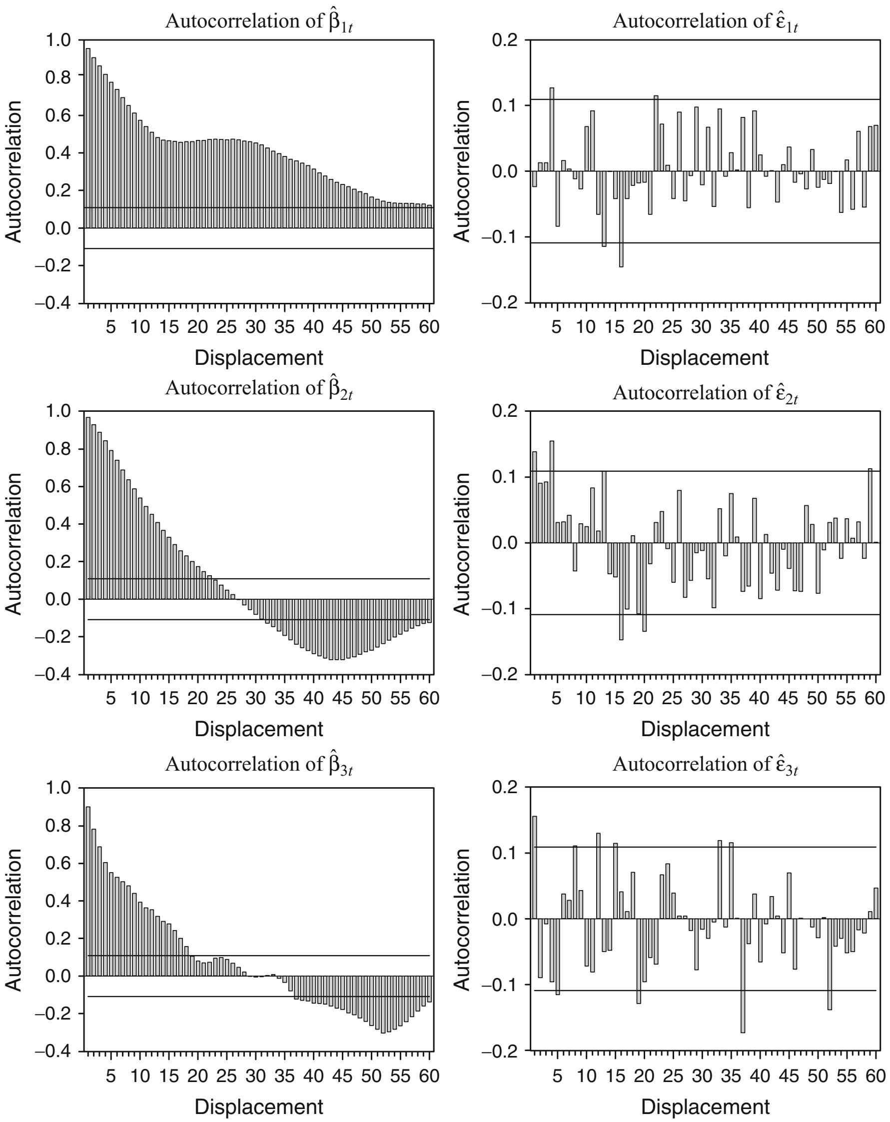


Fig. 39.3 Autocorrelation and residual autocorrelation. This figure presents sample autocorrelations of the level, slope, and curvature factors, as well as the sample autocorrelations of AR (1) models t to the three estimated factors, along with Barletts approximate 95 condence bands (Source: Diebold and Li 2006)

We define forecast errors at $t + h$ as $y_{t+h}(\tau) - \hat{y}_{t+h|t}(\tau)$. Table 39.1 reports the root mean-squared errors of the out-of-sample performance of the DNS model versus the random walk model. An error of 0.235 indicates a mean-squared error of 23.5 basis points in the yield prediction. For 1 month ahead, the random walk is

Table 39.1 Out-of-sample forecasting results

	3-month	1-year	2-year	3-year	5-year	10-year	30-year
1 *month a head*							
Random walk	0.235	**0.259**	**0.290**	0.300	**0.288**	**0.257**	**0.211**
Unrestricted VAR	**0.213**	0.261	0.307	0.302	0.310	0.314	0.419
Univariate AR	0.248	0.282	0.315	0.301	0.311	0.314	0.418
6 *months ahead*							
Random walk	0.910	0.904	0.938	**0.857**	**0.813**	0.655	0.570
Unrestricted VAR	0.909	1.050	1.078	0.999	0.949	0.796	0.717
Univariate AR	**0.694**	**0.890**	**0.908**	0.881	0.827	**0.651**	**0.486**
1 years ahead							
Random walk	1.552	1.505	1.343	1.190	1.065	0.858	0.622
DNS Unrestricted VAR	1.643	1.731	1.641	1.483	1.338	1.123	0.979
DNS Univariate AR	**1.383**	**1.445**	**1.338**	**1.182**	**1.041**	**0.819**	**0.615**

This table presents the results of out-of-sample forecasts using VAR and univariate AR specifications of the DNS factors. I estimate all models recursively from 1987:1 to the time the forecast is made, beginning in 1997:1 and extending through 2002:12. I define forecast errors at $t + h$ as $y_{t+h}(\tau) - \hat{y}_{t+h/t}(\tau)$ and report the root mean-squared errors versus random walk.

hard to beat, but matters improve dramatically. The DNS model with an AR (1) specification produces smaller errors, so it clearly has advantages (see Diebold and Li 2006 for more details). Remember here, we only illustrate the point forecast, and it can be extended easily to interval forecasting since they are useful for risk management.

39.5 Other Applications

Beyond the US Treasury yield curve, the Nelson-Siegel model has also shown success in fitting and forecasting other assets and global yield curves (see Krishnan et al. 2010; Diebold et al. 2008; Hua 2010). Since the model produces three factors that capture the information in the entire term structure, we can analyze the dynamic relationship among them. For example, Hua (2010) analyzes the dynamic interaction between the credit spread term structure and equity option implied volatility term structure.

Moreover, the DNS framework can be easily augmented with other variables of interest. We could also analyze how macroeconomic variables are linked with the interest rate term structure; see Auroba et al. (2006). This can be done easily in a state-space framework. Frequently, yield variation reacts to macrovariables, and macrovariables can also be impacted by the yields. Potentially, we can improve our forecasts of one by incorporating information from the other.

39.6 Generalized Duration Measure

Traditional interest rate risk management focuses on duration and duration management, which only considers parallel shifts of the yield curve. However, in practice, nonparallel shifts do exist and these are a significant source of risk. The DNS model presents a generalized duration component that corresponds to the level, slope, and curvature risk factors.

We can define a bond duration measure as follows. Let the cash flows from bond be $C_1, C_2, \ldots, C_I$, and define the associated maturities to be $\tau_1, \tau_2, \ldots, \tau_I$. We assume that the yield curve is linear in some arbitrary factors f_1, f_2, and f_3,

$$y_t(\tau) = B_1(\tau) f_{1t} + B_2(\tau) f_{2t} + B_3(\tau) f_{3t}. \tag{39.8}$$

This is consistent with the DNS setup. Since the price of the bond can be expressed as,

$$P = \sum_{i=1}^{I} C_i e^{-\tau_i y_t(\tau_i)}, \tag{39.9}$$

for an arbitrary change of the yield curve, the price change is

$$dP = \sum_{i=1}^{I} \frac{\partial P}{\partial y_t(\tau_i)} dy_i(\tau_i) = \sum_{i=1}^{I} \left[C_i e^{-\tau_i y_t(\tau_i)} (-\tau_i) \right] dy_t(\tau_i), \tag{39.10}$$

Where $y_t(\tau_i)$ are treated as independent variables. Rearranging terms, we can express the percentage change in bond price as a function of changes in the factors

$$-\frac{dP}{P} = \sum_{j=1}^{3} \left\{ \sum_{i=1}^{I} \left[\frac{1}{P} C_i e^{-\tau_i y_t(\tau_i)} \tau_i \right] B_j(\tau_i) \right\} df_{it} = \sum_{j=1}^{3} \left\{ \sum_{i=1}^{I} w_i \tau_i B_j(\tau_i) \right\} df_{it}, \tag{39.11}$$

where w_i is the weight associated with C_i.

Since we have decomposed the bond price changes into risk factor changes, the duration component associated with each risk factor is, for $j = 1,2,3$,

$$D_j = \sum_{i=1}^{I} w_i \tau_i B_j(\tau_i). \tag{39.12}$$

Moreover, based on the dynamic Nelson-Siegel model, the vector duration is

$$
\begin{aligned}
D_1 &= \sum_{i=1}^{I} w_i \tau_i \\
D_2 &= \sum_{i=1}^{I} w_i \frac{1 - e^{-\lambda \tau_i}}{\lambda} \\
D_3 &= \sum_{i=1}^{I} w_i \left(\frac{1 - e^{-\lambda \tau_i}}{\lambda} - \tau_i e^{-\lambda \tau_i} \right).
\end{aligned}
\tag{39.13}
$$

The vector duration measure has the following properties:

- D_1, D_2, and D_3 increase with maturity τ.
- D_1, D_2, and D_3 decrease with coupon rate.
- D_1, D_2, and D_3 decrease with yield-to-maturity.

Note that D_1 is exactly the tradition Macaulay duration.

Diebold et al. (2006) apply this vector duration measure as immunization tools in a practical bond portfolio management context. They report that hedging based on the vector duration outperforms hedging based on Macaulay duration in almost all samples and all holding periods. Moreover, it also outperforms polynomial vector duration during unusual market situations such as the monetary regime of the early 80s. Therefore, the vector duration measure seems to be an appealing risk management tool.

39.7 Summary

We have presented the DNS model as a capable and flexible model that can fit the term structure of interest rates in-sample and predict them out-of-sample. It can easily be extended to a different asset class or accommodate other variables into the specifications. Given the advantages of the dynamic Nelson-Siegel model in many dimensions, it is gaining in popularity. Recently, the Board of Governors of the Federal Reserve System has started publishing the daily estimated factors of the Treasury yield curve on their website.[3]

One potential drawback of the DNS model is that it does not rule out arbitrage opportunities. Recently, an arbitrage free version of the model has been developed; see Christensen et al. (2009, 2011).

Appendix 1: One-Step Estimation Method

We illustrate the one-step method to estimate the DNS model. If the dynamics of betas (factors) follow a vector autoregressive process of first order, the model

[3]The data can be found at http://www.federalreserve.gov/econresdata/researchdata.htm

immediately forms a state-space system since ARMA state vector dynamics of any order may be transformed into a state-space form. Thus, the transition equation is

$$\begin{pmatrix} \beta_{1,t} - \mu_{\beta 1} \\ \beta_{2,t} - \mu_{\beta 2} \\ \beta_{3,t} - \mu_{\beta 3} \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} \begin{pmatrix} \beta_{1,t-1} - \mu_{\beta 1} \\ \beta_{2,t-1} - \mu_{\beta 2} \\ \beta_{3,t-1} - \mu_{\beta 3} \end{pmatrix} + \begin{pmatrix} \eta_{1,t} \\ \eta_{2,t} \\ \eta_{3,t} \end{pmatrix}, \tag{39.14}$$

$t = 1, \ldots, T$. The measurement equation, which relates N yields to the three unobserved factors, is

$$\begin{pmatrix} y_t(\tau_1) \\ y_t(\tau_2) \\ \vdots \\ y_t(\tau_N) \end{pmatrix} = \begin{pmatrix} 1 & \dfrac{1 - e_1^{\tau}\lambda}{\tau_1\lambda} & \dfrac{1 - e^{-\tau_1\lambda}}{\tau_1\lambda} - e^{-\tau_1\lambda} \\ 1 & \dfrac{1 - e_2^{\tau}\lambda}{\tau_2\lambda} & \dfrac{1 - e^{-\tau_2\lambda}}{\tau_2\lambda} - e^{-\tau_2\lambda} \\ \vdots & \vdots & \vdots \\ 1 & \dfrac{1 - e_1^{\tau}\lambda}{\tau_N\lambda} & \dfrac{1 - e^{-\tau_N\lambda}}{\tau_N\lambda} - e^{-\tau_N\lambda} \end{pmatrix} \begin{pmatrix} \beta_{1,t} \\ \beta_{2,t} \\ \beta_{3,t} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \\ \vdots \\ \varepsilon_{\mathrm{N},t} \end{pmatrix}, \tag{39.15}$$

$t = 1, \ldots, T$. In matrix notation, we rewrite the state-space system as measurement equation,

$$y_t = \Gamma f_t + \varepsilon_t; \tag{39.16}$$

state equation,

$$(f_t - \mu) = A(f_{t-1} - \mu) + \eta_t. \tag{39.17}$$

For optimality of the Kalman filter, we assume that the white noise transaction and measurement errors be orthogonal to each other.

$$\begin{pmatrix} \eta_t \\ \varepsilon_t \end{pmatrix} \sim WN\left(\begin{bmatrix} 0 \\ 0 \end{bmatrix}, \begin{bmatrix} Q & 0 \\ 0 & H \end{bmatrix} \right), \tag{39.18}$$

$$E\left(f_0 \eta_t'\right) = 0, \tag{39.19}$$

$$E\left(f_0 \varepsilon_t'\right) = 0. \tag{39.20}$$

The state-space set up with the application of the Kalman filter delivers maximum-likelihood estimates and smoothed underlying factors, where all parameters are estimated simultaneously. Such representation also allows for heteroskedasticity, missing data, or heavy-tailed measurement errors. Moreover, other useful variables, such as macroeconomic variables, can be augmented into the state equation to understand the dynamic interactions between the yield curve and the macroeconomy.

Appendix 2: Two-Step Estimation Method

The estimation for the two-step procedure requires us to choose a λ value first. Once the λ is fixed, the values of the two regressors (factor loadings) can be computed, so ordinary least squares can be applied to estimate the betas (factors) at each period t. Doing so is not only simple and convenient, but also eliminates the potential for numerical optimization challenges. The question is: What is the appropriate value of λ. Recall that λ determines the maturity at which the loading achieves its maximum. For example, for Treasury data, 2 or 3 years are commonly considered medium term, so a simple average of the two is 30-month. The λ value that maximizes the loadings on the medium-term factor at exactly 30-month is 0.0609. For a different market, the medium maturity could be different, so the corresponding λ value could be different as well.

References

Aruoba, B. S., Diebold, F. X., & Rudebusch, G. D. (2006). The macroeconomy and yield curve: A dynamic latent factor approach. *Journal of Econometrics, 127*, 309–338.

Christensen, J. H. E., Diebold, F. X., & Rudebusch, G. D. (2009). An arbitrage-free generalized Nelson-Siegel term structure model. *Econometrics Journal, 12*, 33–64.

Christensen, J. H. E., Diebold, F. X., & Rudebusch, G. D. (2011). The affine arbitrage-free class of Nelson-Siegel term structure models. *Journal of Econometrics, 164*, 4–20.

Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica, 53*, 385–407.

Diebold, F. X., Ji, L., & Li, C. (2006). A three-factor yield curve model: Non-affine structure, systematic risk sources, and generalized duration. In L. R. Klein (Ed.), *Macroeconomics, finance and econometrics: Essays in memory of Albert Ando* (pp. 240–274). Cheltenham: Edward Elgar.

Diebold, F. X., & Li, C. (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics, 127*, 337–364.

Diebold, F. X., Li, C., & Yue, V. (2008). Global yield curve dynamics and interactions: A generalized Nelson-Siegel approach. *Journal of Econometrics, 146*, 351–363.

Duffie, D., & Kan, R. (1996). A yield-factor model of interest rates. *Mathematical Finance, 6*, 379–406.

Fama, E., & Bliss, R. (1987). The information in long-maturity forward rates. *American Economic Review, 77*, 680–692.

Heath, D., Jarrow, R., & Morton, A. (1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica, 60*, 77–105.

Hua, J. (2010). *Option implied volatilities and corporate bond yields: A dynamic factor approach* (Working Paper). Baruch College, City University of New York.

Hull, J., & White, A. (1990). Pricing interest-rate-derivative securities. *Review of Financial Studies, 3*, 573–592.

Krishnan, C. N. V., Ritchken, P., & Thomson, J. (2010). Predicting credit spreads. *Journal of Financial Intermediation, 19*(4), 529–563.

Nelson, C. R., & Siegel, A. F. (1987). Parsimonious modeling of yield curve. *Journal of Business, 60*, 473–489.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics, 5*, 177–188.

The Intertemporal Relation Between Expected Return and Risk on Currency

40

Turan G. Bali and Kamil Yilmaz

Contents

Abstract

The literature has so far focused on the risk-return trade-off in equity markets and ignored alternative risky assets. This paper examines the presence and significance of an intertemporal relation between expected return and risk in the foreign exchange market. The paper provides new evidence on the intertemporal capital asset pricing model by using high-frequency intraday data on currency and by presenting significant time variation in the risk aversion parameter. Five-minute returns on the spot exchange rates of the US dollar vis-à-vis six major currencies (the euro, Japanese yen, British pound sterling, Swiss franc, Australian dollar, and Canadian dollar) are used to test the existence and significance of a daily risk-return trade-off in the FX market based on the

T.G. Bali (✉)
McDonough School of Business, Georgetown University, Washington, DC, USA
e-mail: tgb27@georgetown.edu

K. Yilmaz
College of Administrative Sciences and Economics, Koc University, Istanbul, Turkey
e-mail: kyilmaz@ku.edu.tr

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_40,

GARCH, realized, and range volatility estimators. The results indicate a positive but statistically weak relation between risk and return on currency.

Our empirical analysis relies on the maximum likelihood estimation of the GARCH-in-mean models as described in Appendix 1. We also use the seemingly unrelated (SUR) regressions and panel data estimation to investigate the significance of a time-series relation between expected return and risk on currency as described in Appendix 2.

Keywords
GARCH • GARCH-in-mean • Seemingly unrelated regressions (SUR) • Panel data estimation • Foreign exchange market • ICAPM • High-frequency data • Time-varying risk aversion • High-frequency data • Daily realized volatility

40.1 Introduction

Merton's (1973) intertemporal capital asset pricing model (ICAPM) indicates that the conditional expected excess return on a risky market portfolio is a linear function of its conditional variance plus a hedging component that captures the investor's motive to hedge for future investment opportunities. Merton (1980) shows that the hedging demand component becomes negligible under certain conditions and the equilibrium relation between risk and return is defined as

$$E_t(R_{t+1}) = \beta \cdot E_t\left(\sigma^2_{t+1}\right) \tag{40.1}$$

where $E_t(R_{t+1})$ and $E_t(\sigma^2_{t+1})$ are, respectively, the conditional mean and variance of excess returns on a risky market portfolio and $\beta > 0$ is the risk aversion parameter of market investors. Equation 40.1 establishes the dynamic relation that investors require a larger risk premium at times when the market is riskier.

Many studies investigate the significance of an intertemporal relation between expected return and risk in the aggregate stock market. However, the existing literature has not yet reached an agreement on the existence of a positive risk-return trade-off for stock market indices.[1] Due to the fact that the conditional mean and volatility of the market portfolio are not observable, different approaches, different data sets, and different sample periods used by previous studies in estimating the conditional mean and variance are largely responsible for the contradictory empirical evidence.

The prediction of Merton (1973, 1980) that expected returns should be related to conditional risk applies not only to the stock market portfolio but also to any risky portfolio. However, earlier studies have so far focused on the risk-return trade-off in

[1]See French et al. (1987), Campbell (1987), Nelson (1991), Campbell and Hentschel (1992), Chan et al. (1992), Glosten et al. (1993), Scruggs (1998), Harvey (2001), Goyal and Santa-Clara (2003), Brandt and Kang (2004), Ghysels et al. (2005), Bali and Peng (2006), Christoffersen and Diebold (2006), Guo and Whitelaw (2006), Lundblad (2007), and Bali (2008).

equity markets and ignored other risky financial assets. Although there are a few studies testing the significance of a time-series relation between risk and return in international equity markets, the focus is generally on the US stock market. It is also important to note that earlier studies assume a constant risk-return trade-off and ignore time variation in the risk aversion parameter β.[2] This paper examines the intertemporal relation between expected return and risk in currency markets. The paper not only investigates ICAPM in the foreign exchange market but examines the significance of time-varying risk aversion as well.

The foreign exchange market includes the trading of one currency against another between large banks, central banks, currency speculators, multinational corporations, governments, and other financial markets and institutions. The FX market is an interbank or inter-dealer network first established in 1971 when many of the world's major currencies moved towards floating exchange rates. It is considered an over-the-counter (OTC) market, meaning that transactions are conducted between two counterparties that agree to trade via telephone or electronic network. Because foreign exchange is an OTC market where brokers/dealers negotiate directly with one another, there is no central exchange or clearing house.[3]

The FX market has grown rapidly since the early 1990s. According to the triennial central bank surveys conducted by the Bank for International Settlements (BIS), the April 2007 data show an unprecedented rise in activity in traditional foreign exchange markets compared to 2004. As shown in Table 40.1, average daily turnover rose to US $3.1 trillion in April 2007, an increase of 69 % (compared to April 2004) at current exchange rates and 63 % at constant exchange rates.[4] Since April 2001, average daily turnover in foreign exchange markets worldwide (adjusted for cross-border and local double-counting and evaluated at April 2007 exchange rates) increased by 58 % and 69 % between two consecutive triennial surveys. Comparing the average daily turnovers of US $500 billion in 1988 and US $3.1 trillion in 2007 indicates that trading volume in FX markets increased by more than five times over the past two decades.

The FX market has become the world's largest financial market, and it is not uncommon to see over US $3 trillion traded each day. By contrast, the New York Stock Exchange (NYSE) – the world's largest equity market with daily trading volumes in the US $60–80 billion dollar range – is positively dwarfed when compared to the FX market. Daily turnover in FX markets is now more than ten times the size of the combined daily turnover on all the world's equity markets.

---

[2]A few exceptions are Chou et al. (1992), Harvey (2001), and Lettau and Ludvigson (2010).

[3]As FX trading has evolved, several locations have emerged as market leaders. Currently, London contributes the greatest share of transactions with over 32 % of the total trades. Other trading centers – listed in order of volume – are New York, Tokyo, Zurich, Frankfurt, Hong Kong, Paris, and Sydney. Because these trading centers cover most of the major time zones, FX trading is a true 24-h market that operates 5 days a week.

[4]In addition to "traditional" turnover of US $3.1 trillion in global foreign exchange market, US $2.1 trillion was traded in currency derivatives.

Table 40.1 Reported foreign exchange market turnover by currency pair[a] (Daily averages in April, in billions of US dollars and percent)

	2001		2004		2007	
	Amount	% share	Amount	% share	Amount	% share
US dollar/euro	354	30	503	28	840	27
US dollar/yen	231	20	298	17	397	13
US dollar/British pound	125	11	248	14	361	12
US dollar/Australian dollar	47	4	98	5	175	6
US dollar/Swiss franc	57	5	78	4	143	5
US dollar/Canadian dollar	50	4	71	4	115	4
US dollar/other	195	17	295	16	628	21
Euro/yen	30	3	51	3	70	2
Euro/sterling	24	2	43	2	64	2
Euro/Swiss franc	12	1	26	1	54	2
Euro/other	21	2	39	2	112	4
Other currency pairs	26	2	42	2	122	4
All currency pairs	1,173	100	1,794	100	3,081	100

[a]Adjusted for local and cross-border double-counting

Even when combining the US bond and equity markets, total daily volumes still do not come close to the values traded on the currency market.

The FX market is unique because of its trading volumes; the extreme liquidity of the market; the large number of, and variety of, traders in the market; its geographical dispersion; its long trading hours (24 h a day except on weekends); the variety of factors that affect exchange rates; the low margins of profit compared with other markets of fixed income (but profits can be high due to very large trading volumes); and the use of leverage.

Earlier studies have so far focused on the US stock market when investigating the ICAPM. However, with an average daily trading volume of US $3 trillion per day, Forex is far and away the most enormous financial market in the world, dwarfing the trading volumes of other markets. We contribute to the existing literature by examining for the first time the significance of an intertemporal relation between expected return and risk on currency. We also test whether aggregate risk aversion in the FX market changes through time.

We utilize 5-min returns on the spot exchange rates of the US dollar vis-à-vis six major currencies (the euro, Japanese yen, British pound sterling, Swiss franc, Australian dollar, and Canadian dollar) to construct the daily returns, realized volatility, and range volatility estimators. Then, using the intraday data-based daily returns as well as the GARCH, realized, and range-based volatility measures, we test for the presence and significance of a risk-return trade-off in the FX market. By sampling the return process more frequently, we improve the accuracy of the conditional volatility estimate and measure the risk-return relationship at the daily level. When we assume a constant risk-return trade-off in currency markets, we find a positive but statistically weak relation between expected return and risk on currency.

We estimate the dependence of expected returns on the lagged realized variance over time using rolling regressions. This also allows us to check whether our results are driven by a particular sample period. Two different rolling regression approaches provide strong evidence on the time variation of risk aversion parameters for all currencies considered in the paper. However, the direction of a relationship between expected return and risk is not clear for the entire FX market.

The paper is organized as follows. Section 40.2 provides the descriptive statistics for the daily and 5-min returns on exchange rates as well as the daily realized and range-based volatility measures. Section 40.3 explains the estimation methodology. Section 40.4 presents the empirical results on a constant risk-return trade-off in the FX market. Section 40.5 examines the significance of time-varying risk aversion. Section 40.6 investigates whether the covariances of individual exchange rates with the FX market are priced in currency market. Section 40.7 concludes the paper.

40.2 Data

To test the significance of a risk-return trade-off in currency markets, we use daily returns on the spot exchange rates of the US dollar vis-à-vis six major currencies: the euro (EUR), Japanese yen (JPY), British pound sterling (GBP), Swiss franc (CHF), Australian dollar (AUD), and Canadian dollar (CAD). According to the BIS (2007) study, on the spot market the most heavily traded currency pairs were EUR/USD (27 %), JPY/USD (13 %), GBP/USD (12 %), AUD/USD (6 %), CHF/USD (5 %), and CAD/USD (4 %). As reported in Table 40.2, the US dollar has been the dominant currency in both the spot and the forward and the swap transactions. Specifically, the US currency was involved in 88.7 % of transactions, followed by the euro (37.2 %), the Japanese yen (20.3 %), the pound sterling (16.9 %), the Swiss franc (6.1 %), Australian dollar (5.5 %), and Canadian dollar (4.2 %). The sum of the six major currencies (EUR, JPY, GBP, CHF, AUD, CAD) accounts for a market share approximately equal to that of the US dollar (90.2 %).[5]

The raw 5-min data on six exchange rates (EUR/USD, JPY/USD, GBP/USD, CHF/USD, AUD/USD, and CAD/USD) are obtained from Olsen and Associates. The full sample covers 2,282 days, from January 1, 2002 to March 31, 2008. Following Bollerslev and Domowitz (1993) and Andersen et al. (2001b), we define the day as starting at 21:05 pm on one night and ending at 21:00 pm the next night. The total number of 5-min observations for each exchange rate is therefore equal to $2{,}282 \times 288 = 657{,}216$. However, we are not able to use all of these observations

[5]Note that volume percentages should add up to 200 %; 100 % for all the sellers and 100 % for all the buyers. As shown in Table 40.2, the market shares of seven major currencies add up to 180 %. The remaining 20 % of the total (200 %) market turnover has been accounted by other currencies from Europe and from other parts of the world.

Table 40.2 Most traded currencies: currency distribution of reported FX market turnover

Currency	Symbol	% daily share
US dollar	USD ($)	88.7 %
Euro	EUR (€)	37.2 %
Japanese yen	JPY (¥)	20.3 %
British pound sterling	GBP (£)	16.9 %
Swiss franc	CHF (Fr)	6.1 %
Australian dollar	AUD ($)	5.5 %
Canadian dollar	CAD ($)	4.2 %
Swedish krona	SEK (kr)	2.3 %
Hong Kong dollar	HKD ($)	1.9 %
Norwegian krone	NOK (kr)	1.4 %
Other		15.5 %
Total		200 %

because the trading activity in FX markets slows down substantially during the weekends and the major US official holidays. Following Andersen et al. (2001b), along with the weekends, we removed the following holidays from our sample: Christmas (December 24–26), New Year's (December 31–January 2), July 4th, Good Friday, Easter Monday, Memorial Day, Labor Day, Thanksgiving Day, and the day after. In addition to official holidays and weekends, we removed 3 days (March 4, 2002, April 14, 2003, and January 30, 2004) from our sample as these days contained the longest zero or constant 5-min return sequences that might contaminate the daily return and variance estimates. As a result, we end up with a total of 1,556 daily observations.

Panel A of Table 40.3 presents the mean, median, maximum, minimum, standard deviation, skewness, kurtosis, and autoregressive of order one, AR(1), statistics for daily returns on the six exchange rates. The standard errors of the skewness and kurtosis estimates provide evidence that the empirical distributions of returns on exchange rates are generally symmetric and fat tailed. More specifically, the skewness measures are statistically insignificant for all currencies, except for the Japanese yen. The kurtosis measures are statistically significant without any exception. The Jarque-Bera, JB $= n[(S^2/6) + (K–3)^2/24]$, is a formal statistic with the Chi-square distribution for testing whether the returns are normally distributed, where n denotes the number of observations, S is skewness, and K is kurtosis. The JB statistics indicate significant departures from normality for the empirical return distributions of six exchange rates. As expected, daily returns on exchange rates are not highly persistent, as shown by the negative AR (1) coefficients which are less than 0.10 in absolute value. Although the economic significance of the AR(1) coefficients is low, they are statistically significant at the 5 % or 1 % level for all currencies, except for the British pound and Australian dollar.

The daily intertemporal relation between expected return and risk on currency is tested using the daily realized variance of returns on exchange rates. In very early work, the daily realized variance of asset returns is measured

Table 40.3 Descriptive statistics

Panel A. Descriptive statistics for the daily returns on exchange rates

	EUR	JPY	GBP	CHF	AUD	CAD
Mean	−0.00033	−0.000076	−0.00020	−0.00029	−0.00034	−0.00025
Median	−0.00024	0.00004	−0.00029	−0.00012	−0.00064	−0.00032
Maximum	0.01963	0.02515	0.01756	0.02126	0.03496	0.02563
Minimum	−0.01837	−0.02686	−0.02048	−0.02223	−0.02259	−0.01734
Std. dev.	0.00563	0.00580	0.00504	0.00635	0.00681	0.00520
Skewness	0.0783	−0.1522**	0.0776	−0.0306	0.5377**	0.1091
Kurtosis	3.6142**	4.2687**	3.4508**	3.5184**	4.5571**	3.8636**
Jarque-Bera	26.05**	110.36**	14.74**	17.67**	232.18**	51.44**
AR(1)	−0.0643**	−0.0504*	−0.0117	−0.0788**	0.0164	−0.0642**

Panel B. Descriptive statistics for the daily realized variance measures

	EUR	JPY	GBP	CHF	AUD	CAD
Mean	3.47×10^{-5}	4.07×10^{-5}	2.77×10^{-5}	4.31×10^{-5}	6.00×10^{-5}	3.60×10^{-5}
Median	3.13×10^{-5}	3.39×10^{-5}	2.50×10^{-5}	3.77×10^{-5}	4.86×10^{-5}	3.09×10^{-5}
Maximum	0.000253	0.000505	0.000156	0.000367	0.000904	0.00022
Minimum	6.51×10^{-6}	4.05×10^{-6}	5.42×10^{-6}	7.54×10^{-6}	1.48×10^{-5}	5.12×10^{-6}
Std. dev.	1.93×10^{-5}	3.12×10^{-5}	1.37×10^{-5}	2.48×10^{-5}	4.39×10^{-5}	2.30×10^{-5}
Skewness	2.5435**	6.2585**	2.2476**	3.1791**	6.8653**	2.5195**
Kurtosis	19.957**	71.480**	13.655**	28.010**	104.839**	13.591**
Jarque-Bera	20319.0**	314192.0**	8670.8**	43174.2**	684614.4**	8918.8**
AR(1)	0.50**	0.55**	0.51**	0.49**	0.62**	0.64**

Panel C. Descriptive statistics for the daily range variance measures

	EUR	JPY	GBP	CHF	AUD	CAD
Mean	2.76×10^{-5}	3.15×10^{-5}	2.30×10^{-5}	3.56×10^{-5}	4.20×10^{-5}	2.63×10^{-5}
Median	1.91×10^{-5}	2.14×10^{-5}	1.61×10^{-5}	2.52×10^{-5}	2.68×10^{-5}	1.83×10^{-5}

(*continued*)

Table 40.3 (continued)

Panel C. Descriptive statistics for the daily range variance measures						
	EUR	JPY	GBP	CHF	AUD	CAD
Maximum	0.000204	0.000558	0.000189	0.000364	0.000777	0.000297
Minimum	1.37×10^{-6}	3.58×10^{-7}	1.46×10^{-6}	1.55×10^{-6}	2.58×10^{-6}	1.66×10^{-6}
Std. dev.	2.62×10^{-5}	3.69×10^{-5}	2.09×10^{-5}	3.41×10^{-5}	5.20×10^{-5}	2.70×10^{-5}
Skewness	2.2886**	5.2166**	2.3097**	2.5960**	5.4669**	3.3357**
Kurtosis	9.953**	50.502**	11.119**	13.832**	51.593**	22.051**
Jarque-Bera	4492.6**	153348.5**	5657.0**	9354.9**	160839.7**	26417.1**
AR(1)	0.09**	0.25**	0.19**	0.12*	0.34**	0.28**

This table reports statistics for daily returns on US dollar exchange rates (Panel A), daily realized variance (Panel B), and daily range variance (Panel C) for six major currencies: the euro (EUR), Japanese yen (JPY), British pound sterling (GBP), Swiss franc (CHF), Australian dollar (AUD), and Canadian dollar (CAD). Mean, median, maximum, minimum, standard deviation, skewness, kurtosis, and AR(1) statistics are reported for each currency. Standard errors of skewness and kurtosis estimates, computed under the null hypothesis that the returns are normally distributed, are $\sqrt{6/n}$ and $\sqrt{24/n}$, respectively. Jarque-Bera, JB $= n[(S^2/6) + (K-3)^2/24]$, is a formal statistic for testing whether the returns are normally distributed, where n denotes the number of observations, S is skewness, and K is kurtosis. The JB statistic distributed as the Chi-square with 2° of freedom measures the difference of the skewness and kurtosis of the series with those from the normal distribution. The sample period is from January 3, 2002 to March 31, 2008, yielding a total of 1,556 daily observations

*, ** denote statistical significance at the 5 % and 1 % level, respectively

by the squared daily returns, where the asset return is defined as the natural logarithm of the ratio of consecutive daily closing prices. A series of papers by Andersen et al. (2001a, b, 2003, 2004) indicate that these traditional measures are poor estimators of day-by-day movements in volatility, as the idiosyncratic component of daily returns is large. They demonstrate that the realized volatility measures based on intraday data provide a dramatic reduction in noise and a radical improvement in temporal stability relative to realized volatility measures based on daily returns. Andersen et al. (2003) show formally that the concept of realized variance is, according to the theory of quadratic variation and under suitable conditions, an asymptotically unbiased estimator of the integrated variance, and thus it is a canonical and natural measure of daily return volatility.

Following the recent literature on integrated volatility, we use the high-frequency intraday data to construct the daily realized variance of exchange rates. To set forth notation, let P_t denote the time t ($t \geq 0$) exchange rate with the unit interval t corresponding to 1 day. The discretely observed time-series process of logarithmic exchange rate returns with q observations per day, or a return horizon of $1/q$, is then defined by

$$R_{(q),t} = \ln P_t - \ln P_{t-1/q} \tag{40.2}$$

where $t = 1/q, 2/q, \ldots$. We calculate the daily realized variance of exchange rates using the intraday high-frequency (5-min) return data as

$$VAR_t^{\text{realized}} = \sum_{i=0}^{q_i - 1} R^2_{(q),t-i/q} \tag{40.3}$$

where $q_{i,t}$ is the number of 5-min intervals on day t and $R_{i,t}$ is the logarithmic exchange rate return in 5-min interval i on date t.

On a regular trading day, there are 288 5-min intervals. The exchange rate of the most recent record in a given 5-min interval is taken to be the exchange rate of that interval. A 5-min return is then constructed using the logarithmic exchange rate difference for a 5-min interval. With 1,556 days in our full sample, we end up with using a total of $1{,}556 \times 288 = 448{,}128$ 5-min return observations to calculate daily return and variance estimates.

Panel B of Table 40.3 presents the summary statistics of the daily realized-variances of exchange rate returns. The average daily realized variance is 6×10^{-5} for AUD/USD, 4.31×10^{-5} for CHF/USD, 4.07×10^{-5} for JPY/USD, 3.60×10^{-5} for CAD/USD, 3.47×10^{-5} for EUR/USD, and 2.77×10^{-5} for GBP/USD. These measures correspond to an annualized volatility of 12.30 % for AUD/USD, 10.42 % for CHF/USD, 10.13 % for JPY/USD, 9.52 % for CAD/USD, 9.35 % for EUR/USD, and 8.35 % for GBP/USD. A notable point in Panel B is that the daily realized variances are highly persistent, as shown by the AR(1) coefficients which are in the range of 0.49–0.64. Consistent with Andersen et al. (2001a, b), the distributions of realized variances are skewed to the right and have much thicker tails than the corresponding normal distribution.

Market microstructure noises in transaction data such as the bid-ask bounce may influence our risk measures based on the realized volatility and GARCH volatility forecasts, even though the data we use contain very liquid financial time series and thus are least subject to biases created by market microstructure effects. An alternative volatility measure that utilizes information contained in the high-frequency intraday data is Parkinson's (1980) range-based estimator of the daily integrated variance:

$$VAR_t^{\text{range}} = 0.361 \left[\ln\left(P_t^{\max}\right) - \ln\left(P_t^{\min}\right)\right]^2 \tag{40.4}$$

where $P_t^{\max}$ and $P_t^{\min}$ are the maximum and minimum values of the exchange rate on day t. Alizadeh et al. (2002) and Brandt and Diebold (2006) show that the range-based volatility estimator is highly efficient, approximately Gaussian, and robust to certain types of microstructure noise such as bid-ask bounce. In addition, range data are available for many assets over a long sample period.

Panel C of Table 40.3 presents the summary statistics of the daily range variances of exchange rate returns. The average daily range variance is 4.20×10^{-5} for AUD/USD, 3.56×10^{-5} for CHF/USD, 3.15×10^{-5} for JPY/USD, 2.63×10^{-5} for CAD/USD, 2.76×10^{-5} for EUR/USD, and 2.0×10^{-5} for GBP/USD. These measures correspond to an annualized volatility of 10.29 % for AUD/USD, 9.47 % for CHF/USD, 8.91 % for JPY/USD, 8.14 % for CAD/USD, 8.34 % for EUR/USD, and 7.61 % for GBP/USD. These results indicate that the daily realized volatility estimates are somewhat higher than the daily range volatilities. Another notable point in Panel C is that the daily range variances are less persistent than the daily realized variances. Specifically, the AR(1) coefficients are in the range of 0.09–0.34 for the daily range variances. Similar to our findings for the daily realized variances, the distributions of range variances are skewed to the right and have much thicker tails than the corresponding normal distribution.

40.3 Estimation Methodology

The following GARCH-in-mean process is used with the conditional normal density to model the intertemporal relation between expected return and risk on currency:

$$R_{t+1} \equiv \alpha + \beta \cdot \sigma_{t+1|t}^2 + \varepsilon_{t+1} \tag{40.5}$$

$$\varepsilon_{t+1} = z_{t+1} \cdot \sigma_{t+1|t}, z_{t+1} \sim N(0,1), E(\varepsilon_{t+1}) = 0 \tag{40.6}$$

$$E\left(\varepsilon_{t+1}^2 | \Omega_t\right) = \sigma_{t+1|t}^2 = \gamma_0 + \gamma_1 \varepsilon_t^2 + \gamma_2 \sigma_t^2, \tag{40.7}$$

$$f\left(R_{t+1}; \mu_{t+1|t}, \sigma_{t+1|t}\right) = \frac{1}{\sqrt{2\pi\sigma_{t+1|t}^2}} \exp\left[-\frac{1}{2}\left(\frac{R_{t+1} - \mu_{t+1|t}}{\sigma_{t+1|t}}\right)^2\right] \tag{40.8}$$

where R_{t+1} is the daily return on exchange rates for period $t+1$, $\mu_{t+1|t} \equiv \alpha + \beta \cdot \sigma^2_{t+1|t}$ is the conditional mean for period $t + 1$ based on the information set up to time t, $\varepsilon_{t+1} = z_{t+1} \cdot \sigma_{t+1|t}$ is the error term with $E(\varepsilon_{t+1}) = 0$, $\sigma_{t+1|t}$, is the conditional standard deviation of daily returns on currency, and $z_{t+1} \sim N(0,1)$ is a random variable drawn from the standard normal density and can be viewed as information shocks in the FX market. $\sigma^2_{t+1|t}$ is the conditional variance of daily returns based on the information set up to time t denoted by Ω_t. The conditional variance, $\sigma^2_{t+1|t}$, follows a GARCH(1,1) process as defined by Bollerslev (1986) to be a function of the last period's unexpected news (or information shocks), z_t, and the last period's variance, $\sigma_t{}^2$. $f(R_{t+1};\mu_{t+1|t},\sigma_{t+1|t})$ is the conditional normal density function of R_{t+1} with the conditional mean of $\mu_{t+1|t}$ and conditional variance of $\sigma^2_{t+1|t}$. Our focus is to examine the magnitude and statistical significance of the risk aversion parameter β in Eq. 40.5.

Campbell (1987) and Scruggs (1998) point out that the approximate relationship in Eq. 40.1 may be misspecified if the hedging term in ICAPM is important. To make sure that our results from estimating Eq. 40.5 are not due to model misspecification, we added to the specifications a set of control variables that have been used in the literature to capture the state variables that determine changes in the investment opportunity set. Several studies find that macroeconomic variables associated with business cycle fluctuations can predict the stock market.[6] The commonly chosen variables include Treasury bill rates, federal funds rate, default spread, term spread, and dividend-price ratios. We study how variations in the fed funds rate, default spread, and term spread affect the intertemporal risk-return relation.[7] Earlier studies also control for the lagged return in the conditional mean specification.

We obtain daily data on the federal funds rate, 3-month Treasury bill, 10-year Treasury bond yields, and BAA-rated and AAA-rated corporate bond yields from the H.15 database of the Federal Reserve Board. The federal funds (*FED*) rate is the interest rate at which a depository institution lends immediately available funds (balances at the Federal Reserve) to another depository institution overnight. It is a closely watched barometer of the tightness of credit market conditions in the banking system and the stance of monetary policy. In addition to the fed funds rate, we use the term and default spreads as control variables. The term spread (*TERM*) is calculated as the difference between the yields on the 10-year Treasury bond and the 3-month Treasury bill. The default spread (*DEF*) is computed as the difference between the yields on the BAA-rated and AAA-rated corporate bonds. We test the significance of the risk aversion parameter, β, after controlling for macroeconomic variables and lagged return:

[6]See Keim and Stambaugh (1986), Chen et al. (1986), Campbell and Shiller (1988), Fama and French (1988, 1989), Campbell (1987, 1991), Ghysels et al (2005), and Guo and Whitelaw (2006).

[7]We could not include the aggregate dividend yield (or the dividend-price ratio) because the data on dividends are available only at the monthly frequency while our empirical analyses are based on the daily data.

$$R_{t+1} \equiv \alpha + \beta \cdot \sigma^2_{t+1|t} + \lambda_1 \cdot FED_t + \lambda_2 \cdot DEF_t + \lambda_3 \cdot TERM_t + \lambda_4 \cdot R_t + \varepsilon_{t+1}$$

$$R_{t+1} \equiv \alpha + \beta \cdot \sigma^2_{t+1|t} + \lambda_1 \cdot FED_t + \lambda_2 \cdot DEF + \lambda_3 \cdot TERM_t + \lambda_4 \cdot R_t + \varepsilon_{t+1} \tag{40.9}$$

$$\varepsilon_{t+1} = z_{t+1} \cdot \sigma_{t+1|t}, z_{t+1} \sim N(0,1), E(\varepsilon_{t+1}) = 0 \tag{40.10}$$

$$E\left(\varepsilon^2_{t+1}|\Omega_t\right) = \sigma^2_{t+1|t} = \gamma_0 + \gamma_1\varepsilon^2_t + \gamma_2\sigma^2_t \tag{40.11}$$

Earlier studies that investigate the daily risk-return trade-off generally rely on the GARCH-in-mean methodology. In risk-return regressions, it is not common to use the realized variance measures obtained from the intraday data. In this paper, we first generate the daily realized variance based on the 5-min returns on exchange rates and then estimate the following risk-return regression:

$$R_{t+1} = \alpha + \beta \cdot E_t\left[VAR^{\text{realized}}_{t+1}\right] + \varepsilon_{t+1} \tag{40.12}$$

where R_{t+1} is the 1-day ahead return on exchange rate and $E_t[VAR^{\text{realized}}_{t+1}]$ is proxied by the lagged realized variance measure, i.e., $E_t[VAR^{\text{realized}}_{t+1}]=VAR^{\text{realized}}_t$ defined in Eq. 40.3. As reported in Panel B of Table 40.3, VAR^{realized}_t has significant persistence measured by the first-order serial correlation that makes VAR^{realized}_t a reasonable proxy for the 1-day ahead expected realized variance. The slope coefficient β in Eq. 40.12, according to Merton's (1973) ICAPM, is the relative risk aversion coefficient which is expected to be positive and statistically significant.

To control for macroeconomic variables and lagged returns that may potentially affect the fluctuations in the FX market, we estimate the risk aversion coefficient, β, after controlling for the federal funds rate, term spread, default spread, and lagged return:

$$R_{t+1} \equiv \alpha + \beta \cdot VAR^{\text{realized}}_t + \lambda_1 \cdot FED_t + \lambda_2 \cdot DEF_t + \lambda_3 \cdot TERM_t + \lambda_4 \cdot R_t + \varepsilon_{t+1} \tag{40.13}$$

and test the statistical significance of β.

In addition to the GARCH-in-mean and realized volatility models, we use the range-based volatility estimator with and without control variables to test the significance of risk aversion β:

$$R_{t+1} = \alpha + \beta \cdot VAR^{\text{range}}_t + \varepsilon_{t+1} \tag{40.14}$$

$$R_{t+1} \equiv \alpha + \beta \cdot VAR^{\text{range}}_t + \lambda_1 \cdot FED_t + \lambda_2 \cdot DEF_t + \lambda_3 \cdot TERM_t + \lambda_4 \cdot R_t + \varepsilon_{t+1} \tag{40.15}$$

where VAR^{range}_t is the Parkinson's (1980) range-based estimator of the daily integrated variance defined in Eq. 40.4.

The uncovered interest rate parity indicates that the appreciation (or depreciation) rate of a currency is related to the interest rate differential of two countries.[8] Therefore, the hedging demand component of the ICAPM is proxied by the short-term interest rates of the two countries. Specifically, the intertemporal relation is tested based on the GARCH-in-mean, realized, and range volatility estimators along with the London Interbank Offer Rate (LIBOR) for the US and the corresponding foreign country:

$$R_{t+1} \equiv \alpha + \beta \cdot \sigma^2_{t+1|t} + \lambda_1 \cdot LIBOR_t^{\mathrm{US}} + \lambda_2 \cdot LIBOR_t^{\mathrm{foreign}} + \lambda_3 \cdot R_t + \varepsilon_{t+1} \tag{40.16}$$

$$R_{t+1} \equiv \alpha + \beta \cdot VAR_t^{\mathrm{realized}} + \lambda_1 \cdot LIBOR_t^{\mathrm{US}} + \lambda_2 \cdot LIBOR_t^{\mathrm{foreign}} + \lambda_3 \cdot R_t + \varepsilon_{t+1} \tag{40.17}$$

$$R_{t+1} \equiv \alpha + \beta \cdot VAR_t^{\mathrm{range}} + \lambda_1 \cdot LIBOR_t^{\mathrm{US}} + \lambda_2 \cdot LIBOR_t^{\mathrm{foreign}} + \lambda_3 \cdot R_t + \varepsilon_{t+1} \tag{40.18}$$

where $LIBOR_t^{\mathrm{US}}$ and $LIBOR_t^{\mathrm{foreign}}$ are the LIBOR rates for the US and the corresponding foreign country. To control for a potential first-order serial correlation in daily returns on exchange rates, we include the lagged return (R_t) to the conditional mean specifications.

40.4 Empirical Results

Table 40.4 presents the maximum likelihood parameter estimates and the t-statistics in parentheses for the GARCH-in-mean model. The risk aversion parameter (β) is estimated to be positive for all currencies considered in the paper, but the parameter estimates are not statistically significant, except for the British pound and the Canadian dollar. Specifically, β is estimated to be 5.18 for the euro, 4.42 for the Japanese yen, 29.07 for the British pound, 0.87 for the Swiss franc, 11.04 for the Australian dollar, and 22.40 for the Canadian dollar. Based on the Bollerslev-Wooldridge (1992) heteroscedasticity consistent covariance t-statistics reported in Table 40.4, the risk aversion coefficient has a t-statistic of 1.83 for the Canadian dollar and t-statistic of 1.77 for the British pound. Although we do not have a strong statistical significance, we can interpret this finding as a positive risk-return trade-off in the US/Canadian dollar and US dollar/lb exchange rate markets. Overall, these results indicate a positive but statistically weak relation between expected return and risk on currency.

[8] Assuming that the interest rate is 5 % per annum in the US and 2 % per annum in Japan, the uncovered interest rate parity predicts that the US dollar would depreciate against the Japanese yen by 3 %.

Table 40.4 Daily risk-return trade-off in foreign exchange markets based on the GARCH-in-mean model

Parameters	EUR	JPY	GBP	CHF	AUD	CAD
α	−0.0005	−0.0002	−0.0009	−0.0003	−0.0008	−0.0008
	(−1.46)	(−0.30)	(−2.25)	(−0.72)	(−2.26)	(−2.70)
β	5.1772	4.4206	29.065	0.8693	11.042	22.399
	(0.47)	(0.28)	(1.77)	(0.08)	(1.28)	(1.83)
γ_0	1.09×10^{-7}	1.69×10^{-6}	3.37×10^{-7}	2.60×10^{-7}	6.97×10^{-7}	2.45×10^{-7}
	(0.65)	(1.76)	(2.32)	(1.00)	(0.92)	(2.21)
γ_1	0.0301	0.0587	0.0430	0.0331	0.0541	0.0420
	(4.25)	(4.57)	(4.51)	(4.24)	(3.81)	(4.82)
γ_2	0.9672	0.8922	0.9443	0.9617	0.9318	0.9502
	(116.10)	(35.85)	(73.13)	(107.35)	(50.04)	(93.08)

The following GARCH-in-mean process is used with conditional normal density to model the intertemporal relation between expected return and risk on currency

$$
\begin{aligned}
R_{t+1} &\equiv \alpha + \beta \cdot \sigma^2_{t+1|t} + \varepsilon_{t+1} \\
\varepsilon_{t+1} &= z_{t+1} \cdot \sigma_{t+1|t}, z_{t+1} \sim N(0,1), E(\varepsilon_{t+1}) = 0 \\
E\left(\varepsilon^2_{t+1}|\Omega_t\right) &= \sigma^2_{t+1|t} = \gamma_0 + \gamma_1 \varepsilon^2_t + \gamma_2 \sigma^2_t
\end{aligned}
$$

where R_{t+1} is the daily return on exchange rates for period t+1, $\mu_{t+1|t} \equiv \alpha + \beta \cdot \sigma^2_{t+1|t}$ is the conditional mean for period t+1 based on the information set up to time t denoted by Ω_t, $\varepsilon_{t+1} = z_{t+1} \cdot \sigma_{t+1|t}$ is the error term with $E(\varepsilon_{t+1}) = 0$, $\sigma_{t+1|t}$ is the conditional standard deviation of daily returns on currency, and $z_{t+1} \sim N(0,1)$ is a random variable drawn from the standard normal density and can be viewed as information shocks in FX markets. $\sigma^2_{t+1|t}$ is the conditional variance of daily returns based on the information set up to time t denoted by Ω_t. The conditional variance, $\sigma^2_{t+1|t}$, follows a GARCH(1,1) process defined as a function of the last period's unexpected news (or information shocks), z_t, and the last period's variance, $\sigma_t^{\ 2}$. The table presents the maximum likelihood parameter estimates and the t-statistics in parentheses

Another notable point in Table 40.4 is the significance of volatility clustering. For all currencies, the conditional volatility parameters (γ_1, γ_2) are positive, between zero and one, and highly significant. The results indicate the presence of rather extreme conditionally heteroskedastic volatility effects in the exchange rate process because the GARCH parameters, γ_1 and γ_2, are found to be not only highly significant, but also the sum ($\gamma_1 + \gamma_2$) is close to one for all exchange rates considered in the paper. This implies the existence of substantial volatility persistence in the FX market.

Table 40.5 reports the daily risk aversion parameter estimates and their statistical significance for each currency after controlling for macroeconomic variables and lagged return. The risk-return coefficient estimates are similar to our earlier findings in Table 40.4. The relationship between expected return and conditional risk is positive but statistically weak for all exchange rates, except for the British pound and the Canadian dollar where we have a risk aversion parameter of 36.51 with t-stat. =1.89 for the British pound and 27.86 with t-stat. = 2.15 for the Canadian dollar. These results indicate that controlling for the hedging demand component of the ICAPM does not alter our findings.

Table 40.5 Daily risk-return trade-off in foreign exchange markets based on the GARCH-in-mean model with control variables

Parameters	EUR	JPY	GBP	CHF	AUD	CAD
α	−0.0003	0.0024	−0.0012	0.0006	0.0007	−0.0013
	(−0.13)	(1.54)	(−0.73)	(0.29)	(0.39)	(−0.95)
β	6.9731	15.986	36.514	17.913	16.478	29.329
	(0.68)	(0.96)	(1.89)	(1.17)	(1.56)	(2.15)
λ_1	-2.19×10^{-4}	-3.76×10^{-4}	-1.00×10^{-4}	-2.43×10^{-4}	-3.11×10^{-4}	-4.64×10^{-5}
	(−0.66)	(−1.13)	(−0.35)	(−0.69)	(−0.90)	(−0.18)
λ_2	5.58×10^{-5}	-7.24×10^{-4}	8.06×10^{-4}	1.00×10^{-4}	8.66×10^{-5}	7.33×10^{-4}
	(0.07)	(−0.96)	(1.13)	(0.09)	(0.10)	(1.03)
λ_3	-5.33×10^{-4}	-6.79×10^{-4}	-2.71×10^{-4}	-5.82×10^{-4}	-5.64×10^{-4}	-1.62×10^{-4}
	(−1.45)	(−1.65)	(−0.82)	(−1.43)	(−1.27)	(−0.50)
λ_4	−0.055	−0.042	−0.008	−0.062	−0.010	−0.038
	(−2.04)	(−1.48)	(−0.30)	(−2.48)	(−0.35)	(−1.45)
γ_0	1.14×10^{-7}	1.64×10^{-6}	3.13×10^{-7}	2.37×10^{-7}	7.32×10^{-7}	2.42×10^{-7}
	(0.99)	(2.94)	(2.30)	(1.83)	(2.63)	(2.15)
γ_1	0.0311	0.0581	0.0408	0.0331	0.0566	0.0417
	(4.56)	(4.41)	(4.52)	(3.81)	(4.03)	(4.80)
γ_2	0.9660	0.8938	0.9475	0.9622	0.9285	0.9506
	(97.28)	(33.74)	(77.76)	(88.65)	(53.11)	(92.60)

The following GARCH-in-mean process is used with control variables to estimate the intertemporal relation between expected return and risk on currency

$$R_{t+1} \equiv \alpha + \beta \cdot \sigma^2_{t+1|t} + \lambda_1 \cdot FED_t + \lambda_2 \cdot DEF_t + \lambda_3 \cdot TERM_t + \lambda_4 \cdot R_t + \varepsilon_{t+1}$$

$$\varepsilon_{t+1} = z_{t+1} \cdot \sigma_{t+1|t}, z_{t+1} \sim N(0,1), E(\varepsilon_{t+1}) = 0$$

$$E\left(\varepsilon^2_{t+1} | \Omega_t\right) = \sigma^2_{t+1|t} = \gamma_0 + \gamma_1 \varepsilon^2_t + \gamma_2 \sigma^2_t$$

where FED_t, DEF_t, and $TERM_t$ are macroeconomic variables that proxy for the hedging demand component of ICAPM, and R_t is the lagged daily return. The federal funds rate (*FED*) is the interest rate at which a depository institution lends immediately available funds (balances at the Federal Reserve) to another depository institution overnight. The term spread (*TERM*) is calculated as the difference between the yields on the 10-year Treasury bond and the 3-month Treasury bill. The default spread (*DEF*) is computed as the difference between the yields on the BAA-rated and AAA-rated corporate bonds. The table presents the maximum likelihood parameter estimates and the t-statistics in parentheses

Table 40.5 shows that the slope coefficient (λ_4) on the lagged return is negative for all currencies, but it is statistically significant only for the euro (with t-stat. = –2.04) and the Swiss franc (with t-stat. = –2.48).[9] We find a negative but insignificant first-order serial correlation for the Japanese yen, British pound, Australian dollar, and Canadian dollar.

Table 40.6 presents the parameter estimates and their Newey and West (1987) adjusted *t*-statistics from the risk-return regressions with realized daily variance. Panel A reports results without the control variables and tests whether the realized variance obtained from the sum of squared 5-min returns can predict 1-day ahead returns on exchange rates. The risk aversion parameter (β) is estimated to be positive for five out of six currencies considered in the paper, but only two of these parameter estimates are statistically significant at the 10 % level. Specifically, β is estimated to be 11.39 for the euro, 7.91 for the Japanese yen, 7.91 for the British pound, 13.78 for the Swiss franc, –1.09 for the Australian dollar, and 11.89 for the Canadian dollar. Based on the Newey-West (1987) t-statistics reported in Table 40.6, the Swiss franc has a risk aversion parameter of 13.78 (t-stat. = 2.21) and the euro has a risk aversion coefficient of 11.39 (t-stat. = 1.77). These results indicate that the daily realized variance measures obtained from intraday data positively predict future returns on exchange rates, but the link between risk and return is generally statistically insignificant.

Panel B of Table 40.6 presents the risk aversion coefficient estimates after controlling for the federal funds rate, term spread, default spread, and lagged return. Similar to our findings in Panel A, the risk aversion parameter is estimated to be 18.76 with t-stat. = 2.42 for the euro, 8.77 with t-stat. = 1.80 for the Japanese yen, and 18.89 with t-stat. = 2.70 for the Swiss franc, indicating a positive and significant link between the realized variance and the 1-day ahead returns on the US dollar/euro, US dollar/yen, and US dollar/Swiss franc exchange rates. There is also a positive but statistically weak relation for the British pound and the Canadian dollar.

Table 40.7 reports the parameter estimates and their Newey-West t-statistics from the risk-return regressions with the daily range variance of Parkinson (1980). As shown in both panels, with and without control variables, the risk aversion parameter (β) is estimated to be positive but statistically insignificant, except for the marginal significance of β for the Canadian dollar in Panel B. These results provide evidence that the daily range volatility obtained from the intraday data positively predict future returns on exchange rates, but there is no significant relation between risk and return on currency.

The estimates in Tables 40.6 and 40.7 present a negative and significant autocorrelation for the euro, Japanese yen, Swiss franc, and Canadian dollar. The first-order autocorrelation coefficient is negative but statistically insignificant for the British pound and the Australian dollar.

[9]Jegadeesh (1990), Lehmann (1990), and Lo and MacKinlay (1990) provide evidence for the significance of short-term reversal (or negative autocorrelation) in short-term stock returns.

Table 40.6 Daily risk-return trade-off in foreign exchange markets based on the realized variance

Panel A. Daily risk-return trade-off without control variables						
Parameters	EUR	JPY	GBP	CHF	AUD	CAD
α	−0.00073	−0.00040	−0.00042	−0.00088	−0.00027	−0.00068
	(−2.84)	(−1.59)	(−1.77)	(−2.90)	(−0.77)	(−2.25)
β	11.393	7.9064	7.9145	13.777	−1.0895	11.885
	(1.77)	(1.61)	(1.07)	(2.21)	(−0.21)	(1.38)
Panel B. Daily risk-return trade-off with control variables						
Parameters	EUR	JPY	GBP	CHF	AUD	CAD
α	0.00124	0.00257	0.00080	0.00169	0.00030	−0.00074
	(0.82)	(1.58)	(0.57)	(0.96)	(0.17)	(−0.49)
β	18.759	8.7656	10.429	18.886	−0.0858	13.608
	(2.42)	(1.80)	(1.30)	(2.70)	(−0.01)	(1.55)
λ_1	−0.00028	−0.00035	−0.00025	−0.00033	4.4×10^{-6}	−0.00007
	(−0.99)	(−1.06)	(−0.90)	(−1.03)	(0.01)	(−0.25)
λ_2	−0.00051	−0.00112	0.00013	−0.00076	−0.00059	0.00041
	(−0.72)	(−1.33)	(0.18)	(−0.91)	(−0.63)	(0.60)
λ_3	−0.00054	−0.00054	−0.00041	0.00065	−0.00004	−0.00012
	(−1.54)	(−1.36)	(−1.19)	(1.67)	(−0.07)	(−0.37)
λ_4	−0.068	−0.048	−0.013	−0.077	−0.017	−0.071
	(−2.81)	(−1.81)	(−0.49)	(−3.00)	(−0.56)	(−2.88)

The following regression is estimated with and without control variables to test the significance of the intertemporal relation between expected return and risk on currency

$$R_{t+1} \equiv \alpha + \beta \cdot VAR_t^{\text{realized}} + \lambda_1 \cdot FED_t + \lambda_2 \cdot DEF_t + \lambda_3 \cdot TERM_t + \lambda_4 \cdot R_t + \varepsilon_{t+1}$$

where VAR_t^{realized} is the daily realized variance computed as the sum of squared 5-min returns on exchange rates. The table presents the parameter estimates and their Newey and West (1987) t-statistics in parentheses

An interesting observation in Tables 40.5, 40.6, and 40.7 is that the slope coefficients (λ_1, λ_2, λ_3) on the lagged macroeconomic variables are found to be statistically insignificant, except for some marginal significance for the term spread in the regressions with the Swiss franc. Although one would think that unexpected news in macroeconomic variables could be viewed as risks that would be rewarded in the FX market, we find that the changes in federal funds rate and term and default spreads do not affect time-series variation in daily exchange rate returns. Our interpretation is that it would be very difficult for macroeconomic variables to explain daily variations in exchange rates. If we examined the risk-return trade-off at lower frequency (such as monthly or quarterly frequency), we might observe significant impact of macroeconomics variables on monthly or quarterly variations in exchange rates.

Panel A of Table 40.8 presents the maximum likelihood parameter estimates and the t-statistics in parentheses for the GARCH-in-mean model with LIBOR rates for the US and the corresponding foreign country. The risk aversion parameter (β) is estimated to be positive for all currencies, but the parameter estimates are

Table 40.7 Daily risk-return trade-off in foreign exchange markets based on the range volatility

Panel A. Daily risk-return trade-off without control variables						
Parameters	EUR	JPY	GBP	CHF	AUD	CAD
α	−0.00051	−0.00022	−0.00038	−0.00053	−0.00053	−0.00048
	(−2.54)	(−1.11)	(−2.15)	(−2.28)	(−2.15)	(−2.67)
β	6.6224	4.6907	7.8004	6.8730	4.4630	8.4727
	(1.32)	(1.15)	(1.31)	(1.56)	(1.05)	(1.61)
Panel B. Daily risk-return trade-off with control variables						
Parameters	EUR	JPY	GBP	CHF	AUD	CAD
α	0.00179	0.00250	0.00088	0.00238	0.00042	−0.00053
	(1.22)	(1.57)	(0.65)	(1.40)	(0.23)	(−0.37)
β	8.6053	5.0792	8.3969	7.3411	5.3794	9.2256
	(1.58)	(1.24)	(1.36)	(1.52)	(1.14)	(1.69)
λ_1	−0.00036	−0.00030	−0.00026	−0.00045	−0.00010	−0.00005
	(−1.26)	(−0.95)	(−0.93)	(−1.41)	(−0.25)	(−0.16)
λ_2	−0.00041	−0.00106	0.00015	−0.00049	−0.00038	0.00033
	(−0.57)	(−1.29)	(0.22)	(−0.59)	(−0.42)	(0.48)
λ_3	−0.00056	−0.00049	−0.00040	0.00070	−0.00019	−0.00009
	(−1.59)	(−1.25)	(−1.18)	(1.81)	(−0.40)	(−0.28)
λ_4	−0.067	−0.051	−0.012	−0.078	−0.026	−0.067
	(−2.77)	(−1.89)	(−0.45)	(−3.13)	(−0.87)	(−2.71)

The following regression is estimated with and without control variables to test the significance of the intertemporal relation between expected return and risk on currency

$$R_{t+1} \equiv \alpha + \beta \cdot VAR_t^{\text{range}} + \lambda_1 \cdot FED_t + \lambda_2 \cdot DEF_t + \lambda_3 \cdot TERM_t + \lambda_4 \cdot R_t + \varepsilon_{t+1}$$

where $VAR_t^{\text{range}} = 0.361[\ln(P_t^{\max}) - \ln(P_t^{\min})]^2$ is Parkinson's (1980) range-based estimator of the daily integrated variance. The table presents the parameter estimates and their Newey-West (1987) t-statistics in parentheses

statistically significant only for the British pound, Australian dollar, and Canadian dollar. Specifically, β is estimated to be 30.87 for the British pound, 17.75 for the Australian dollar, and 32.90 for the Canadian dollar. Based on the Bollerslev-Wooldridge heteroscedasticity consistent covariance t-statistics reported in Table 40.8, the risk aversion coefficient has a t-statistic of 1.82 for the British pound, 1.84 for the Australian dollar, and 2.36 for the Canadian dollar. Although we do not have a strong statistical significance, we can interpret this finding as a positive risk-return trade-off in the US dollar/British pound, US/Australian dollar, and US/Canadian dollar markets. Overall, the results indicate a positive but statistically weak relation between expected return and risk on currency.

Another point worth mentioning in Panel A is that the slope coefficients on the US LIBOR rate are estimated to be positive and statistically significant at the 5 % level for the euro, Japanese yen, and Swiss franc and significant at the 10 % level for the Canadian dollar. As expected, the slope coefficients on the LIBOR rates of the corresponding foreign country turn out to be negative but statistically insignificant.

Table 40.8 Daily risk-return trade-off in foreign exchange markets with LIBOR interest rates

Panel A. GARCH-in-mean						
Parameters	EUR	JPY	GBP	CHF	AUD	CAD
α	−0.0004	−0.0011	−0.0019	−0.0004	−0.0004	0.0006
	(−0.62)	(−1.82)	(−1.97)	(−0.31)	(−0.27)	(0.58)
β	6.7962	13.422	30.871	18.036	17.753	32.898
	(0.66)	(0.84)	(1.82)	(1.33)	(1.84)	(2.36)
λ_1	2.16×10^{-4}	2.86×10^{-4}	1.41×10^{-5}	3.30×10^{-4}	2.57×10^{-4}	1.95×10^{-4}
	(1.97)	(2.53)	(0.12)	(2.19)	(1.47)	(1.66)
λ_2	-2.61×10^{-4}	-8.94×10^{-4}	1.83×10^{-4}	-3.44×10^{-4}	-3.35×10^{-4}	-4.81×10^{-4}
	(−1.10)	(−1.32)	(0.78)	(−1.13)	(−0.78)	(−1.73)
λ_3	−0.055	−0.042	−0.008	−0.062	−0.010	−0.039
	(−2.22)	(−1.62)	(−0.31)	(−2.46)	(−0.36)	(−1.48)
γ_0	1.15×10^{-7}	1.64×10^{-6}	3.11×10^{-7}	2.64×10^{-7}	7.35×10^{-7}	2.44×10^{-7}
	(1.22)	(1.30)	(2.25)	(0.73)	(0.99)	(2.13)
γ_1	0.0300	0.0575	0.0408	0.0338	0.0563	0.0424
	(4.44)	(4.49)	(4.51)	(3.72)	(3.93)	(4.87)
γ_2	0.9672	0.8946	0.9476	0.9608	0.9287	0.9498
	(96.57)	(32.83)	(77.50)	(79.92)	(50.57)	(92.31)
Panel B. Realized variance						
Parameters	EUR	JPY	GBP	CHF	AUD	CAD
α	−0.00074	−0.00108	−0.00150	−0.00171	−0.00118	−0.00024
	(−1.08)	(−2.52)	(−1.61)	(−2.86)	(−0.77)	(−0.37)
β	17.558	10.272	10.499	18.730	−0.708	15.199
	(2.24)	(1.95)	(1.31)	(2.63)	(−0.11)	(1.69)
λ_1	2.5×10^{-4}	3.1×10^{-4}	-3.14×10^{-6}	4.5×10^{-4}	4.6×10^{-6}	1.4×10^{-4}
	(2.39)	(2.47)	(−0.03)	(2.96)	(0.026)	(1.00)
λ_2	-3.3×10^{-4}	-1.6×10^{-3}	2.2×10^{-4}	-6.5×10^{-4}	1.6×10^{-4}	-3.0×10^{-4}
	(−1.81)	(−2.01)	(0.90)	(−2.19)	(0.42)	(−1.03)
λ_3	−0.069	−0.048	−0.013	−0.078	−0.016	−0.073
	(−2.85)	(−1.79)	(−0.51)	(−3.10)	(−0.53)	(−2.91)
Panel C. Range variance						
Parameters	EUR	JPY	GBP	CHF	AUD	CAD
α	−0.00015	−0.00079	−0.00136	−0.00090	−0.00079	−0.00018
	(−0.26)	(−2.13)	(−1.49)	(−1.86)	(−0.52)	(−0.28)
β	8.181	5.558	8.416	7.400	5.249	9.761
	(1.49)	(1.29)	(1.37)	(1.51)	(1.03)	(1.77)
λ_1	1.9×10^{-4}	2.8×10^{-4}	-1.1×10^{-5}	3.5×10^{-4}	7.26×10^{-5}	9.5×10^{-5}
	(−1.92)	(2.32)	(−0.09)	(2.51)	(0.46)	(0.68)

(*continued*)

Table 40.8 (continued)

Panel C. Range variance						
Parameters	EUR	JPY	GBP	CHF	AUD	CAD
λ_2	-3.4×10^{-4}	-1.3×10^{-3}	2.1×10^{-4}	-6.3×10^{-4}	1.08×10^{-6}	-1.9×10^{-4}
	(−1.86)	(−1.84)	(0.88)	(−2.20)	(0.003)	(−0.63)
λ_3	−0.068	−0.051	−0.012	−0.080	−0.025	−0.067
	(−2.82)	(−1.87)	(−0.47)	(−3.22)	(−0.85)	(−2.71)

The following regressions with GARCH-in-mean, realized variance, and range variance are estimated to test the significance of the intertemporal relation between expected return and risk on currency

$$R_{t+1} \equiv \alpha + \beta \cdot \sigma^2_{t+1|t} + \lambda_1 \cdot LIBOR_t^{\mathrm{US}} + \lambda_2 \cdot LIBOR_t^{\mathrm{foreign}} + \lambda_3 \cdot R_t + \varepsilon_{t+1},$$
$$R_{t+1} \equiv \alpha + \beta \cdot VAR_t^{\mathrm{realized}} + \lambda_1 \cdot LIBOR_t^{\mathrm{US}} + \lambda_2 \cdot LIBOR_t^{\mathrm{foreign}} + \lambda_3 \cdot R_t + \varepsilon_{t+1},$$
$$R_{t+1} \equiv \alpha + \beta \cdot VAR_t^{\mathrm{range}} + \lambda_1 \cdot LIBOR_t^{\mathrm{US}} + \lambda_2 \cdot LIBOR_t^{\mathrm{foreign}} + \lambda_3 \cdot R_t + \varepsilon_{t+1},$$

where $LIBOR_t^{\mathrm{US}}$ and $LIBOR_t^{\mathrm{foreign}}$ are the LIBOR rates for the US and the corresponding foreign country

Panel B of Table 40.8 reports the parameter estimates and their Newey-West adjusted t-statistics from the risk-return regressions with daily realized variance after controlling for the LIBOR rates and the lagged return. The results indicate a positive and significant link between the realized variance and the 1-day ahead returns on the euro, Japanese yen, Swiss franc, and Canadian dollar. There is also a positive but statistically weak relation for the British pound.

Panel C of Table 40.8 shows the parameter estimates and their Newey-West t-statistics from the risk-return regressions with the daily range variance of Parkinson (1980). With LIBOR rates and the lagged return, the risk aversion parameter (β) is estimated to be positive for all currencies but statistically significant only for the Canadian dollar. Overall, the results provide evidence that after controlling for the interest rate differential of two countries, there is a positive but statistically weak relation between risk and return on currency.

Similar to our earlier findings from the GARCH-in-mean model, Panels B and C of Table 40.8 show that the slope coefficients on the US LIBOR rate are generally positive, whereas the slopes on the corresponding foreign LIBOR rates are negative with a few exceptions.

Many studies fail to identify a statistically significant intertemporal relation between risk and return of the stock market portfolios. French et al. (1987) find that the coefficient estimate is not significantly different from zero when they use past daily returns to estimate the monthly conditional variance.[10] Chan et al. (1992) employ a bivariate GARCH-in-mean model to estimate the conditional variance,

[10]When testing monthly risk-return trade-off, French et al. (1987) use the monthly realized variance obtained from the sum of squared daily returns within a month.

and they also fail to obtain a significant coefficient estimate for the United States. Campbell and Hentchel (1992) use the quadratic GARCH (QGARCH) model of Sentana (1995) to determine the existence of a risk-return trade-off within an asymmetric GARCH-in-mean framework. Their estimate is positive for one sample period and negative for another sample period, but neither is statistically significant. Glosten et al. (1993) use monthly data and find a negative but statistically insignificant relation from two asymmetric GARCH-in-mean models. Based on semi-nonparametric density estimation and Monte Carlo integration, Harrison and Zhang (1999) find a significantly positive risk and return relation at a 1-year horizon, but they do not find a significant relation at shorter holding periods such as 1 month. Using a sample of monthly returns and implied and realized volatilities for the S and P 500 index, Bollerslev and Zhou (2006) find an insignificant intertemporal relation between expected return and realized volatility, whereas the relation between return and implied volatility turns out to be significantly positive.

Several studies find that the intertemporal relation between risk and return is negative (e.g., Campbell 1987; Breen et al. 1989; Turner et al. 1989; Nelson 1991; Glosten et al. 1993; Harvey 2001; Brandt and Kang 2004). Some studies do provide evidence supporting a positive and significant relation between expected return and risk on stock market portfolios (e.g., Bollerslev et al. 1988; Scruggs 1998; Ghysels et al. 2005; Bali and Peng 2006; Guo and Whitelaw 2006; Lundblad 2007; Bali 2008).

Merton's (1973) ICAPM provides a theoretical model that gives a positive equilibrium relation between the conditional first and second moments of excess returns on the aggregate market portfolio. However, Abel (1988), Backus and Gregory (1993), and Gennotte and Marsh (1993) develop models in which a negative relation between expected return and volatility is consistent with equilibrium. As summarized above, there has been a lively debate on the existence and direction of a risk-return trade-off, and empirical studies are still not in agreement for the stock market portfolios. The empirical results presented in Tables 40.4, 40.5, 40.6, 40.7, and 40.8 indicate that the intertemporal relation between expected return and risk on currency is positive but in most cases statistically insignificant. Hence, our findings from the FX market are in line with some of the earlier studies that investigated the significance of a risk-return trade-off for the stock market.

40.5 Time-Varying Risk Aversion in the Foreign Exchange Market

Chou et al. (1992), Harvey (2001), and Lettau and Ludvigson (2010) suggest that the risk-return relation for the stock market may be time varying. In the existing literature, there is no study investigating the presence and significance of time-varying risk aversion in the FX market. We have so far assumed a constant risk-return trade-off in currency markets and found a positive but statistically insignificant relation between expected return and risk on exchange rates.

We now estimate the dependence of expected returns on the lagged realized variance over time using rolling regressions. This also allows us to check whether our results are driven by a particular sample period. We estimate the risk-return relation specified in Eqs. 40.12 and 40.13 for the six exchange rates with rolling samples. We used two different rolling regression approaches. The first one uses a fixed rolling window of 250 business days (i.e., approximately 1 year), whereas the second one starts with the in-sample period of 250 business days and then extends the sample by adding each daily observation to the estimation while keeping the start date constant.

Figure 40.1 plots the estimated relative risk aversion parameters (β) and their statistical significance over time from the fixed rolling window of 250 days.[11] Specifically, the first 250 daily return observations of exchange rates and their realized variances (from 1/3/2002 to 1/8/2003) are used for estimation of the relative risk aversion parameter for 1/8/2003. The sample is then rolled forward by removing the first observation of the sample and adding one to the end, and another 1-day ahead risk-return relationship is measured. This recursive estimation procedure is repeated until the sample is exhausted on March 31, 2008. The estimated relative risk aversion parameter over the fixed rolling sample period represents the average degree of risk aversion over that sample period. Computation of the relative risk aversion parameters using a rolling window of data allows us to observe the time variation in an investors' average risk aversion.

A common observation in Fig. 40.1 is that there is a strong time-series variation in the risk aversion estimates for all currencies considered in the paper. The first panel in Fig. 40.1 indicates that in the US dollar/euro FX market, the aggregate risk aversion is generally positive with some exceptions in the second half of 2006 and from May to August 2007. For the out-of-sample period of January 2003 to March 2008, only 208 out of 1,306 daily risk aversion estimates are negative. Based on the Newey-West adjusted t-statistics, all of these negative risk aversion estimates are statistically insignificant. 143 (291) out of 1,098 positive risk aversion estimates turn out to be statistically significant at least at the 5 % level (10 % level). These results indicate a positive but statistically insignificant time-varying risk aversion in the US dollar/euro market.

The second panel in Fig. 40.1 displays that in the Japanese yen market, the aggregate risk aversion is generally positive, but there are quite a lot of days in which we observe a negative relation between expected return and risk in the US dollar/yen market. 431 out of 1,306 daily risk aversion estimates are negative, but about one third is statistically significant at the 10 % level. 185 (314) out of 875 positive risk aversion estimates turn out to be statistically significant at least at the 5 % level (10 % level). These results indicate that there is a positive but not strong time-varying risk aversion in the US dollar/yen exchange rate market.

[11] Since the time-varying risk aversion coefficients from estimating Eqs. 40.12 and 40.13 with and without control variables turn out to be very similar, we only report results from the full specification of Eq. 40.13. Time-varying risk aversion estimates obtained from the parsimonious specification of Eq. 40.12 are available from the authors upon request.

Third panel in Fig. 40.1 shows that in the US dollar/lb sterling market, the risk aversion is generally positive, but there is a long period of time in which we observe a negative relation between expected return and risk in the US dollar/lb market. Specifically, 872 out of 1,306 daily risk aversion estimates are positive, but only 5 out of 872 are marginally significant. Similarly, only 46 out of 434 negative risk aversion estimates turn out to be statistically significant at the 10 % level. These results provide evidence that although there is a significant time variation in the aggregate risk aversion, it is not clear whether the currency trade generates a larger or smaller risk premium at times when the US dollar/lb FX market is riskier.

The fourth panel in Fig. 40.1 indicates that in the Swiss franc market, the risk aversion is estimated to be positive throughout the sample period (2002–2008), except for a few months in 2006. Only 71 out of 1,306 daily risk aversion estimates are negative, but none of them is statistically significant. 353 (467) out of 1,235 positive risk aversion estimates turn out to be statistically significant at least at the 5 % level (10 % level). These results indicate a positive and relatively strong time-varying risk aversion, implying that the currency trade generates a larger risk premium at times when the US dollar/Swiss franc trade becomes riskier.

The fifth panel in Fig. 40.1 indicates that in the Australian dollar market 736 out of 1,306 daily risk aversion estimates are positive, but none of them is statistically significant. Only 65 out of 570 negative risk aversion estimates turn out to be marginally significant at the 10 % level. The figure indicates a strong time-varying risk aversion, but there is no significantly positive or negative relation between risk and return in the US/Australian dollar exchange rate market.

The last panel in Fig. 40.1 demonstrates that in the US/Canadian dollar market, for slightly more than half of the sample, the risk aversion is estimated to be positive and slightly less than half of the sample it turns out to be negative. However, based on the t-statistics of these risk aversion estimates, there is no evidence for a significantly positive or negative link between expected return and risk on currency. Only 35 out of 757 positive risk aversion coefficients and only 46 out of 549 negative risk aversion parameters are found to be significant at the 10 % level. Although there is a significant time-series variation in the aggregate risk aversion, trading in the US/Canadian dollar FX market does not provide clear evidence for a larger or smaller risk premium at times when the market is riskier.

Figure 40.2 plots the estimated relative risk aversion parameters (β) and their statistical significance over time from the rolling regressions with a fixed starting date. Specifically, the first 250 daily return observations of exchange rates and their realized variances (from 1/3/2002 to 1/7/2003) are used for estimation of the relative risk aversion parameter for 1/8/2003. The sample is then extended by adding one observation to the end (from 1/3/2002 to 1/8/2003), and the 1-day ahead risk-return relation is measured for 1/9/2003. This recursive estimation procedure is repeated until March 31, 2008.

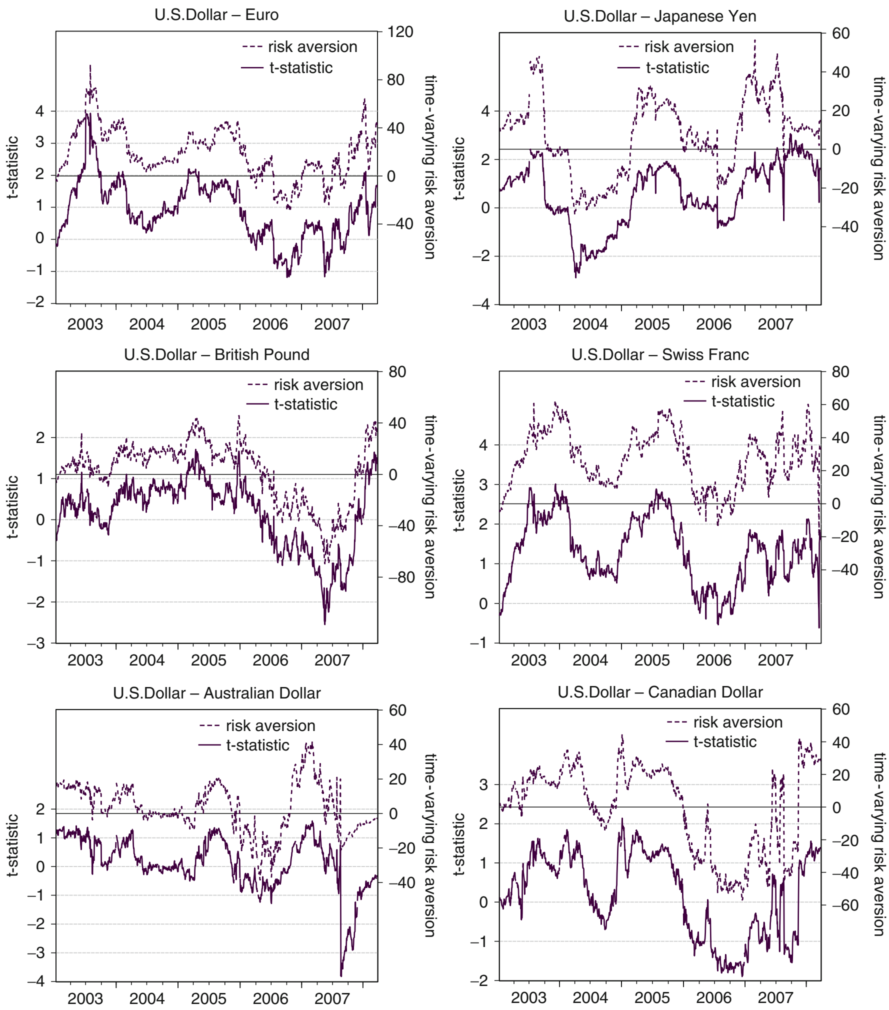


Fig. 40.1 Rolling regression estimates from the fixed length window of 250 days

Similar to our findings from the fixed rolling window regressions, Fig. 40.2 provides evidence for a significant time variation in the risk aversion estimates for all currencies considered in the paper. The first panel in Fig. 40.2 shows that in the US dollar/euro market, the aggregate risk aversion is positive with a few exceptions in January 2003. Only 16 out of 1,306 risk aversion estimates are negative, but none of these estimates is statistically significant based on the Newey-West t-statistics. 795 (870) out of 1,290 positive risk aversion estimates turn out to be statistically significant at least at the 5 % level (10 % level). These results indicate a positive and strong time-varying risk aversion in the US dollar/euro market.

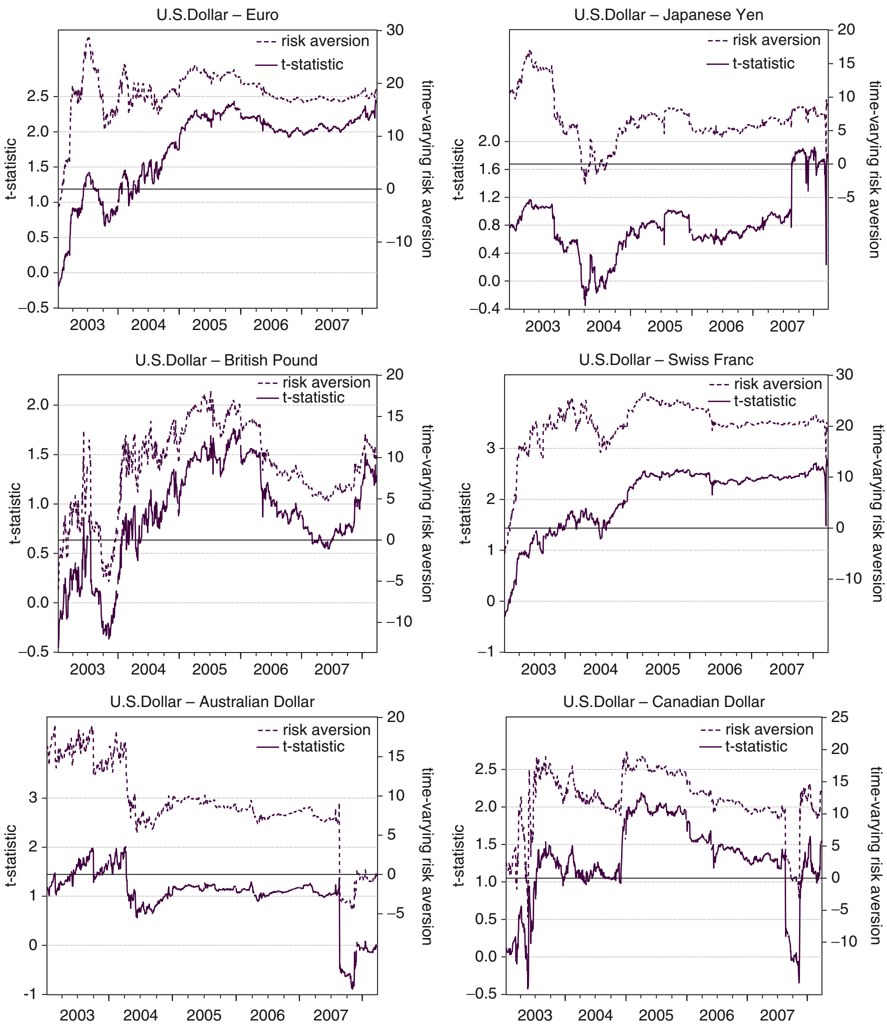


Fig. 40.2 Rolling regression estimates from the windows with fixed starting point

The second panel in Fig. 40.2 shows that in the US dollar/yen FX market, the aggregate risk aversion is positive with a few exceptions from March to June 2004. Only 68 out of 1,306 risk aversion estimates are negative and all of them are statistically insignificant. 129 out of 1,238 positive risk aversion estimates turn out to be marginally significant at the 10 % level. These results imply a positive but statistically weak time-varying risk aversion in the US dollar/yen market.

The third panel in Fig. 40.2 depicts that in the pound sterling market, the risk aversion is positive throughout the sample, except for a short period of time in 2003. Only 90 out of 1,306 risk aversion estimates are negative, but they are not statistically significant. Although there is a significant time variation in the risk aversion and most of the risk-return coefficients are positive, only 41 out of 1,216 positive risk aversion estimates turn out to be significant at the 10 % level. Therefore, it is not clear whether the currency trade generates a larger or smaller risk premium at times when the US dollar/lb market is riskier.

The fourth panel in Fig. 40.2 provides evidence that in the Swiss franc market, the risk aversion is estimated to be positive throughout the sample period (2002–2008), except for a few days in January 2003. Only 20 out of 1,306 risk aversion estimates are negative but statistically insignificant. 816 (934) out of 1,286 positive risk aversion estimates turn out to be statistically significant at least at the 5 % level (10 % level). These results suggest a positive and strong time-varying risk aversion, implying that the currency trade generates a larger risk premium at times when the US dollar/Swiss franc exchange rate market is riskier.

The fifth panel in Fig. 40.2 shows that in the Australian dollar market only 138 out of 1,306 risk aversion estimates are negative with no statistical significance even at the 10 % level. Only 106 out of 1,168 positive risk aversion coefficients are found to be marginally significant at the 10 % level. Although there is significant time variation in the aggregate risk aversion, the results do not suggest a strong positive or negative link between expected return and risk in the US/Australian dollar market.

The last panel in Fig. 40.2 demonstrates that in the US/Canadian dollar market, the risk aversion is estimated to be positive, except for a few days in May, October, and November 2003. Similar to our earlier findings, only 41 out of 1,308 risk aversion estimates are negative with very low t-statistics. However, based on the statistical significance of positive risk aversion estimates, there is no evidence for a strong positive link between expected return and risk on currency either. Only 276 out of 1,265 positive risk aversion coefficients are found to be significant at the 10 % level. Although there is a significant time-series variation in the aggregate risk aversion, trading in the US/Canadian dollar FX market does not provide clear evidence for a larger or smaller risk premium at times when the market is riskier.

40.6 Testing Merton's (1973) ICAPM in Currency Market

Merton's (1973) ICAPM implies the following equilibrium relation between risk and return for any risky asset i:

$$\mu_i - r = A \cdot \sigma_{im} + B \cdot \sigma_{ix}, \tag{40.19}$$

where r is the risk-free interest rate, $\mu_i - r$ is the expected excess return on the risky asset i, σ_{im} denotes the covariance between the returns on the risky asset i and the market portfolio m, and σ_{ix} denotes a $(1 \times k)$ row of covariances between the returns on risky asset i and the k state variables x. A denotes the average relative risk aversion of market investors, and B measures the market's aggregate reaction to shifts in a k-dimensional state vector that governs the stochastic investment opportunity. Equation 40.19 states that in equilibrium, investors are compensated in terms of expected return for bearing market risk and for bearing the risk of unfavorable shifts in the investment opportunity set.

Merton (1980) shows that the intertemporal hedging demand component ($B \cdot \sigma_{ix}$) is economically and statistically smaller than the market risk component ($A \cdot \sigma_{im}$) of ICAPM. While testing the significance of A and B at daily frequency, Bali and Engle (2010) provide supporting evidence for Merton (1980) that the conditional covariances of individual stocks with the market portfolio have positive and statistically significant loading, whereas the innovations in state variables are not priced in the stock market. That is, the conditional covariances of stock returns with the unexpected news in state variables have insignificant loadings.

We examine Merton's (1973) ICAPM based on the following system of equations:

$$\begin{aligned} R_{i,t+1} &= C_i + A \cdot \sigma_{im,t} + \varepsilon_{i,t+1} \\ R_{m,t+1} &= C_m + A \cdot \sigma^2_{m,t} + \varepsilon_{m,t+1} \end{aligned} \tag{40.20}$$

where the expected conditional covariance of individual exchange rates with the currency market, $E_t(\sigma_{im,t+1})$, is represented by the 1-day lagged realized covariance, i.e., $E_t(\sigma_{im,t+1}) = \sigma_{im,t}$. Similarly, the expected conditional variance of the currency market, $E_t(\sigma_{m,t+1}{}^2)$, is represented by the 1-day lagged realized variance, i.e., $E_t(\sigma_{m,t+1}{}^2) = \sigma_{m,t}{}^2$.[12]

The currency market portfolio is measured by the "value-weighted" average returns on EUR, JPY, GBP, CHF, AUD, and CAD. The weights are obtained from the "US Dollar Index." Just as the Dow Jones Industrial Average reflects the general state of the US stock market, the US Dollar Index (USDX) reflects the general assessment of the US dollar. USDX does it through exchange rates averaging of US dollar and six most tradable global currencies. The weights are 57.6 % for EUR, 13.6 % for JPY, 11.9 % for GBP, 9.1 % CAD, 4.2 % for AUD, and 3.6 % for CHF. In our empirical analysis, daily returns on the currency market, $R_{m,t+1}$, are calculated by multiplying daily returns on the six exchange rates by the aforementioned weights.

We estimate the system of Eq. 40.20 using an ordinary least square (OLS) as well as a weighted least square method that allows us to place constraints on coefficients across equations. We constrain the slope coefficient (A) on the lagged realized variance-covariance matrix ($\sigma_{im,t}$, $\sigma_{m,t}{}^2$) to the same value across all the currencies for cross-sectional consistency. We allow the intercepts (C_i, C_m) to differ across all the currencies. Under the null hypothesis of the ICAPM, the intercepts should be jointly zero, and the common slope coefficient (A) should be positive and statistically significant. We use insignificant estimates of A and the deviations from zero of the intercept estimates as a test against the validity and sufficiency of

[12]Daily realized covariances between the exchange rates and the currency market and daily realized variance of the currency market are computed using 5-min returns in a day.

the ICAPM. In addition to the OLS panel estimates, we compute the t-statistics of the parameter estimates accounting for heteroskedasticity and autocorrelation as well as contemporaneous cross-correlations in the error terms. This estimation methodology for the system of Eq. 40.20 can be regarded as an extension of the seemingly unrelated regression (SUR) method.

Table 40.9 presents the OLS and SUR panel regression estimates of the currency-specific intercepts, common slope coefficients on the lagged realized variance-covariance matrix, and their t-statistics. The parameters and their t-statistics are estimated using the daily returns on the currency market and the six exchange rates. The last row reports the Wald statistics with p-values from testing the joint hypothesis that all intercepts equal zero: H_0: $C_1 = C_2 = \ldots = C_6 = C_m = 0$. A notable point in Table 40.9 is that the common slope coefficient (A) is positive and statistically significant. Specifically, the risk aversion coefficient on the realized variance-covariance matrix is estimated to be 23.33 with a t-statistic of 5.40. After correcting for heteroscedasticity, autocorrelation, and contemporaneous cross-correlations, the SUR estimate of the risk aversion coefficient turns out to be 15.80 with t-stat. = 2.57. These results indicate a positive and significant relation between risk and return on the currency market. Another notable point in Table 40.9 is that for both the OLS and SUR estimates, the Wald statistics reject the hypothesis that all intercepts equal zero. This implies that the market risk alone cannot explain the entire time-series variation in exchange rates.

According to the original ICAPM of Merton (1973), the relative risk aversion coefficient (A) is restricted to the same value across all risky assets, and it is positive and statistically significant. The common slope estimates in Table 40.9 provide empirical support for the positive risk-return trade-off.

We now test whether the slopes on (σ_{im}, ${\sigma_m}^2$) are different across currencies. We examine the sign and statistical significance of different slope coefficients (A_i, A_m) on (σ_{im}, ${\sigma_m}^2$) in the following system of equations:

$$\begin{aligned} R_{i,t+1} &= C_i + A_i \cdot \sigma_{im,t} + \varepsilon_{i,t+1}, \\ R_{m,t+1} &= C_m + A_m \cdot \sigma^2_{m,t} + \varepsilon_{m,t+1} \end{aligned} \tag{40.21}$$

To determine whether there is a common slope coefficient (A) that corresponds to the average relative risk aversion, we first estimate the currency-specific slope coefficients (A_i, A_m) and then test the joint hypothesis that H_0: $A_1 = A_2 = \ldots = A_6 = A_m$.

Table 40.10 presents the OLS and SUR parameter estimates using daily returns on the six exchange rates and the value-weighted currency market index. As compared to Eq. 40.20, we have additional six-slope coefficients to estimate in Eq. 40.21. As shown in Table 40.10, all of the slope coefficients (A_i, A_m) are estimated to be positive and highly significant without any exception. These results indicate a positive and significant intertemporal relation between risk and return on the currency market. We examine the cross-sectional consistency of the intertemporal relation by testing the equality of slope coefficients based on the Wald statistics. As reported in Table 40.10, the Wald statistics, from testing

Table 40.9 Testing Merton's (1973) ICAPM with a common slope coefficient

	OLS panel regression		SUR panel regression	
	Intercept	t-stat.	Intercept	t-stat.
AUD	0.00076	4.69	0.00063	3.15
EUR	0.00097	5.24	0.00077	3.53
GBP	0.00062	3.82	0.00049	2.95
CAD	−0.00049	−3.24	−0.00041	−2.77
CHF	−0.00085	−4.72	−0.00066	−2.98
JPY	−0.00045	−2.79	−0.00032	−1.79
Market	−0.00078	−4.50	−0.00061	−3.35
Risk aversion	Slope	t-stat.	Slope	t-stat.
	23.33	5.40	15.80	2.57
H_0: Intercepts = 0	Wald	p-value	Wald	p-value
	52.20	0.00	17.98	0.0121

Entries report the OLS and SUR panel regression estimates based on the following system of equations

$$R_{i,t+1} = C_i + A \cdot \sigma_{im,t} + \varepsilon_{i,t+1},$$
$$R_{m,t+1} = C_m + A \cdot \sigma^2_{m,t} + \varepsilon_{m,t+1},$$

where $\sigma_{im,t}$ is the 1-day lagged realized covariance between the exchange rate and the currency market. ${\sigma_{m,t}}^2$ is the 1-day lagged realized variance of the currency market. A is a common slope coefficient on the lagged realized variance-covariance matrix. (C_i, C_m) denotes currency-specific intercepts for AUD, EUR, GBP, CAD, CHF, JPF, and the currency market. The last row reports the Wald statistics with p-values from testing the joint hypothesis that all intercepts equal zero: H_0: $C_1 = C_2 = \ldots = C_6 = C_m = 0$

the joint hypothesis that H_0: $A_1 = A_2 = \ldots = A_6 = A_m$, is 1.41 for OLS and 5.24 for SUR, which fail to reject the null hypothesis. These results indicate the equality of positive slope coefficients across all currencies, which empirically validates the ICAPM.

40.7 Conclusion

There is an ongoing debate in the literature about the intertemporal relation between stock market risk and return and the extent to which expected stock returns are related to expected market volatility. Recently, some studies have provided evidence for a significantly positive link between risk and return in the aggregate stock market, but the risk-return trade-off is generally found to be insignificant and sometimes even negative. This paper is the first to investigate the presence and significance of an intertemporal relation between expected return and risk in the foreign exchange market. The paper provides new evidence on the ICAPM by using high-frequency intraday data on currency and by presenting significant time variation in the risk aversion parameter. We utilize daily and 5-min returns on the spot exchange rates of the US dollar vis-à-vis six major currencies (the euro, Japanese yen, British pound sterling, Swiss franc, Australian dollar, and Canadian dollar)

Table 40.10 Testing Merton's (1973) ICAPM with different slope coefficients

	OLS panel regression				SUR panel regression			
	Intercept	t-stat.	Slope	t-stat.	Intercept	t-stat.	Slope	t-stat.
AUD	0.00084	2.95	28.51	2.08	0.00072	2.85	21.01	1.90
EUR	0.00080	2.80	17.12	1.83	0.00071	3.13	13.62	2.06
GBP	0.00060	2.35	22.19	1.73	0.00030	1.38	4.60	0.46
CAD	−0.00063	−3.04	35.92	2.47	−0.00041	−2.23	15.67	1.36
CHF	−0.00082	−2.66	21.92	2.10	−0.00070	−2.80	17.32	2.26
JPY	−0.00048	−1.89	25.25	2.06	−0.00050	−2.30	26.29	2.77
Market	−0.00075	−3.08	21.92	2.31	−0.00059	−3.14	14.53	2.29
H_0: Intercepts = 0	Wald	p-value			Wald	p-value		
	51.40	0.00			15.15	0.0341		
H_0: Equal slopes	Wald	p-value			Wald	p-value		
	1.41	0.9655			5.24	0.5131		

Entries report the OLS and SUR panel regression estimates based on the following system of equations

$$R_{i,t+1} = C_i + A_i \cdot \sigma_{im,t} + \varepsilon_{i,t+1},$$
$$R_{m,t+1} = C_m + A_m \cdot \sigma^2_{m,t} + \varepsilon_{m,t+1},$$

where $\sigma_{im,t}$ is the 1-day lagged realized covariance between the exchange rate and the currency market. $\sigma^2_{m,t}$ is the 1-day lagged realized variance of the currency market. (A_i, A_m) denotes currency-specific slope coefficients on the lagged realized variance-covariance matrix. (C_i, C_m) denotes currency-specific intercepts for AUD, EUR, GBP, CAD, CHF, JPF, and the currency market. The Wald statistics with p-values are reported from testing the joint hypothesis that all intercepts equal zero: H_0:$C_1 = C_2 = \ldots = C_6 = C_m = 0$. The last row presents the Wald statistics from testing the equality of currency-specific slope coefficients H_0: $A_1 = A_2 = \ldots = A_6 = A_m$

and test the existence and significance of a risk-return trade-off in the FX market using the GARCH, realized, and range-based volatility measures. The maximum likelihood parameter estimates of the GARCH-in-mean model and the risk-return regressions with daily realized and range volatility estimators indicate that the intertemporal relation between risk and return is generally positive but statistically weak in the FX market.

We provide strong evidence on the time variation of risk aversion parameters for all currencies considered in the paper. However, the direction of a relationship between expected return and risk is not clear. The results indicate a positive but not strong time-varying risk aversion in the US dollar/euro exchange rate market. The risk-return regressions with realized variance provide evidence for a positive but statistically weak risk aversion estimates in the US dollar/yen market. Although there is a significant time variation in risk aversion estimates for the British pound, it is not clear whether the currency trade generates a larger or smaller risk premium at times when the US dollar/lb market is riskier. The risk aversion parameter is estimated to be positive but marginally significant throughout the sample period for the Swiss franc, implying that the currency trade generally yields a larger risk premium at times when the US dollar/Swiss franc market is riskier. For most of the sample, the risk-return coefficients are estimated to be positive but statistically insignificant for the Canadian dollar, suggesting that the intertemporal relation between risk and return is flat for the US/Canadian dollar market.

Appendix 1: Maximum Likelihood Estimation of GARCH-in-Mean Models

Modeling and estimating the volatility of financial time series has been high on the agenda of financial economists since the early 1980s. Engle (1982) put forward the Autoregressive Conditional Heteroskedastic (ARCH) class of models for conditional variances which proved to be extremely useful for analyzing financial return series. Since then an extensive literature has been developed for modeling the conditional distribution of stock prices, interest rates, exchange rates, and futures prices. Following the introduction of ARCH models by Engle (1982) and their generalization by Bollerslev (1986), there have been numerous refinements of this approach to estimating conditional volatility. Most of the refinements have been driven by empirical regularities in financial data.

Engle (1982) introduces *ARCH*(*p*) model:

$$R_{t+1} \equiv \alpha + \varepsilon_{t+1} \tag{40.22}$$

$$\begin{aligned} &\varepsilon_{t+1} = z_{t+1} \cdot \sigma_{t+1|t}, z_{t+1} \sim N(0,1), \\ &E(\varepsilon_{t+1}) = 0 \end{aligned} \tag{40.23}$$

$$E\left(\varepsilon_{t+1}^2|\Omega_t\right) = \sigma_{t+1|t}^2 = \gamma_0 + \gamma_1\varepsilon_t^2 + \gamma_2\varepsilon_{t-1}^2 + \ldots + \gamma_p\varepsilon_{t-p}^2 \tag{40.24}$$

$$f\left(R_{t+1};\mu,\sigma_{t+1|t}\right)=\frac{1}{\sqrt{2\pi\sigma^2_{t+1|t}}}\exp\left[-\frac{1}{2}\left(\frac{R_{t+1}-\mu}{\sigma_{t+1|t}}\right)^2\right] \tag{40.25}$$

where R_{t+1} is the daily return for period t+1, $\mu = \alpha$ is the constant conditional mean, $\varepsilon_{t+1} = z_{t+1} \cdot \sigma_{t+1|t}$ is the error term with $E(\varepsilon_{t+1}) = 0$, $\sigma_{t+1|t}$ is the conditional standard deviation of daily returns, and $z_{t+1} \sim N(0,1)$ is a random variable drawn from the standard normal density and can be viewed as information shocks or unexpected news in the market. ${\sigma_{t+1|t}}^2$ is the conditional variance of daily returns based on the information set up to time t denoted by Ω_t. The conditional variance, ${\sigma_{t+1|t}}^2$, follows an ARCH(p) process which is a function of the last period's unexpected news (or information shocks). $f(R_{t+1};\mu,\sigma_{t+1|t})$ is the conditional normal density function of R_{t+1} with the conditional mean of μ and conditional variance of ${\sigma_{t+1|t}}^2$.

Given the initial values of ε_t and, the parameters in Eqs 40.22 and 40.24 can be estimated by maximizing the log-likelihood function over the sample period. The conditional normal density in Eq. 40.25 yields the following log-likelihood function:

$$\text{Log}L_{ARCH}=-\frac{n}{2}\ln(2\pi)-\frac{n}{2}\ln\sigma_{t+1|t}-\frac{1}{2}\sum_{t=1}^{n}\left(\frac{R_{t+1}-\mu}{\sigma_{t+1|t}}\right)^2 \tag{40.26}$$

Bollerslev (1986) extends the original work of Engle (1982) and defines the current conditional variance as a function of the last period's unexpected news as well as the last period's conditional volatility:

$$R_{t+1}\equiv\alpha+\varepsilon_{t+1} \tag{40.27}$$

$$\varepsilon_{t+1}=z_{t+1}\cdot\sigma_{t+1|t},z_{t+1}\sim N(0,1)E(\varepsilon_{t+1})=0 \tag{40.28}$$

$$E\left(\varepsilon^2_{t+1}|\Omega_t\right)=\sigma^2_{t+1|t}=\gamma_0+\gamma_1\varepsilon^2_t+\gamma_2\sigma^2_t \tag{40.29}$$

where the conditional variance, ${\sigma_{t+1|t}}^2$, in Eq. 40.29 follows a GARCH(1,1) process as defined by Bollerslev (1986) to be a function of the last period's unexpected news (or information shocks), z_t, and the last period's variance, σ_t^2. The parameters in Eqs. 40.27, 40.28, and 40.29 are estimated by maximizing the conditional log-likelihood function in Eq. 40.26.

Engle et al. (1987) introduce the ARCH-in-mean model in which the conditional mean of financial time series is defined as a function of the conditional variance. In our empirical investigation of the ICAPM for exchange rates, we use the following GARCH-in-mean process to model the intertemporal relation between expected return and risk on currency

$$R_{t+1}\equiv\alpha+\beta\cdot\sigma^2_{t+1|t}+\varepsilon_{t+1} \tag{40.30}$$

$$\begin{aligned}&\varepsilon_{t+1}=z_{t+1}\cdot\sigma_{t+1|t};z_{t+1}\sim N(0,1);\\&E(\varepsilon_{t+1})=0\end{aligned} \tag{40.31}$$

$$E\left(\varepsilon^2_{t+1}|\Omega_t\right)=\sigma^2_{t+1|t}=\gamma_0+\gamma_1\varepsilon^2_t+\gamma_2\sigma^2_t \tag{40.32}$$

$$f\left(R_{t+1};\mu_{t+1|t},\sigma_{t+1|t}\right) = \frac{1}{\sqrt{2\pi\sigma^2_{t+1|t}}}\exp\left[-\frac{1}{2}\left(\frac{R_{t+1}-\mu_{t+1|t}}{\sigma_{t+1|t}}\right)^2\right] \tag{40.33}$$

where R_{t+1} is the daily return on exchange rates for period t+1, $\mu_{t+1|t} \equiv \alpha + \beta \cdot \sigma^2_{t+1|t}$ is the conditional mean for period t+1 based on the information set up to time t, $\varepsilon_{t+1} = z_{t+1} \cdot \sigma_{t+1|t}$ is the error term with $E(\varepsilon_{t+1}) = 0$, $\sigma_{t+1|t}$ is the conditional standard deviation of daily returns on currency, and $z_{t+1} \sim N(0,1)$ is a random variable drawn from the standard normal density and can be viewed as information shocks in the FX market. $\sigma^2_{t+1|t}$ is the conditional variance of daily returns based on the information set up to time t denoted by Ω_t. The conditional variance, $\sigma^2_{t+1|t}$, follows a GARCH(1,1) process as defined by Bollerslev (1986) to be a function of the last period's unexpected news (or information shocks), z_t, and the last period's variance, σ_t^2. $f(R_{t+1};\mu_{t+1|t},\sigma_{t+1|t})$ is the conditional normal density function of R_{t+1} with the conditional mean of $\mu_{t+1|t}$ and conditional variance of $\sigma^2_{t+1|t}$.

Given the initial values of ε_t and, the parameters in Eqs. 40.30 and 40.32 can be estimated by maximizing the log-likelihood function over the sample period. The conditional normal density in Eq. 40.33 yields the following log-likelihood function

$$\begin{aligned}\text{Log}L_{ARCH} = &-\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma_{t+1|t} \\ &-\frac{1}{2}\sum_{t=1}^{n}\left(\frac{R_{t+1}-\mu_{t+1|t}}{\sigma_{t+1|t}}\right)^2\end{aligned} \tag{40.34}$$

where the conditional mean $\mu_{t+1|t} \equiv \alpha + \beta \cdot \sigma^2_{t+1|t}$ has two parameters and the conditional variance $\sigma_{t+1|t}{}^2 = \gamma_0 + \gamma_1\varepsilon_t{}^2 + \gamma_2\sigma_t^2$ has three parameters. Maximizing the log-likelihood in Eq. 40.34 yields the parameter estimates (α, β, γ_0, γ_1, γ_2).

The interested reader may wish to consult Enders (2009), Chap. 3, and Tsay (2010), Chap. 3, for comprehensive analysis of ARCH/GARCH models and their maximum likelihood estimation. Chapter 3 in Enders (2009) provides a detailed coverage of the basic ARCH and GARCH models, as well as the GARCH-in-mean processes and multivariate GARCH in some detail. Chapter 3 in Tsay (2010) provides a detailed coverage of the ARCH, GARCH, GARCH-M, the exponential GARCH, and Threshold GARCH models.

Appendix 2: Estimation of a System of Regression Equations

Consider a system of n equations, of which the typical ith equation is

$$y_i = X_i\beta_i + u_i \tag{40.35}$$

where y_i is a $N \times 1$ vector of time-series observations on the i th dependent variable, X_i is a $N \times k_i$ matrix of observations of k_i independent variables, β_i is a $k_i \times 1$ vector of unknown coefficients to be estimated, and u_i is a $N \times 1$ vector of random disturbance terms with mean zero. Parks (1967) proposes an estimation procedure that allows the error term to be both serially and cross-sectionally correlated. In particular, he assumes that the elements of the disturbance vector u follow an AR(1) process

$$u_{it} = \rho_i u_{it-1} + \varepsilon_{it}; \rho_i < 1 \tag{40.36}$$

where ε_{it} is serially independently but contemporaneously correlated:

$$\mathrm{Cov}\left(\varepsilon_{it}\varepsilon_{jt}\right) = \sigma_{ij}, \forall i, j, \text{and } \mathrm{Cov}\left(\varepsilon_{it}\varepsilon_{js}\right) = 0, \text{for } s \neq t \tag{40.37}$$

Equation 40.35 can then be written as

$$y_i = X_i\beta_i + P_i\varepsilon_i, \tag{40.38}$$

with

$$P_i = \begin{bmatrix} \left(1-\rho_i^2\right)^{-1/2} & 0 & 0 & \dots & 0 \\ \rho_i\left(1-\rho_i^2\right)^{-1/2} & 1 & 0 & \dots & 0 \\ \rho_i^2\left(1-\rho_i^2\right)^{-1/2} & \rho_i & 0 & \dots & 0 \\ \cdot & & & & \\ \cdot & & & & \\ \cdot & & & & \\ \rho_i^{N-1}\left(1-\rho_i^2\right)^{-1/2} & \rho_i^{N-2} & \rho_i^{N-3} & \dots & 1 \end{bmatrix} \tag{40.39}$$

Under this setup, Parks presents a consistent and asymptotically efficient three-step estimation technique for the regression coefficients. The first step uses single equation regressions to estimate the parameters of autoregressive model. The second step uses single equation regressions on transformed equations to estimate the contemporaneous covariances. Finally, the Aitken estimator is formed using the estimated covariance,

$$\hat{\beta} = \left(X^T\Omega^{-1}X\right)^{-1}X^T\Omega^{-1}y \tag{40.40}$$

Where $\Omega \equiv E[uu^T]$ denotes the general covariance matrix of the innovation. In my application, I use the aforementioned methodology with the slope coefficients restricted to be the same for all portfolios. In particular, we use the same three-step procedure and the same covariance assumptions as in Eqs. 40.36, 40.37, 40.38, 40.39, and 40.40 to estimate the covariances and to generate the t-statistics for the parameter estimates.

The interested reader may wish to consult Wooldridge (2010), Chaps. 10.4, 10.5, and 10.6 for recent developments on panel data estimation. Chapter 10 in

Wooldridge (2010) presents Basic Linear Unobserved Effects Panel Data Models, Chap. 10.4 provides Random Effects Methods, Chap. 10.5 contains Fixed Effects Methods, and Chap. 10.6 First Differencing Methods. Bali (2008) and Bali and Engle (2010) follow SUR estimation to investigate the empirical validity of the conditional intertemporal capital asset pricing models (ICAPM).

References

Abel, A. B. (1988). Stock prices under time-varying dividend risk: An exact solution in an infinite-horizon general equilibrium model. *Journal of Monetary Economics, 22*, 375–393.

Alizadeh, S., Brandt, M. W., & Diebold, F. X. (2002). Range-based estimation of stochastic volatility models. *Journal of Finance, 57*, 1047–1092.

Andersen, T. G., Bollerslev, T., & Diebold, F. X. (2004). Parametric and nonparametric volatility measurement. In Y. Aït-Sahalia & L. P. Hansen (Eds.), *Handbook of financial econometrics*. Amsterdam: North Holland.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Ebens, H. (2001a). The distribution of realized stock return volatility. *Journal of Financial Economics, 61*, 43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2001b). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association, 96*, 42–55.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica, 71*, 579–626.

Backus, D. K., & Gregory, A. W. (1993). Theoretical relations between risk premiums and conditional variance. *Journal of Business and Economic Statistics, 11*, 177–185.

Bali, T. G. (2008). The intertemporal relation between expected returns and risk. *Journal of Financial Economics, 87*, 101–131.

Bali, T. G., & Engle, R. F. (2010). The intertemporal capital asset pricing model with dynamic conditional correlations. *Journal of Monetary Economics, 57*, 377–390.

Bali, T. G., & Peng, L. (2006). Is there a risk-return tradeoff? Evidence from high frequency data. *Journal of Applied Econometrics, 21*, 1169–1198.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics, 31*, 307–327.

Bollerslev, T., & Domowitz, I. (1993). Trading patterns and prices in the interbank foreign exchange market. *Journal of Finance, 48*, 1421–1443.

Bollerslev, T., Engle, R. F., & Wooldridge, J. M. (1988). A capital asset pricing model with time-varying covariances. *Journal of Political Economy, 96*, 116–131.

Bollerslev, T., & Wooldridge, J. M. (1992). Quasi-maximum likelihood estimation and inference in dynamic models with time varying covariances. *Econometric Reviews, 11*, 143–172.

Bollerslev, T., & Zhou, H. (2006). Volatility puzzles: A simple framework for gauging return-volatility regressions. *Journal of Econometrics, 131*, 123–150.

Brandt, M. W., & Diebold, F. X. (2006). A no-arbitrage approach to range-based estimation of return covariances and correlations. *Journal of Business, 79*, 61–73.

Brandt, M. W., & Kang, Q. (2004). On the relationship between the conditional mean and volatility of stock returns: A latent VAR approach. *Journal of Financial Economics, 72*, 217–257.

Breen, W., Glosten, L. R., & Jagannathan, R. (1989). Economic significance of predictable variations in stock index returns. *Journal of Finance, 44*, 1177–1189.

Campbell, J. Y. (1987). Stock returns and the term structure. *Journal of Financial Economics, 18*, 373–399.

Campbell, J. Y. (1991). A variance decomposition for stock returns. *Economic Journal, 101*, 157–179.

Campbell, J. Y., & Hentchel, L. (1992). No news is good news: An asymmetric model of changing volatility in stock returns. *Journal of Financial Economics, 31*, 281–318.

Campbell, J. Y., & Shiller, R. (1988). The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies, 1*, 195–228.

Chan, K. C., Karolyi, G. A., & Stulz, R. M. (1992). Global financial markets and the risk premium on U.S. equity. *Journal of Financial Economics, 32*, 137–167.

Chen, N.-F., Roll, R., & Ross, S. A. (1986). Economic forces and the stock market. *Journal of Business, 59*, 383–403.

Chou, R., Engle, R. F., & Kane, A. (1992). Measuring risk aversion from excess returns on a stock index. *Journal of Econometrics, 52*, 201–224.

Christoffersen, P. F., & Diebold, F. X. (2006). Financial asset returns, direction-of-change, forecasting, and volatility dynamics. *Management Science, 52*, 1273–1287.

Enders, W. (2009). *Applied econometric time series* (3rd ed.). New York: Wiley.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica, 50*, 987–1007.

Engle, R. F., Lilien, D. M., & Robins, R. P. (1987). Estimating time varying risk premia in the term structure: The ARCH M model. *Econometrica, 55*, 391–407.

Fama, E. F., & French, K. R. (1988). Dividend yields and expected stock returns. *Journal of Financial Economics, 22*, 3–25.

Fama, E. F., & French, K. R. (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics, 25*, 23–49.

French, K. R., Schwert, G. W., & Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of Financial Economics, 19*, 3–29.

Gennotte, G., & Marsh, T. A. (1993). Variations in economic uncertainty and risk premiums on capital assets. *European Economic Review, 37*, 1021–1041.

Ghysels, E., Santa-Clara, P., & Valkanov, R. (2005). There is a risk-return tradeoff after all. *Journal of Financial Economics, 76*, 509–548.

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess returns on stocks. *Journal of Finance, 48*, 1779–1801.

Goyal, A., & Santa-Clara, P. (2003). Idiosyncratic risk matters! *Journal of Finance, 58,* 975–1008.

Guo, H., & Whitelaw, R. F. (2006). Uncovering the risk-return relation in the stock market. *Journal of Finance, 61*, 1433–1463.

Harrison, P., & Zhang, H. (1999). An investigation of the risk and return relation at long horizons. *Review of Economics and Statistics, 81*, 399–408.

Harvey, C. R. (2001). The specification of conditional expectations. *Journal of Empirical Finance, 8*, 573–638.

Jegadeesh, N. (1990). Evidence of predictable behavior of security returns. *Journal of Finance, 45*, 881–898.

Keim, D. B., & Stambaugh, R. F. (1986). Predicting returns in the stock and bond markets. *Journal of Financial Economics, 17*, 357–390.

Lehmann, B. (1990). Fads, martingales, and market efficiency. *Quarterly Journal of Economics, 105*, 1–28.

Lettau, M., & Ludvigson, S. C. (2010). Measuring and modeling variation in the risk-return tradeoff. In *Handbook of Financial Econometrics* by Y. Ait-Sahalia & L. P. Hansen (Eds.), Vol 1, pp. 617–690. Amsterdam: North Holland.

Lo, A. W., & MacKinlay, A. C. (1990). When are contrarian profits due to stock market overreaction? *Review of Financial Studies, 3*, 175–205.

Lundblad, C. (2007). The risk return tradeoff in the long run: 1836–2003. *Journal of Financial Economics, 85*, 123–150.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica, 41*, 867–887.

Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics, 8*, 323–361.

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica, 59*, 347–370.

Newey, W. K., & West, K. D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica, 55*, 703–708.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business, 53*, 61–65.

Parks, R. W. (1967). Efficient estimation of a system of regression equations when disturbances are both serially and contemporaneously correlated. *Journal of the American Statistical Association, 62*, 500–509.

Scruggs, J. T. (1998). Resolving the puzzling intertemporal relation between the market risk premium and conditional market variance: A two-factor approach. *Journal of Finance, 53*, 575–603.

Sentana, E. (1995). Quadratic ARCH models. *Review of Economic Studies, 62*, 639–661.

Tsay, R. S. (2010). *Analysis of financial time series*. New York: Wiley.

Turner, C. M., Startz, R., & Nelson, C. R. (1989). A markov model of heteroskedasticity, risk, and learning in the stock market. *Journal of Financial Economics, 25*, 3–22.

Wooldridge, J. M. (2010). *Econometric analysis of cross section and panel data* (2nd ed.). Boston: MIT Press.

Quantile Regression and Value at Risk

41

Zhijie Xiao, Hongtao Guo, and Miranda S. Lam

Contents

Abstract

This paper studies quantile regression (QR) estimation of Value at Risk (VaR). VaRs estimated by the QR method display some nice properties. In this paper, different QR models in estimating VaRs are introduced. In particular, VaR

We thank Prof. C.-F. Lee for helpful comments. This project is partially supported by Boston College Research fund.

Z. Xiao (✉)
Department of Economics, Boston College, Chestnut Hill, MA, USA
e-mail: xiaoz@bc.edu

H. Guo • M.S. Lam
Bertolon School of Business, Salem State University, Salem, MA, USA
e-mail: hguo@salemstate.edu; miranda.lam@salemstate.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_41,

estimations based on quantile regression of the QAR models, copula models, ARCH models, GARCH models, and the CaViaR models are systematically introduced. Comparing the proposed QR method with traditional methods based on distributional assumptions, the QR method has the important property in that it is robust to non-Gaussian distributions. Quantile estimation is only influenced by the local behavior of the conditional distribution of the response near the specified quantile. As a result, the estimates are not sensitive to outlier observations. Such a property is especially attractive in financial applications since many financial data like, say, portfolio returns (or log returns) are usually not normally distributed. To highlight the importance of the QR method in estimating VaR, we apply the QR techniques to estimate VaRs in International Equity Markets. Numerical evidence indicates that QR is a robust estimation method for VaR.

Keywords

ARCH • Copula • GARCH • Non-normality • QAR • Quantile regression • Risk management • Robust estimation • Time series • Value at risk

41.1 Introduction

The Value at Risk (VaR) is the loss in market value over a given time horizon that is exceeded with probability τ, where τ is often set at 0.01 or 0.05. In recent years, VaR has become a popular tool in the measurement and management of financial risk. This popularity is spurred both by the need of various institutions for managing risk and by government regulations (see Blankley et al., 2000; Dowd 1998, 2000; Saunders 1999). VaR is an easily interpretable measure of risk that summarizes information regarding the distribution of potential losses. In requiring publicly traded firms to report risk exposure, the Securities and Exchange Commission (SEC) lists VaR as a disclosure method "expressing the potential loss in future earnings, fair values, or cash flows from market movements over a selected period of time and with a selected likelihood of occurrence."

Estimation of VaR has attracted much attention from researchers (Duffie and Pan (1997); Wu and Xiao (2002); Guo et al. (2007)). Many existing methods of VaR estimation in economics and finance are based on the assumption that financial returns have normal (or conditional normal) distributions (usually with ARCH or GARCH effects). Under the assumption of a conditionally normal return distribution, the estimation of conditional quantiles is equivalent to estimating conditional volatility of returns. The massive literature on volatility modeling offers a rich source of parametric methods of this type. However, there is accumulating evidence that financial time series and return distributions are *not* well approximated by Gaussian models. In particular, it is frequently found that market returns display negative skewness and excess kurtosis. Extreme realizations of returns can adversely affect the performance of estimation and inference designed for Gaussian conditions; this is particularly true of ARCH and GARCH models whose estimation of variances is very sensitive to large innovations. For this reason, research attention has recently shifted toward the development of more robust estimators of conditional quantiles.

There is also growing interest in nonparametric estimation of conditional quantiles. However, nearest neighbor and kernel methods are somewhat limited in their ability to cope with more than one or two covariates. Other approaches to estimating VaR include the hybrid method and methods based on extreme value theory.

Quantile regression as introduced by Koenker and Bassett (1978) is well suited to estimating VaR. Value at Risk, as mandated in many current regulatory contexts, is a conditional quantile by definition. This concept is intimately linked to quantile regression estimation.

Quantile regression has now become a popular robust approach for statistical analysis. Just as classical linear regression methods based on minimizing sums of squared residuals enable one to estimate models for conditional mean, quantile regression methods offer a mechanism for estimating models for the conditional quantiles. These methods exhibit robustness to extreme shocks and facilitate distribution-free inference. In recent years, quantile regression estimation for time-series models has gradually attracted more attention. Koenker and Zhao (1996) extended quantile regression to linear ARCH models and estimate conditional quantiles by a linear quantile regression. Engle and Manganelli (1999) have suggested a nonlinear dynamic quantile model where conditional quantiles themselves follow an autoregression, and they call this a Conditional Autoregressive Value at Risk (CaViaR) specification. Computation of the CaViaR model is challenging and grid searching is conventionally used in practice. Koenker and Xiao (2006) investigate quantile autoregressive processes that can capture systematic influences of conditioning variables on the location, scale, and shape of the conditional distribution of the response and therefore constitute a significant extension of classical time-series models in which the effect of conditioning is confined to a location shift. Xiao and Koenker (2009) recently studied quantile regression estimation of GARCH models. GARCH models have proven to be highly successful in modeling financial data and are arguably the most widely used class of models in financial applications. However, quantile regression GARCH models are highly nonlinear and thus complicated to estimate. The quantile estimation problem in GARCH models corresponds to a restricted nonlinear quantile regression, and conventional nonlinear quantile regression techniques are not directly applicable, adding an additional challenge to the already complicated estimation problem. Koenker and Xiao (2009) propose a two-step approach for quantile regression on GARCH models. The proposed method is relatively easy to implement compared to other nonlinear estimation techniques in quantile regression and has good sampling performance in our simulation experiments.

VaRs estimated by the quantile regression approach display some nice properties. For example, they track VaRs estimated from GARCH volatility models well during normal market conditions. However, during market turmoils when large market price drops are followed by either further drops or rebounds, GARCH volatility models tend to predict implausibly high VaRs. This is due to the fact that volatility and VaRs are not synonymous. While large positive and negative return shocks indicate higher volatility, only large negative return shocks indicate higher Value at Risk. GARCH models treat both large positive and negative return shocks as indications of higher volatility. VaRs estimated by the ARCH/GARCH quantile regression model, while predicting higher volatility in the ARCH/GARCH component, assign a much bigger weight to the large negative return shock than the large positive return shock. The resulting estimated VaRs seem to be closer to reality.

In this chapter, we study quantile regression estimation of VaR. Different quantile models in estimating VaR are introduced in this paper. In particular, Value at Risk analysis based on quantile regression of the QAR models, copula models, ARCH models, GARCH models, and the CaViaR models is systematically introduced. To highlight the importance of quantile regression method in estimating VaR, we apply the quantile regression techniques to estimate VaR in International Equity Markets. Numerical evidence indicates that quantile regression is a robust estimation method for VaR.

This chapter is organized as follows. We introduce the traditional VaR estimation methods and quantile regression in Sect. 41.2. The quantile autoregression (QAR) models are given in Sect. 41.3; nonlinear QAR models based on copula and the CaViaR models are also introduced. Section 41.4 introduces quantile regression estimation on ARCH and GARCH models. Section 41.5 contains an empirical application of quantile regression estimation of VaRs. Section 41.6 concludes.

41.2 Traditional Estimation Methods of VaR

For a time series of returns on an asset, $\{r_t\}_{t=1}^{n}$, the τ (or $100\tau\%$) VaR at time t, denoted by VaR_t, is defined by

$$\Pr(r_t < -VaR_t | \mathcal{F}_{t-1}) = \tau \tag{41.1}$$

where $\mathcal{F}_{t-1}$ denotes the information set at time $t-1$, including past values of returns and possibly the value of some covariates X_t.

If we assume that the time series of returns are modeled by

$$r_t = \mu_t + \sigma_t \varepsilon_t,$$

where r_t is the return of an asset at time t and $\mu_t, \sigma_t \in \mathcal{F}_{t-1}$. The random variables ε_t are martingale difference sequences. The Conditional Value at Risk of r_t given $\mathcal{F}_{t-1}$ is

$$VaR_t(\tau) = \mu_t + \sigma_t Q_\varepsilon(\tau),$$

where $Q_\varepsilon(\tau)$ denotes the Unconditional Value at Risk of the error term ε_t. Assuming conditional normality, the 5 % VaR at time t can be computed as

$$VaR_t(0.05) = \mu_t + 1.65\sigma_t$$

where μ_t and σ_t are the conditional mean and conditional volatility for r_t.

RiskMetrics takes a simple and pragmatic approach to modeling the conditional volatility. The forecast for time t variance in RiskMetrics method is a weighted average of the previous forecast, using weight λ, and of the latest squared innovation, using weight $(1-\lambda)$:

$$\sigma_t^2 = \lambda\sigma_{t-1}^2 + (1-\lambda)r_{t-1}^2 \tag{41.2}$$

where the parameter λ is called the decay factor $(1 > \lambda > 0)$. Conceptually λ should be estimated using a maximum likelihood approach. RiskMetrics simply set it optimally at 0.94 for daily data and 0.97 for monthly data. Our analysis is on weekly data and we set λ at 0.95.

There are extensive empirical evidences supporting the use of ARCH and GARCH models in conditional volatility estimation. Bollerslev et al. (1992) provide a nice overview of the issue. Sarma et al. (2000) showed that at the 5 % level, an AR(1)-GARCH(1,1) model is a preferred model under the conditional normality assumption. The AR(1)-GARCH(1,1) model is specified:

$$\begin{aligned} r_{t+1} &= a_0 + a_1 r_t + \in_{t+1} \qquad\qquad \in_{t+1}\big|I_t \sim N\left(0, \sigma_t^2\right) \\ \sigma_t^2 &= \omega_0 + \omega_1\sigma_{t-1}^2 + \omega_2\in_t^2. \end{aligned} \tag{41.3}$$

The conditional mean equation is modeled as an AR(1) process to account for the weakly autoregressive behavior of returns.

41.3 Quantile Regression

Quantile regression was introduced by Koenker and Bassett (1978) and has received a lot of attention in econometrics and statistics research in the past two decades. The *quantile function* of a scalar random variable Y is the inverse of its distribution function. Similarly, the *conditional quantile function* of Y given X is the inverse of the corresponding conditional distribution function, i.e.,

$$Q_Y\left(\tau|X\right) = F_Y^{-1}(\tau|X) = \inf\{y : F_Y(y|X) \geq \tau\},$$

where $F_Y(y|X) = P(Y \leq y|X)$. By definition, the τ VaR at time t is the τ-th conditional quantile of r_t giving information at time $t-1$.

Consider a random variable Y characterized by its distribution function $F(y)$, the τ-th quantile of Y is defined by

$$Q_Y(\tau) = \inf\{y|F(y) \geq \tau\}.$$

If we have a random sample $\{y_1, \ldots, y_n\}$ from the distribution F, the τ-th sample quantile can be defined as

$$\hat{Q}_Y(\tau) = \inf\{y | \hat{F}(y) \geq \tau\},$$

where $\hat{F}$ is the empirical distribution function of the random sample. Note that the above sample quantile may be found by solving the following minimization problem:

$$\min_{b \in \Re} \left[\sum_{t \in \{t: y_t \geq b\}} \tau |y_t - b| + \sum_{t \in \{t: y_t < b\}} (1 - \tau)|y_t - b| \right]. \tag{41.4}$$

Koenker and Bassett (1978) studied the analogue of the empirical quantile function for the linear models and generalized the concept of quantiles to the regression context.

If we consider the linear regression model

$$Y_t = \beta' X_t + u_t, \tag{41.5}$$

where u_t are iid mean zero random variables with quantile function $Q_u(\tau)$ and X_t are k-by-1 vector of regressors including an intercept term and lagged residuals, then, conditional on the regressor X_t, the τ-th quantile of Y is a linear function of X_t:

$$Q_{Y_t}(\tau | X_t) = \beta' X_t + Q_u(\tau) = \beta(\tau)' X_t$$

where $\beta(\tau)' = (\beta_1 + Q_u(\tau), \beta_2, \cdots, \beta_k)$. Koenker and Bassett (1978) show that the τ-th conditional quantile of Y can be estimated by an analogue of Eq. 41.4:

$$\hat{Q}_{Y_t}(\tau | X_t) = X_t' \hat{\beta}(\tau)$$

where

$$\hat{\beta}(\tau) = \arg\min_{\beta \in \Re^k} \left[\sum_{t \in \{t: y_t \geq x_t \beta\}} \tau |y_t - x_t' \beta| + \sum_{t \in \{t: y_t < x_t \beta\}} (1 - \tau)|y_t - x_t' \beta| \right] \tag{41.6}$$

is called as the regression quantiles. Let $\rho_\tau(u) = u(\tau - I(u < 0))$, then

$$\hat{\beta}(\tau) = \arg\min_{\beta \in \Re^k} \sum_t \rho_\tau \left(y_t - x_t' \beta \right).$$

Quantile regression method has the important property that it is robust to distributional assumptions. This property is inherited from the robustness property of the ordinary sample quantiles. Quantile estimation is only influenced by the local behavior of the conditional distribution of the response near the specified quantile. As a result, the estimated coefficient vector $\hat{y}(\tau)$ is not sensitive to outlier observations. Such a property is especially attractive in financial applications since many financial data like, say, portfolio returns (or log returns) are usually heavy tailed and thus not normally distributed.

The quantile regression model has a mathematical programming representation which facilitates the estimation. Notice that the optimization problem (Eq. 41.6) may be reformulated as a linear program by introducing "slack" variables to represent the positive and negative parts of the vector of residuals (see Koenker and Bassett (1978) for a more detailed discussion). Computation of the regression quantiles by standard linear programming techniques is very efficient. It is also straightforward to impose the nonnegativity constraints on all elements of γ. Barrodale and Roberts (1974) proposed the first efficient algorithm for L_1-estimation problems based on modified simplex method. Koenker and d'Orey (1987) modified this algorithm to solve quantile regression problems. For very large quantile regression problems, there are some important new ideas which speed up the performance of computation relative to the simplex approach underlying the original code. Portnoy and Koenker (1997) describe an approach that combines some statistical preprocessing with interior point methods and achieves faster speed over the simplex method for very large problems.

41.4 Autoregressive Quantile Regression Models

41.4.1 The QAR Models

In many finance applications, the time-series dynamics can be more complicated than the classical autoregression where past information (Y_{t-j}) influences only the location of the conditional distribution of Y_t. For example, it is well known that the correlations tend to be larger in bear than in bull markets. Recognizing that the correlation is asymmetric is important for risk management and other financial applications. Any attempt to diagnose or forecast series of this type requires that a mechanism be introduced to capture the empirical features of the series.

An important extension of the classical constant coefficient time-series model is the quantile autoregression (QAR) model (Koenker and Xiao 2006). Given a time series $\{Y_t\}$, let $\mathcal{F}_t$ be the σ-field generated by $\{Y_s, s \leq t\}$; $\{Y_t\}$ is a p-th order QAR process if

$$Q_{Y_t}(\tau|\mathcal{F}_{t-1}) = \theta_0(\tau) + \theta_1(\tau)Y_{t-1} + \cdots + \theta_p(\tau)Y_{t-p}; \tag{41.7}$$

this implies, of course, that the right-hand side of Eq. 41.7 is monotonically increasing in τ. In the above QAR model, the autoregressive coefficients may be τ-dependent and thus can vary over different quantiles of the conditional distribution. Consequently, the conditioning variables not only shift the location of the distribution of Y_t but also may alter the scale and shape of the conditional distribution. The QAR models play a useful role in expanding the modeling territory of the classical autoregressive time-series models, and the classical AR(p) model can be viewed as a special case of QAR by setting $\theta_j(\tau)$ ($j = 1, \ldots, p$) to constants.

Koenker and Xiao (2006) studied the QAR model. The QAR model can be estimated by

$$\hat{\theta}(\tau) = \min_{\theta} \sum_{t} \rho_{\tau}\left(Y_t - \theta^{\top} X_t\right), \tag{41.8}$$

where $X_t = (1, Y_{t-1}, \ldots, Y_{t-p})^{\top}$ and $\theta(\tau) = (\theta_0(\tau), \theta_1(\tau), \ldots, \theta_p(\tau))^{\top}$; they show that under regularity assumptions, the limiting distribution of the QAR estimator is given by

$$\sqrt{n}\left(\hat{\theta}(\tau) - \theta(\tau)\right) \Rightarrow \mathcal{N}\left(0, \tau(1-\tau)\Omega_1^{-1}\Omega_0\Omega_1^{-1}\right),$$

where $\Omega_0 = E(X_t X_t^{\top})$ and $\Omega_1 = \lim n^{-1} \sum_{t=1}^{n} f_{t-1}[F_{t-1}^{-1}(\tau)]X_t X_t^{\top}$.

The QAR models expand the modeling options for time series that display asymmetric dynamics and allow for local persistency. The models can capture systematic influences of conditioning variables on the location, scale, and shape of the conditional distribution of the response and therefore constitute a significant extension of classical constant coefficient linear time-series models.

Quantile varying coefficients indicate the existence of conditional heteroskedasticity. Given the QAR process (Eq. 41.7), let $\theta_0 = \mathrm{E}[\theta_0(U_t)]$, $\theta_1 = \mathrm{E}[\theta_1(U_t)], \ldots,$ $\theta_p = \mathrm{E}[\theta_p(U_t)]$, and

$$V_t = \theta_0(U_t) - \mathrm{E}\theta_0(U_t) + [\theta_1(U_t) - \mathrm{E}\theta_1(U_t)]Y_{t-1} + \cdots + \left[\theta_p(U_t) - \mathrm{E}\theta_p(U_t)\right]Y_{t-p};$$

the QAR process can be rewritten as

$$Y_t = \theta_0 + \theta_1 Y_{t-1} + \cdots + \theta_p Y_{t-p} + V_t \tag{41.9}$$

where V_t is martingale difference sequence. The QAR process is a weak sense AR process with conditional heteroskedasticity.

What's the difference between a QAR process and an AR process with ARCH (or GARCH) errors? In short, the ARCH type model only focuses on the first two moments, while the QAR model goes beyond the second moment and allows for more flexible structure in higher moments. Both models allow for conditional heteroskedasticity and they are similar in the first two moments, but they can be quite different beyond conditional variance.

41.4.2 Nonlinear QAR and Copula-Based Quantile Models

More complicated functional forms with nonlinearity can be considered for the conditional quantile function if we are interested in the global behavior of the time series. If the τ-th conditional quantile function of Y_t is given by

$$Q_{Y_t}(\tau|\mathcal{F}_{t-1}) = H(X_t;\, \theta(\tau)),$$

where X_t is the vector containing lagged Ys, we may estimate the vector of parameters $\theta(\tau)$ (and thus the conditional quantile of Y_t) by the following nonlinear quantile regression:

$$\min_\theta \sum_t \rho_\tau(Y_t - H(X_t,\, \theta)). \tag{41.10}$$

Let $\varepsilon_{t\tau} = y_t - H(x_t, \theta(\tau)), \dot{H}_\theta(x_t, \theta) = \partial H(x_t; \theta)/\partial\theta$; we assume that

$$V_n(\tau) = \frac{1}{n}\sum_t f_t\big(Q_{Y_t}(\tau|X_t)\big)\dot{H}_\theta(X_t, \theta(\tau))\dot{H}_\theta(X_t, \theta(\tau))^\top \xrightarrow{P} V(\tau),$$

$$\Omega_n(\tau) = \frac{1}{n}\sum_t \dot{H}_\theta(X_t, \theta(\tau))\dot{H}_\theta(X_t, \theta(\tau))^\top \xrightarrow{P} \Omega(\tau),$$

and

$$\frac{1}{\sqrt{n}}\sum_t \dot{H}_\theta(x_t,\, \theta(\tau))Y_\tau(\varepsilon_{t\tau}) \Rightarrow N(0,\;\; \tau(1-\tau)\Omega(\tau)),$$

where $V(\tau)$ and $\Omega(\tau)$ are non-singular; then under appropriate assumptions, the nonlinear QAR estimator $\hat{\theta}(\tau)$ defined as solution of Eq. 41.10 is root-n consistent and

$$\sqrt{n}\Big(\hat{\theta}(\tau) - \theta(\tau)\Big) \Rightarrow N\Big(0, \tau(1-\tau)V(\tau)^{-1}\Omega(\tau)V(\tau)^{-1}\Big). \tag{41.11}$$

In practice, one may employ parametric copula models to generate nonlinear-in-parameters QAR models (see, e.g., Bouyé and Salmon 2008; Chen et al. 2009). Copula-based Markov models provide a rich source of potential nonlinear dynamics describing temporal dependence and tail dependence. If we consider, for example, a first-order strictly stationary Markov process, $\{Y_t\}_{t=1}^n$, whose probabilistic properties are determined by the joint distribution of Y_{t-1} and Y_t, say, $G^*(y_{t-1}, y_t)$, and suppose that $G^*(y_{t-1}, y_t)$ has continuous marginal distribution function $F^*(\cdot)$, then by Sklar's Theorem, there exists a unique copula function $C^*(\cdot, \cdot)$ such that

$$G^*(y_{t-1}, y_t) \equiv C^*(F^*(y_{t-1}),\;\; F^*(y_t)),$$

where the copula function $C^*(\cdot,\cdot)$ is a bivariate probability distribution function with uniform marginals. Differentiating $C^*(u, v)$ with respect to u and evaluating at $u = F^*(x)$, $v = F^*(y)$, we obtain the conditional distribution of Y_t given $Y_{t-1} = x$:

$$\Pr[Y_t < y | Y_{t-1} = x] = \left.\frac{\partial C^*(u,v)}{\partial u}\right|_{u=F^*(x),\, v=F^*(y)} \equiv C_1^*(F^*(x), F^*(y)).$$

For any $\tau \in (0, 1)$, solving $\tau = \Pr[Y_t < y | Y_{t-1} = x] \equiv C_1^*(F^*(x), F^*(y))$ for y (in terms of τ), we obtain the τ-th conditional quantile function of Y_t given $Y_{t-1} = x$:

$$Q_{Y_t}(\tau|x) = F^{*-1}\left(C_1^{*-1}(\tau; F^*(x))\right),$$

where $F^{*-1}(\cdot)$ signifies the inverse of $F^*(\cdot)$ and $C_1^{*-1}(\cdot;u)$ is the partial inverse of $C_1^*(u,v)$ with respect to $v = F^*(y_t)$.

In practice, neither the true copula function $C^*(\cdot,\cdot)$ nor the true marginal distribution function $F^*(\cdot)$ of $\{Y_t\}$ is known. If we model both parametrically by $C(\cdot,\cdot;\alpha)$ and $F(y;\beta)$, then the τ-th conditional quantile function of Y_t, $Q_{Y_t}(\tau|x)$, becomes a function of the unknown parameters α and β, i.e.,

$$Q_{Y_t}(\tau|x) = F^{-1}\left(C_1^{-1}(\tau; F(x,\beta), \alpha), \beta\right).$$

Denoting $\theta = (\alpha', \beta')'$ and $h(x, \alpha, \beta) \equiv C_1^{-1}(\tau; F(x, \beta), \alpha)$, we will write

$$Q_{Y_t}(\tau|x) = F^{-1}(h(x,\alpha,\beta),\beta) \equiv H(x;\theta). \tag{41.12}$$

For example, if we consider the Clayton copula:

$$C(u,v;\alpha) = [u^{-\alpha} + v^{-\alpha} - 1]^{-1/\alpha}, \text{ where } \alpha > 0,$$

one can easily verify that the τ-th conditional quantile function of U_t given u_{t-1} is

$$Q_{U_t}\left(\tau|u_{t-1}\right) = \left[\left(\tau^{-\alpha/(1+\alpha)} - 1\right)u_{t-1}^{-\alpha} + 1\right]^{-1/\alpha}.$$

See Bouyé and Salmon (2008) for additional examples of copula-based conditional quantile functions.

Although the quantile function specification in the above representation assumes the parameters to be identical across quantiles, we may permit the estimated parameters to vary with τ, thus extending the original copula-based QAR models to capture a wide range of systematic influences of conditioning variables on the conditional distribution of the response. By varying the choice of the copula specification, we can induce a wide variety of nonlinear QAR(1) dependence, and the choice of the marginal enables us to consider a wide range of possible tail behavior as well. In many financial time-series applications, the nature of the temporal dependence varies over the quantiles of the conditional distribution. Chen et al. (2009) studied asymptotic properties of the copula-based nonlinear quantile autoregression.

41.4.3 The CaViaR Models

Quantile-based method provides a local approach to directly model the dynamics of a time series at a specified quantile. Engle and Manganelli (2004) propose the Conditional Autoregressive Value at Risk (CaViaR) specification for the τ-th conditional quantile of u_t:

$$Q_{u_t}(\tau|\mathcal{F}_{t-1}) = \beta_0 + \sum_{i=1}^{p} \beta_i Q_{u_{t-i}}(\tau|\mathcal{F}_{t-i-1}) + \sum_{j=1}^{q} \alpha_j \ell\left(X_{t-j}\right), \tag{41.13}$$

where $X_{t-j} \in \mathcal{F}_{t-j}$, $\mathcal{F}_{t-j}$ is the information set at time $t-j$. A natural choice of X_{t-j} is the lagged u. When we choose $X_{t-j} = |u_{t-j}|$, we obtain GARCH-type CaViaR models:

$$Q_{u_t}(\tau|\mathcal{F}_{t-1}) = \beta_0 + \sum_{i=1}^{p} \beta_i Q_{u_{t-i}}(\tau|\mathcal{F}_{t-i-1}) + \sum_{j=1}^{q} \alpha_j \left|u_{t-j}\right|.$$

If $X_{t-j} = 0$, we obtain an autoregressive model for the VaRs:

$$Q_{u_t}(\tau|\mathcal{F}_{t-1}) = \beta_0 + \sum_{i=1}^{p} \beta_i Q_{u_{t-i}}(\tau|\mathcal{F}_{t-i-1}).$$

Engle and Manganelli (2004) discussed many choices of $\ell(X_{t-j})$, leading to different specifications of the CaViaR model.

41.5 Quantile Regression of Conditional Heteroskedastic Models

41.5.1 ARCH Quantile Regression Models

ARCH and GARCH models have proven to be highly successful in modeling financial data. Estimators of volatilities and quantiles based on ARCH and GARCH models are now widely used in finance applications. Consider the following linear ARCH(p) process:

$$u_t = \sigma_t \cdot \varepsilon_t, \sigma_t = \gamma_0 + \gamma_1 |u_{t-1}| + \cdots + \gamma_p \left|u_{t-p}\right|, \tag{41.14}$$

where $0 < \gamma_0 < \infty$, $\gamma_1, \ldots, \gamma_p \geq 0$ and ε_t are iid(0,1) random variables with pdf $f(\cdot)$ and CDF $F(\cdot)$. Let $Z_t = (1,|u_{t-1}|, \ldots,|u_{t-q}|)^\top$ and $\gamma(\tau) = (\gamma_0 F^{-1}(\tau), \gamma_1 F^{-1}(\tau), \ldots, \gamma_q F^{-1}(\tau))^\top$; the conditional quantiles of u_t is given by

$$Q_{u_t}(\tau|\mathcal{F}_{t-1}) = \gamma_0(\tau) + \gamma_1(\tau)|u_{t-1}| + \cdots + \gamma_p(\tau)\left|u_{t-p}\right| = \gamma(\tau)^\top Z_t$$

and can be estimated by the following linear quantile regression of u_t on Z_t:

$$\min_\gamma \sum_t \rho_\tau\left(u_t - \gamma^\top Z_t\right), \tag{41.15}$$

where $\gamma = (\gamma_0, \gamma_1, \cdots, \gamma_q)^\top$. The asymptotic behavior of the above quantile regression estimator is given by Koenker and Zhao (1996). In particular, suppose that u_t is given by model (Eq. 41.14), f is bounded and continuous, and $f(F^{-1}(t)) > 0$ for any $0 < \tau < 1$. In addition, if $\mathrm{E}|u_t|^{2+\delta} < \infty$, then the regression quantile $\hat{\gamma}(\tau)$ of Eq. 41.15 has the following Bahadur representation:

$$\sqrt{n}(\hat{\gamma}(\tau) - \gamma(\tau)) = \frac{\Sigma_1^{-1}}{f\left(F^{-1}(\tau)\right)} \frac{1}{\sqrt{n}} \sum_{t=1}^n Z_t^\top y_\tau(\varepsilon_{t\tau}) + o_p(1)$$

where $\Sigma_1 = \mathrm{E}Z_t Z_t'/\sigma_t$ and $\varepsilon_{t\tau} = \varepsilon_t - F^{-1}(\tau)$. Consequently,

$$\sqrt{n}(\hat{\gamma}(\tau) - \gamma(\tau)) = N\left(0,\ \frac{\tau(1-\tau)}{f\left(F^{-1}(\tau)\right)^2} \Sigma_1^{-1}\Sigma_0\Sigma_1^{-1}\right),\ \text{with}\ \Sigma_0 = \mathrm{E}Z_t Z_t'.$$

In many applications, conditional heteroskedasticity is modeled on the residuals of a regression. For example, we may consider the following AR-ARCH model:

$$Y_t = \alpha' X_t + u_t \tag{41.16}$$

where $X_t = (1, Y_{t-1}, \ldots,Y_{t-p})^\top$, $\alpha = (\alpha_0, \alpha_1, \ldots, \alpha_p)^\top$, and u_t is a linear ARCH(p) process given by model (Eq. 41.14). The conditional quantiles of Y_t is then given by

$$Q_{Y_t}(\tau|\mathcal{F}_{t-1}) = \alpha' X_t + \gamma(\tau)^\top Z_t. \tag{41.17}$$

One way to estimate the above model is to construct a *joint* estimation of α and $\gamma(\tau)$ based on *nonlinear* quantile regression. Alternatively, we may consider a two-step procedure that estimates α in the first step and then estimates $\gamma(\tau)$ based on the estimated residuals. The two-step procedure is usually less efficient because the preliminary estimation of α may affect the second-step estimation of $\gamma(\tau)$, but it is computationally much simpler and is widely used in empirical applications.

41.5.2 GARCH Quantile Regression Models

ARCH models are easier to estimate, but cannot parsimoniously capture the persistent influence of long past shocks comparing to the GARCH models. However, quantile regression GARCH models are highly nonlinear and thus complicated to estimate. In particular, the quantile estimation problem in GARCH models corresponds to a restricted nonlinear quantile regression, and conventional nonlinear quantile regression techniques are not directly applicable.

Xiao and Koenker (2009) studied quantile regression estimation of the following linear GARCH(p, q) model:

$$u_t = \sigma_t \cdot \varepsilon_t, \tag{41.18}$$

$$\sigma_t = \beta_0 + \beta_1\sigma_{t-1} + \cdots + \beta_p\sigma_{t-p} + \gamma_1|u_{t-1}| + \cdots + \gamma_q|u_{t-q}|. \tag{41.19}$$

Let $\mathcal{F}_{t-1}$ represents information up to time $t-1$; the τ-th conditional quantile of u_t is given by

$$Q_{u_t}(\tau|\mathcal{F}_{t-1}) = \theta(\tau)^{\top} Z_t, \tag{41.20}$$

where $Z_t = (1, \sigma_{t-1}, \ldots, \sigma_{t-p}, |u_{t-1}|, \ldots, |u_{t-q}|)^{\top}$ and $\theta(\tau)^{\top} = (\beta_0, \beta_1, \ldots, \beta_p, \gamma_1, \ldots, \gamma_q)F^{-1}(\tau)$.

Since Z_t contains $\sigma_{t-k}(k = 1, \cdots, q)$ which in turn depends on unknown parameters $\theta = (\beta_0, \beta_1, \ldots, \beta_p, \gamma_1, \ldots, \gamma_q)$, we may write Z_t as $Z_t(\theta)$ to emphasize the nonlinearity and its dependence on θ. If we use the following nonlinear quantile regression

$$\min_{\theta} \sum_t \rho_\tau\left(u_t - \theta^{\top} Z_t(\theta)\right), \tag{41.21}$$

for a fixed τ in isolation, consistent estimate of θ cannot be obtained since it ignores the global dependence of the σ_{t-k}'s on the entire function $\theta(\cdot)$. If the dependence structure of u_t is characterized by (1) and (1), we can consider the following restricted quantile regression instead of Eq. 41.21:

$$\left(\hat{\pi}, \hat{\theta}\right) = \begin{cases} \arg\min_{\pi,\theta} \sum_i \sum_t \rho_{\tau_i}\left(u_t - \pi_i^{\top} Z_t(\theta)\right) \\ \text{s.t. } \pi_i = \theta(\tau_i) = \theta F^{-1}(\tau_i). \end{cases}$$

Estimation of this global restricted nonlinear quantile regression is complicated. Xiao and Koenker (2009) propose a simpler two-stage estimator that both incorporates the global restrictions and also focuses on the local approximation around the specified quantile. The proposed estimation consists of the following two steps: (i) The first step considers a global estimation to incorporate the global dependence of the latent σ_{t-k}'s on θ. (ii) Then, using results from the first step, we focus on the specified quantile to find the best local estimate for the conditional quantile. Let

$$A(L) = 1 - \beta_1 L - \cdots - \beta_p L^p, \quad B(L) = \gamma_1 + \cdots + \gamma_q L^{q-1};$$

under regularity assumptions ensuring that $A(L)$ is invertible, we obtain an ARCH(∞) representation for σ_t:

$$\sigma_t = a_0 + \sum_{j=1}^{\infty} a_j \left| u_{t-j} \right|. \tag{41.22}$$

For identification, we normalize $a_0 = 1$. Substituting the above ARCH(∞) representation into (1) and (1), we have

$$u_t = \left(a_0 + \sum_{j=1}^{\infty} a_j \left| u_{t-j} \right| \right) \varepsilon_t, \tag{41.23}$$

and

$$Q_{u_t}(\tau | \mathcal{F}_{t-1}) = \alpha_0(\tau) + \sum_{j=1}^{\infty} \alpha_j(\tau) \left| u_{t-j} \right|,$$

where $\alpha_j(\tau) = a_j Q_{\varepsilon_t}(\tau), j = 0, 1, 2, \ldots$.

Let $m = m(n)$ be a truncation parameter; we may consider the following truncated quantile autoregression:

$$Q_{u_t}(\tau | \mathcal{F}_{t-1}) \approx a_0(\tau) + a_1(\tau)|u_{t-1}| + \cdots + a_m(\tau)|u_{t-m}|.$$

By choosing m suitably small relative to the sample size n, but large enough to avoid serious bias, we obtain a sieve approximation for the GARCH model.

One could estimate the conditional quantiles simply using a sieve approximation:

$$\check{Q}_{u_t}(\tau \,|\, \mathcal{F}_{t-1}) = \hat{a}_0(\tau) + \hat{\alpha}_1(\tau) \,|\, u_{t-1} \,| + \cdots + \hat{a}_m(\tau) \,|\, u_{t-m} \,|,$$

where $\hat{a}_j(\tau)$ are the quantile autoregression estimates. Under regularity assumptions

$$\check{Q}_{u_t}(\tau \,|\, \mathcal{F}_{t-1}) = Q_{u_t}(\tau \,|\, \mathcal{F}_{t-1}) + O_p(m/\sqrt{n}).$$

However, Monte Carlo evidence indicates that the simple sieve approximation does not directly provide a good estimator for the GARCH model, but it serves as an adequate preliminary estimator. Since the first step estimation focuses on the global model, it is desirable to use information over multiple quantiles in estimation.

Combining information over multiple quantiles helps us to obtain globally coherent estimate of the scale parameters.

Suppose that we estimate the m-th order quantile autoregression

$$\widetilde{\alpha}(\tau) = \arg\min_{\alpha} \sum_{t=m+1}^{n} \rho_\tau\left(u_t - \alpha_0 - \sum_{j=1}^{m} \alpha_j |u_{t-j}|\right) \tag{41.24}$$

at quantiles $(\tau_1, \ldots, \tau_K)$ and obtain estimates $\widetilde{\alpha}(\tau_k)$, $k = 1, \ldots, K$. Let $\widetilde{\alpha}_0 = 1$ in accordance with the identification assumption. Denote

$$\mathbf{a} = [a_1, \ldots, a_m, q_1, \ldots, q_K]^\top, \quad \boldsymbol{\overline{\pi}} = \left[\widetilde{\alpha}(\tau_1)^\top, \ldots, \widetilde{\alpha}(\tau_K)^\top\right]^\top,$$

where $q_k = Q_{\varepsilon_t}(\tau_k)$, and

$$\phi(\mathbf{a}) = g \otimes \alpha = [q_1, a_1 q_1, \ldots, a_m q_1, \ldots, q_K, a_1 q_K, \ldots, a_m q_K]^\top,$$

where $g = [q_1, \ldots, q_K]^\top$ and $\alpha = [1, a_1, a_2, \ldots, a_m]^\top$; we consider the following estimator for the vector a that combines information over the K quantile estimates based on the restrictions $\alpha_j(\tau) = a_j Q_{\varepsilon_t}(\tau)$:

$$\widetilde{\mathbf{a}} = \arg\min_{\mathbf{a}} (\boldsymbol{\overline{\pi}} - \phi(\mathbf{a}))^\top A_n (\boldsymbol{\overline{\pi}} - \phi(\mathbf{a})), \tag{41.25}$$

where A_n is a $(K(m + 1)) \times (K(m + 1))$ positive definite matrix. Denoting $\widetilde{\mathbf{a}} = (\widetilde{a}_0, \ldots, \widetilde{a}_m)$, σ_t can be estimated by

$$\widetilde{\sigma}_t = \widetilde{a}_0 + \sum_{j=1}^{m} \widetilde{a}_j |u_{t-j}|.$$

In the second step, we perform a quantile regression of u_t on $\widetilde{Z}_t = \left(1, \widetilde{\sigma}_{t-1}, \ldots \widetilde{\sigma}_{t-p}, |u_{t-1}|, \ldots, |u_{t-q}|\right)^\top$ by

$$\min_{\theta} \sum_t \rho_\tau\left(u_t - \theta^\top \widetilde{Z}_t\right); \tag{41.26}$$

the two-step estimator of $\theta(\tau)^\top = (\beta_0(\tau), \beta_1(\tau), \ldots, \beta_p(\tau), \gamma_1(\tau), \ldots, \gamma_q(\tau))$ is then given by the solution of Eq. 41.26, $\widehat{\theta}(\tau)$, and the τ-th conditional quantile of u_t can be estimated by

$$\hat{Q}_{u_t}(\tau|\mathcal{F}_{t-1}) = \hat{\theta}(\tau)^\top \widetilde{Z}_t.$$

Iteration can be applied to the above procedure for further improvement.

Let $\widetilde{\alpha}(\tau)$ be the solution of Eq. 41.24; then under appropriate assumptions, we have

$$\|\widetilde{\alpha}(\tau)-\alpha(\tau)\|^2 = O_p(m/n), \tag{41.27}$$

and for any $\lambda \in \mathcal{R}^{m+1}$,

$$\frac{\sqrt{n}\lambda^\top(\widetilde{\alpha}(\tau)-\alpha(\tau))}{\sigma_\lambda} \Rightarrow N(0,1),$$

where $\sigma_\lambda^2 = f_\varepsilon(F_\varepsilon^{-1}(\tau))^{-2}\lambda^\top D_n^{-1}\sum_n(\tau)D_n^{-1}\lambda$, and

$$D_n = \left[\frac{1}{n}\sum_{t=m+1}^{n}\frac{x_t x_t^\top}{\sigma_t}\right], \Sigma_n(\tau) = \frac{1}{n}\sum_{t=m+1}^{n} x_t x_t^T Y_\tau^2(u_{t\tau}),$$

where $x_t = (1,|u_{t-1}|, \ldots, |u_{t-m}|)^\top$.

Define

$$G = \left.\frac{\partial\phi(\mathbf{a})}{\partial\mathbf{a}^\top}\right|_{\mathbf{a}=\mathbf{a}_0} = \dot{\phi}(\mathbf{a}_0) = [g\otimes J_m \vdots I_K\otimes\alpha_0], g_0 = \begin{bmatrix} Q_{\varepsilon_t}(\tau_1) \\ \cdots \\ Q_{\varepsilon_t}(\tau_K)\end{bmatrix},$$

where g_0 and α_0 are the true values of vectors $g = [q_1, \ldots, q_K]^\top$ and $\alpha = [1, a_1, a_2, \ldots, a_m]^\top$, and

$$J_m = \begin{bmatrix} 0 & \cdots & 0 \\ 1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & 1 \end{bmatrix}$$

is an $(m+1)\times m$ matrix and I_K is a K-dimensional identity matrix; under regularity assumptions, the minimum distance estimator $\widetilde{\mathbf{a}}$ solving (Eq. 41.25) has the following asymptotic representation:

$$\sqrt{n}(\hat{\mathbf{a}}-\mathbf{a}_0) = \left[G^\top A_n G\right]^{-1} G^\top A_n\sqrt{n}(\overline{\boldsymbol{\pi}}-\boldsymbol{\pi}) + o_p(1)$$

where

$$\sqrt{n}(\overline{\boldsymbol{\pi}}-\boldsymbol{\pi}) = -\frac{1}{\sqrt{n}}\sum_{t=m+1}^{n}\begin{bmatrix}\left(D_n^{-1}x_t\dfrac{y_{\tau_1}(u_{t\tau_1})}{f_\varepsilon\left(F_\varepsilon^{-1}(\tau_1)\right)}\right) \\ \cdots \\ \left(D_n^{-1}x_t\dfrac{y_{\tau_m}(u_{t\tau_m})}{f_\varepsilon\left(F_\varepsilon^{-1}(\tau_m)\right)}\right)\end{bmatrix} + o_p(1),$$

and the two-step estimator $\hat{\theta}(\tau)$ based on Eq. 41.26 has asymptotic representation:

$$\sqrt{n}\left(\hat{\theta}(\tau)-\theta(\tau)\right)=-\frac{1}{f_{\varepsilon}\left(F_{\varepsilon}^{-1}(\tau)\right)}\Omega^{-1}\left\{\frac{1}{\sqrt{n}}\sum_{t}Z_t y_{\tau}(u_{t\tau})\right\}+\Omega^{-1}\Gamma\sqrt{n}(\tilde{\alpha}-\alpha)+o_p(1),$$

where $a=[a_1, a_2, \ldots, a_m]^{\top}$, $\Omega=E[Z_tZ_t^{\top}/\sigma_t]$, and

$$\Gamma=\sum_{k=1}^{p}\theta_k C_k, C_k=\mathrm{E}\left[(|u_{t-k-1}|,\ldots,|u_{t-k-m}|)\frac{Z_t}{\sigma_t}\right].$$

41.6 An Empirical Application

41.6.1 Data and the Empirical Model

In this section, we apply the quantile regression method to five major world equity market indexes. The data used in our application are the weekly return series, from September 1976 to June 2008, of five major world equity market indexes: the US S&P 500 Composite Index, the Japanese Nikkei 225 Index, the UK FTSE 100 Index, the Hong Kong Hang Seng Index, and the Singapore Strait Times Index. The FTSE 100 Index data are from January 1984 to June 2008. Table 41.1 reports some summary statistics of the data.

The mean weekly returns of the five indexes are all over 0.1 % per week, with the Hang Seng Index producing an average return of 0.23 % per week, an astonishing

Table 41.1 Summary statistics of the data

	S&P 500	Nikkei 225	FTSE 100	Hang Seng	Singapore ST
Mean	0.0015	0.0010	0.0017	0.0023	0.0012
Std. Dev.	0.0199	0.0253	0.0237	0.0376	0.0291
Max	0.1002	0.1205	0.1307	0.1592	0.1987
Min	−0.1566	−0.1289	−0.2489	−0.5401	−0.4551
Skewness	−0.4687	−0.2982	−1.7105	−3.0124	−1.5077
Excess kurtosis	3.3494	2.9958	12.867	9.8971	19.3154
AC(1)	−0.0703	−0.0306	0.0197	0.0891	0.0592
AC(2)	0.0508	0.0665	0.0916	0.0803	0.0081
AC(3)	0.0188	0.0328	−0.0490	−0.0171	0.0336
AC(4)	−0.0039	−0.0418	−0.0202	−0.0122	0.0099
AC(5)	−0.0189	−0.0053	−0.0069	−0.0386	0.0519
AC(10)	−0.0446	−0.0712	0.0138	−0.0345	−0.0227

This table shows the summary statistics for the weekly returns of five major equity indexes of the world. AC(k) denotes autocorrelation of order k. The source of the data is the online data service Datastream increase in the index level over the sample period. In comparison, the average return of Nikkei 225 index is only 0.1 %. The Hang Seng's phenomenal rise does not come without risk. The weekly sample standard deviation of the index is 3.76 %, the highest of the five indexes. In addition, over the sample period the Hang Seng suffered four larger than 15 % drop in weekly index level, with maximum loss reaching 35 %, and there were 23 weekly returns below −10 %! As has been documented extensively in the literature, all five indexes display negative skewness and excess kurtosis. The excess kurtosis of Singapore Strait Times Index reached 19.31, to a large extent driven by the huge 1 week loss of 47.47 % during the 1987 market crash. The autocorrelation coefficients for all five indexes are fairly small. The Hang Seng Index seems to display the strongest autocorrelation with the AR(1) coefficient equal to 0.0891.

We consider an AR-linear ARCH model in the empirical analysis. Thus, the return process is modeled as

$$r_t = \alpha_0 + \alpha_1 r_{t-1} + \cdots + \alpha_s r_{t-s} + u_t, \tag{41.28}$$

where

$$u_t = \sigma_t \varepsilon_t, \sigma_t = \gamma_0 + \gamma_1 |u_{t-1}| + \cdots + \gamma_q |u_{t-q}|,$$

and the τ Conditional VaR of u_t is given by

$$Q_{u_t}(\tau|\mathcal{F}_{t-1}) = \gamma(\tau)' Z_t$$

$$\gamma(\tau)' = (\gamma_0(\tau), \gamma_1(\tau), \ldots, \gamma_q(\tau)), \text{ and } Z_t = (1, |u_{t-1}|, \ldots, |u_{t-q}|)'.$$

For each time series, we first conduct model specification analysis and choose the appropriate lags for the mean equation and the quantile ARCH component. Based on the selected model, we use Eq. 41.28 to obtain a time series of residuals. The residuals are then used in the ARCH VaR estimation using a quantile regression.

41.6.2 Model Specification Analysis

We conduct sequential tests for the significance of the coefficients on lags. The inference procedures we use here are asymptotic inferences. For estimation of the covariance matrix, we use the robust HAC (Heteroskedastic and Autocorrelation Consistent) covariance matrix estimator of Andrews (1991) with the data-dependent automatic bandwidth parameter estimator recommended in that paper. First of all, we choose the lag length in the autoregression,

$$r_t = \alpha_0 + \alpha_1 r_{t-1} + \cdots + \alpha_s r_{t-s} + u_t,$$

using a sequential test of significance on lag coefficients. The maximum lag length that we start with is $s = 9$, and the procedure is repeated until a rejection occurs. Table 41.2 reports the sequential testing results for the S&P 500 index. The t-statistics of all the coefficients are listed for nine rounds of the test. We see that the t-statistic of the coefficient with the maximum number of lags does not become significant until $s = 1$, the ninth round. The preferred model is an AR(1) model. The selected mean equations for all five indexes are reported in Table 41.4.

Our next task is to select the lag length in the ARCH effect

$$u_t = \left(\gamma_0 + \gamma_1 |u_{t-1}| + \cdots + \gamma_q |u_{t-q}|\right)\varepsilon_t.$$

Again, a sequential test is conducted. To calculate the t-statistic, we need to estimate $\omega^2 = \tau(1 - \tau)/f(F^{-1}(\tau))^2$. There are many studies on estimating $f(F^{-1}(t))$, including Siddiqui (1960), Bofinger (1975), Sheather and Maritz (1983), and Welsh (1987). Notice that

$$\frac{dF^{-1}(t)}{dt} = \frac{1}{f\left(F^{-1}(t)\right)}; \tag{41.29}$$

following Siddiqui (1960), we may estimate (Eq. 41.29) by a simple difference quotient of the empirical quantile function. As a result,

$$f\left(\widehat{F^{-1}(t)}\right) = \frac{2h_n}{\hat{F}^{-1}(t + h_n) - \hat{F}^{-1}(t - h_n)} \tag{41.30}$$

where $\hat{F}^{-1}(t)$ is an estimate of $F^{-1}(t)$ and h_n is a bandwidth which goes to zero as $n \to \infty$. A bandwidth choice has been suggested by Hall and Sheather (1988) based on Edgeworth expansion for studentized quantiles. This bandwidth is of order $n^{-1/3}$ and has the following representation:

$$h_{HS} = z_\alpha^{2/3}\left[1.5s(t)/s''(t)\right]^{1/3} n^{-1/3},$$

where z_α satisfies $\Phi(z_a) = 1 - \alpha/2$ for the construction of $1-\alpha$ confidence intervals. In the absence of additional information, $s(t)$ is just the normal density. Starting with $q_{\max} = 10$, a sequential test was conducted and results for the 5 % VaR model of the S&P 500 Index are reported in Table 41.3. We see that in the fourth round, the t-statistic on lag 7 becomes significant. The sequential test stops here, and it suggests that ARCH(7) is appropriate.

Based on the model selection tests, we decide to use the AR(1)-ARCH(7) regression quantile model to estimate 5 % VaR for the S&P 500 index. We also conduct similar tests on the 5 % VaR models for other four indexes. To conserve space we do not report the entire testing process in the paper. Table 41.4 provides a summary of the selected models based on the tests. The mean equations

Table 41.2 VaR model mean specification test for the S&P 500 Index

Round	1st	2nd	3rd	4th	5th	6th	7th	8th	9th
α_0	3.3460	3.3003	3.2846	3.3248	3.2219	3.7304	3.1723	3.0650	3.8125
α_1	−1.6941	−1.7693	−1.8249	−1.9987	−1.9996	−2.0868	−2.1536	−2.097	2.2094
α_2	1.2950	1.3464	1.1555	1.0776	0.0872	1.3106	1.2089	1.0016	
α_3	−0.9235	−0.9565	−0.9774	1.5521	−0.8123	−0.8162	−0.9553		
α_4	−1.0414	−1.0080	−0.9947	−1.0102	−0.9899	−0.1612			
α_5	−0.7776	−0.7642	−0.7865	−0.8288	−0.7662				
α_6	0.2094	0.5362	0.7166	−0.8931					
α_7	−1.5594	−1.5426	−1.5233						
α_8	−0.8926	−0.8664							
α_9	−0.3816								

This table reports the test results for the VaR model mean equation specification for the S&P 500 Index. The number of lags in the AR component of the ARCH model is selected according to the sequential test. The table reports the *t*-statistic for the coefficient with the maximum lag in the mean equation

Table 41.3 5 % VaR model ARCH specification test for the S&P 500 Index

Round	1st	2nd	3rd	4th
γ_0	−16.856	−15.263	−17.118	−15.362
γ_1	2.9163	3.1891	3.2011	3.1106
γ_2	1.9601	2.658	2.533	2.321
γ_3	1.0982	1.0002	0.9951	1.0089
γ_4	0.6807	0.8954	1.1124	1.5811
γ_5	0.7456	0.8913	0.9016	0.9156
γ_6	0.3362	0.3456	0.4520	0.3795
γ_7	1.9868	2.0197	1.8145	2.1105
γ_8	0.4866	0.4688	1.5631	
γ_9	1.2045	1.0108		
γ_{10}	1.1326			

This table reports the test results for the 5 % VaR model specification for the S&P 500 Index. The number of lags in the volatility component of the ARCH model is selected according to the test. The table reports the *t*-statistic for the coefficient with the maximum lag in the ARCH equation

Table 41.4 ARCH VaR models selected by the sequential test

Index	Mean Lag	5 % ARCH Lag
S&P 500	1	6
Nikkei 225	1	7
FTSE 100	1	6
Hang Seng	3	6
Singapore ST	2	7

This table summarizes the preferred ARCH VaR models for the five global market indexes. The number of lags in the mean equation and the volatility component of the ARCH model is selected according to the test

generally have one or two lags, except the Hang Seng Index, which has a lag of 3 and displays more persistent autoregressive effect.

For the ARCH equations, at least six lags are needed for the indexes.

41.6.3 Estimated VaRs

The estimated parameters for the mean equations for all five indexes are reported in Table 41.5. The constant term for the five indexes is between 0.11 % for the Nikkei and 0.24 % for the Hang Seng. As suggested by Table 41.1, the Hang Seng seems to display the strongest autocorrelation, and this is reflected in the four lags chosen by the sequential test. Table 41.6 reports the estimated quantile regression ARCH parameters for the 5 % VaR model:

USA – S&P 500 Index. The estimated 5 % VaRs generally range between 2.5 % and 5 %, but during very volatile periods they could jump over 10 %, as what happened in October 1987. During high-volatility periods, there is high variation in estimated VaRs.

Table 41.5 Estimated mean equation parameters

Round	S&P 500	Nikkei 225	FTSE 100	Hang Seng	Singapore ST
α_0	0.0019	0.0011	0.0022	0.0024	0.0014
	(0.0006)	(0.0006)	(0.0008)	(0.001)	(0.0009)
α_1	−0.0579	−0.0827	0.0617	0.1110	0.0555
	(0.0233)	(0.0305)	(0.0283)	(0.0275)	(0.0225)
α_2				0.0796	0.0751
				(0.0288)	(0.0288)
α_3				−0.0985	
				(0.0238)	

This table reports the estimated parameters of the mean equation for the five global equity indexes. The standard errors are in parentheses under the estimated parameters

Table 41.6 Estimated ARCH equation parameters for the 5 % VaR model

Parameter	S&P 500	Nikkei 225	FTSE 100	Hang Seng	Singapore ST
γ_0	0.0351	0.0421	0.0346	0.0646	0.0428
	(0.0016)	(0.0023)	(0.0013)	(0.0031)	(0.0027)
γ_1	0.2096	0.0651	0.0518	0.1712	0.1119
	(0.0711)	(0.0416)	(0.0645)	(0.0803)	(0.0502)
γ_2	0.1007	0.1896	0.0588	0.0922	0.1389
	(0.0531)	(0.0415)	(0.0665)	(0.0314)	(0.0593)
γ_3	−0.0101	0.1109	0.0311	0.2054	0.0218
	(0.0142)	(0.0651)	(0.0242)	(0.0409)	(0.0379)
γ_4	0.1466	0.0528	0.0589	0.0671	0.1102
	(0.0908)	(0.0375)	(0.0776)	(0.0321)	(0.0903)
γ_5	0.0105	0.0987	−0.0119	0.0229	0.1519
	(0.0136)	(0.0448)	(0.0123)	(0.0338)	(0.0511)
γ_6	0.0318	0.0155	0.0876	0.0359	0.0311
	(0.0117)	(0.0297)	(0.0412)	(0.0136)	(0.0215)
γ_7		0.2323			0.1123
		(0.0451)			(0.0517)

This table reports the estimated parameters of the ARCH equation for the 5 % VaR model for the five global indexes. The standard errors are in parentheses under the estimated parameters

Japan – Nikkei 225 Index. The estimated VaR series is quite stable and remains at the 4 % and the 7 % level from 1976 till 1982. Then the Nikkei 225 Index took off and appreciated about 450 % over the next 8 years, reaching its highest level at the end of 1989. This quick rise in stock value is accompanied by high risk, manifested here by the more volatile VaR series. In particular, the VaRs fluctuated dramatically, ranging from a low of 3 % to a high of 15 %. This volatility in VaR may reflect both optimistic market outlook at times and worry about high valuation and the possibility of a market crash. That crash did come in 1990, and

10 years later, the Nikkei 225 Index still hovers around at a level which is about half off the record high in 1989. The 1990s is far from a rewarding decade for investors in the Japanese equity market. Average weekly 5 % VaR is about 5 %, and the variation is also very high.

UK – FTSE 100 Index. The 5 % VaR is very stable and averages about 3 %. They stay very much within the 2–4 % band, except on a few occasions, such as the 1987 global market crash.

Hong Kong – Hang Seng Index. The Hang Seng Index produces an average return of 0.23 % per week. The Hang Seng's phenomenal rise does not come without risk. We mentioned above that the weekly sample standard deviation of the index is 3.76 %, the highest of the five indexes. In addition, the Hong Kong stock market has had more than its fair share of the market crashes.

Singapore – Strait Times Index. Interestingly, the estimated VaRs display a pattern very similar to that of the UK FTSE 100 Index, although the former is generally larger than the latter. The higher risk in the Singapore market did not result in higher return over the sample period. Among the five indexes, the Singapore market suffered the largest loss during the 1987 crash, a 47.5 % drop in a week. The market has since recovered much of the loss. Among the five indexes, the Singapore market only outperformed the Nikkei 225 Index over this period.

41.6.4 Performance of the ARCH Quantile Regression Model

In this section we conduct an empirical analysis to compare VaRs estimated by RiskMetrics and regression quantiles and those by volatility models with the conditional normality assumption. There are extensive empirical evidences supporting the use of the GARCH models in conditional volatility estimation. Bollerslev et al. (1992) provide a nice overview of the issue. Therefore, we compare VaR estimated based on RiskMetrics and GARCH(1,1) model and quantile regression based on ARCH.

To measure the relative performance more accurately, we compute the percentage of realized returns that are below the negative estimated VaRs. The results are reported in Table 41.7. The top panel of the table presents the percentages for the VaRs estimated by the ARCH quantile regression model, the middle panel for the VaRs estimated by the GARCH model with the conditional normal return distribution assumption, and the bottom panel for the VaRs estimated by the RiskMetrics method. We estimate VaRs using these methods at 1 %, 2 %, 5 %, 10 %. Now we have a total of four percentage levels. The regression quantile method produces the closest percentage in general. Both the RiskMetrics method and the GARCH method seem to underestimate VaRs for the smaller percentages and overestimate VaRs for the larger percentages.

The five indexes we analyzed are quite different in their risk characteristics as discussed above. The quantile regression approach seems to be relatively robust and can consistently produce reasonably good estimates of the VaRs at different

Table 41.7 VaR model performance comparison

% VaR	1 %	2 %	5 %	10 %
Quantile regression				
S&P 500	1.319	1.925	5.3108	9.656
Nikkei 225	1.350	2.011	5.7210	10.56
FTSE 100	0.714	1.867	5.6019	9.016
Hang Seng	0.799	2.113	4.9011	9.289
GARCH				
S&P 500	1.3996	1.7641	4.0114	7.6151
Nikkei 225	1.4974	1.7927	4.3676	8.4098
FTSE 100	1.1980	1.6133	3.3891	6.7717
Hang Seng	1.8962	2.8658	3.6653	7.6439
RiskMetrics				
S&P 500	0.3790	0.5199	1.1180	3.2563
Nikkei 225	0.5877	0.9814	1.358	4.1367
FTSE 100	0.2979	0.5796	0.9984	3.5625
Hang Seng	0.7798	0.9822	1.4212	4.1936

This table reports the coverage ratios, i.e., the percentage of realized returns that is below the estimated VaRs. The top panel reports the performance of the VaRs estimated by the quantile regression model. The middle panel reports the results for VaRs estimated by the GARCH model based on the conditionally normal return distribution assumption. The bottom panel reports the results for VaRs estimated by the RiskMetrics method

percentage (probability) levels. The GARCH model with the normality assumption, being a good volatility model, is not able to produce good VaR estimates. The quantile regression model does not assume normality and is well suited to hand negative skewness and heavy tails.

41.7 Conclusion

Quantile regression provides a convenient and powerful method of estimating VaR. The quantile regression approach not only provides a method of estimating the conditional quantiles (VaRs) of existing time-series models; it also substantially expands the modeling options for time-series analysis. Estimating Value at Risk using the quantile regression does not assume a particular conditional distribution for the returns. Numerical evidence indicates that the quantile-based methods have better performance than the traditional J. P. Morgan's RiskMetrics method and other methods based on normality. The quantile regression based method provides an important tool in risk management.

There are several existing programs for quantile regression applications. For example, both parametric and nonparametric quantile regression estimations can be implemented by the function **rq**() and **rqss**() in the package **quantreg** in the computing language **R**, and **SAS** now has a suite of procedures modeled closely on the functionality of the R package quantreg.

References

Andrews, D. W. K. (1991). Heteroskedasticity and autocorrelation consistent covariance matrix estimation. *Econometrica, 59*, 817–858.

Barrodale, I., & Roberts, F. D. K. (1974). Solution of an overdetermined system of equations in the l_1 norm. *Communications of the ACM, 17*, 319–320.

Blanley, A., Lamb, R., & Schroeder, R. (2000). Compliance with SEC disclosure requirements about market risk. *Journal of Derivatives, 7*, 39–50.

Bofinger, E. (1975). Estimation of a density function using order statistics. *Australian Journal of Statistics, 17*, 1–7.

Bollerslev, T., Chou, R. Y., & Kroner, K. F. (1992). ARCH modeling in finance. *Journal of Econometrics, 52*, 5–59.

Bouyé, E., & Salmon, M. (2008). *Dynamic copula quantile regressions and tail area dynamic dependence in forex markets* [Manuscript]. Coventry: Financial Econometrics Research Centre, Warwick Business School.

Chen, X., Koenker, R., & Xiao, Z. (2009). Copula-based nonlinear quantile autoregression. *The Econometrics Journal, 12*, 50–67.

Dowd, K. (1998). Beyond value at risk: The new science of risk management. England: Wiley.

Dowd, K. (2000). Assessing VaR accuracy. *Derivatives Quarterly, 6*, 61–63.

Duffie, D., & Pan, J. (1997). An overview of value at risk. *Journal of Derivatives, 4*, 7–49.

Engle, R. F., & Manganelli, S. (1999). *CAViaR: Conditional autoregressive value at risk by regression quantiles* (Working paper). San Diego: University of California.

Engle, R. F., & Manganelli, S. (2004). *CAViaR: Conditional autoregressive value at risk by regression quantiles* (Working paper). San Diego: University of California.

Guo, H., Wu, G., & Xiao, Z. (2007). Estimating value at risk for defaultable bond portfolios by regression quantile. *Journal of Risk Finance, 8*(2), 166–185.

Hall, P., & Sheather, S. J. (1988). On the distribution of a studentized quantile. *Journal of Royal Statistical Society B, 50*, 381–391.

Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica, 84*, 33–50.

Koenker, R., & d'Orey, V. (1987). Computing regression quantiles. *Applied Statistics, 36*, 383–393.

Koenker, R., & Xiao, Z. (2006). Quantile autoregression. *Journal of the American Statistical Association, 101*(475), 980–1006.

Koenker, R., & Zhao, Q. (1996). Conditional quantile estimation and inference for ARCH models. *Econometric Theory, 12*, 793–813.

Portnoy, S. L., & Koenker, R. (1997). The gaussian hare and the Laplacian tortoise: Computability of squared-error vs. absolute-error estimator. *Statistical Science, 12*, 279–300.

Sarma, M., Thomas, S., & Shah, A. (2000). *Performance evaluation of alternative VaR models* (Working paper). Indira Gandhi Institute of Development Research.

Saunders, A. (1999). *Financial institutions management: A modern perspective, Irwin series in finance*. McGraw-Hill, New York.

Sheather, S. J., & Maritz, J. S. (1983). An estimate of the asymptotic standard error of the sample median. *Australian Journal of Statistics, 25*, 109–122.

Siddiqui, M. (1960). Distribution of quantiles in samples from a bivariate population. *Journa of Research Natural Bureau Standards, 64B*, 145–150.

Welsh, A. (1987). Asymptotically efficient estimation of the sparsity function at a point. *Statistics and Probability Letters, 6*, 427–432.

Wu, G., & Zhijie Xiao (2002). An analysis of risk measures. *Journal of Risk, 4*(4), 53–75.

Xiao, Z., & Koenker, R. (2009). Conditional quantile estimation and inference for GARCH models. *Journal of the American Statistical Association, 104*(488), 1696–1712.

Earnings Quality and Board Structure: Evidence from South East Asia

42

Kin-Wai Lee

Contents

Abstract

Using a sample of listed firms in Southeast Asian countries, this paper examines the association among board structure and corporate ownership structure in affecting earnings quality. I find that the negative association between separation

K.-W. Lee
Division of Accounting, Nanyang Business School, Nanyang Technological University, Singapore, Singapore
e-mail: akwlee@ntu.edu.sg

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_42,

of control rights from cash flow rights and earnings quality varies systematically with board structure. I find that the negative association between separation of control rights from cash flow rights and earnings quality is less pronounced in firms with high equity ownership by outside directors. I also document that in firms with high separation of control rights from cash flow rights, those firms with higher proportion of outside directors on the board have higher earnings quality. Overall, my results suggest that outside directors' equity ownership and board independence are associated with better financial reporting outcome, especially in firms with high expected agency costs arising from misalignment of control rights and cash flow rights.

The econometric method employed is regressions of panel data. In a panel data setting, I address both cross-sectional and time-series dependence. Gow et al. (2010, *The Accounting Review 85*(2), 483–512) find that in the presence of both cross-sectional and time-series dependence, the two-way clustering method which allows for both cross-sectional and time-series dependence produces well-specified test statistics. Following Gow et al. (2010, *The Accounting Review 85*(2), 483–512), I employ the two-way clustering method where the standard errors are clustered by both firm and year in my regressions of panel data. Johnston and DiNardo (1997, *Econometrics method*. New York: Mc-Graw Hill) and Greene (2000, *Econometrics analysis*. Upper Saddle River: Prentice-Hall) are two econometric textbooks that contain a detailed discussion of the econometrics issues relating to panel data.

Keywords

Earnings quality • Board structure • Corporate ownership structure • Panel data regressions • Cross-sectional and time-series dependence • Two-way clustering method of standard errors

42.1 Introduction

In Asia, corporate ownership concentration is high, and many listed firms are mainly controlled by a single large shareholder (La Porta et al. 1999; Claessens et al. 2000). Asian firms also show a high divergence between control rights and cash flow rights, which allows the largest shareholder to control a firm's operations with a relatively small direct stake in its cash flow rights. Control is often increased beyond ownership stakes through pyramid structures, cross-holdings among firms, and dual class shares (Claessens et al. 2000). It is argued that concentrated ownership facilitates transactions in weak property rights environment by providing the controlling shareholders the power and incentive to negotiate and enforce contracts with various stakeholders (Shleifer and Vishny 1997). As a result of concentrated ownership, the main agency problem in listed firms in Asia is the conflict of interest between the controlling shareholder and minority shareholder. Specifically, controlling shareholder has incentives to expropriate the wealth of minority shareholders by engaging in rent-seeking activities and to mask their private benefits of control by supplying low-quality financial accounting information. Empirical evidence also shows that the quality and credibility of financial accounting information are lower in firms with high separation of control rights and cash flow rights (Fan and Wong 2002; Haw et al. 2004).

An important question is how effective are corporate governance mechanisms in mitigating the agency problems in Asia firms, especially in improving corporate transparency in firms with concentrated ownership. Controlling shareholders in Asia typically face limited disciplinary pressures from the market for corporate control because hostile takeovers are infrequent (La Porta et al. 1999; Fan and Wong 2002). Furthermore, controlling shareholders face little monitoring pressure from analysts because analysts are less likely to follow firms with potential incentives to withhold or manipulate information, such as when the family/management group is the largest control rights blockholder (Lang et al. 2004). In these environments, external corporate governance mechanisms, in particular the market for corporate control and analysts' scrutiny, exert limited disciplinary pressure on controlling shareholders. Consequently, internal corporate governance mechanisms such as the board of directors may be important to mitigate the agency costs associated with the ownership structure of Asian firms. Thus, the primary research questions in this paper are: (1) Do board of directors play a corporate governance role over the financial reporting process in listed firms in Asia? (2) How does the board of director affect financial reporting quality in firms with high expected agency costs arising from the separation of control rights and cash flow rights?

Specifically, this paper examines the relation among outside directors' equity ownership, board independence, and separation of control rights from cash flow rights of controlling shareholder in affecting earnings quality. My empirical strategy is as follows: First, I examine the main effect between earnings quality and (i) outside directors' equity ownership, (ii) the proportion of outside directors on the board, and (iii) the separation of control rights from cash flow rights of the largest ultimate shareholder. This sheds light on my first research question on whether the board of directors plays a corporate governance role over the financial reporting process in listed firms in Asia. Second, I examine the (i) interaction between outside directors' equity ownership and the separation of control rights from cash flow rights and (ii) interaction between the proportion of outside directors on the board and the separation of control rights from cash flow rights, in shaping earnings quality. This addresses the second research question on the effect of board structure (in particular, board independence and equity ownership of outside directors) on financial reporting quality in firms with high expected agency costs arising from the separation of control rights and cash flow rights.

In this paper, I focus on two important attributes of board monitoring – outside directors' equity ownership and board independence – and their association with financial reporting quality. These attributes are important for two reasons.

First, prior research generally finds that in developed economies such as the United States and the United Kingdom, there is a positive association between board independence and earnings quality (Dechow et al. 1996; Klein 2002; Peasnell et al. 2005). However, there is limited evidence on the effect of board independence on the financial accounting process in Asia. My paper attempts to fill this gap. Second, recent research on the monitoring incentives of the board suggests that equity ownership of outside directors plays an important role in mitigating managerial entrenchment (Perry 2000; Ryan and Wiggins 2004). An implication of this stream of research is that even in firms with high board independence, entrenched managers can weaken the monitoring incentives of outside directors by reducing their equity-based compensation. In other words, board independence that is not properly augmented with incentive compensation may hamper the monitoring effectiveness of independent directors over management.

My sample consists of 2,875 firm-year observations during the period 2004–2008 in five Asian countries comprising Indonesia, Malaysia, the Philippines, Singapore, and Thailand. These countries provide a good setting to test the governance potential of the board of directors because shareholders in these countries typically suffer from misaligned managerial incentives, ineffective legal protection, and underdeveloped markets for corporate control (La Porta et al. 1999; Claessens et al. 2000; Fan and Wong 2002).

I measure earnings quality with three financial reporting metrics: (i) discretionary accruals, (ii) mapping of accruals to cash flow, and (iii) informativeness of reported earnings. My results are robust across alternative earnings quality metrics. I find that earnings quality is higher when outside directors have higher equity ownership. This result suggests that internal monitoring of the quality and credibility of accounting information is improved through aligning shareholders' and directors' incentives. Consistent with the monitoring role of outside directors (Fama and Jensen 1983), I also find that earnings quality is positively associated with the proportion of outside directors on the board. This result supports the notion that outside directors have incentives to be effective monitors in order to maintain the value of their reputational capital. Consistent with prior studies (Fan and Wong 2002; Haw et al. 2004), I also document that earnings quality is negatively associated with the separation of control rights from cash flow rights of the largest ultimate shareholder.

More importantly, I document that the negative association between separation of control rights from cash flow rights and earnings quality is less pronounced in firms with high equity ownership by outside directors. This result suggests equity ownership improves the incentives and monitoring intensity of outside directors in firms with high expected agency costs arising from the divergence of control rights from cash flow rights. Furthermore, my result indicates the negative association between separation of control rights from cash flow rights and earnings quality is mitigated by the higher proportion of outside directors on the board. This result provides evidence supporting the corporate governance role of outside directors in constraining managerial discretion over financial accounting process in firms with high levels of misalignment between control rights and cash flow rights. Collectively, my results suggest that strong internal governance structures can alleviate agency problems between the controlling shareholder and minority shareholders. More generally, my results highlight the interplay between board structure and corporate ownership structure in shaping earnings quality.

I perform several robustness tests. My results are robust across different economies. In addition, year-by-year regressions yield qualitatively similar results, suggesting my inferences are not time-period specific. I also include additional country-level institutional variables such as legal origin, country investor protection, and enforcement of shareholder rights. My results are qualitatively similar. Specifically, after controlling for country-level legal institutions, firm-specific internal governance mechanisms, namely – outside directors' equity ownership and board independence – continue to be important in mitigating the negative effects of the divergence between control rights and cash flow rights on earnings quality.

My study has several contributions. First, prior studies find that the divergence of control rights from cash flow rights reduces the informativeness of reported earnings (Fan and Wong 2002) and induces earnings management (Haw et al. 2004). I extend these studies by demonstrating two specific channels at the firm level – equity ownership by outside directors and proportion of outside directors on the board – that mitigate the negative association between earnings quality and divergence of control rights from cash flow rights. This result suggests that the board of directors play an important corporate governance role to alleviate agency problems in firms with entrenched insiders. My findings also complement Fan and Wong's (2005) result that given concentrated ownership, a controlling owner may introduce some monitoring or bonding mechanisms that limit his ability to expropriate minority shareholders and hence mitigate agency conflicts. In the Fan and Wong's study, high-quality external auditors alleviate agency problems in firms with concentrated ownership, whereas in my study, strong board of directors augmented with proper monitoring incentives mitigate agency problems in firms with concentrated ownership.

Second, my results suggest that there is an incremental role for firm-specific internal governance mechanisms, beyond country-level institutions, in improving the quality of financial information. Haw et al. (2004) find that earnings management that is induced by the divergence between control rights and cash flow rights is less pronounced in countries where (i) legal institutions protect minority shareholder rights (such as legal tradition, minority shareholder rights, efficiency of judicial system, or disclosure system) and (ii) in countries with effective extralegal institutions (such as the effectiveness of competition law, diffusion of the press, and tax compliance). My study shows that after controlling for both country-level legal and extralegal institutions, firm-specific internal governance mechanisms, namely, outside directors' incentive compensation and board independence, continue to be important in constraining management opportunism over the financial reporting process in firms with high expected agency costs arising from the divergence between control rights and cash flow rights. To the extent that changes in country-level legal institutions are relatively more costly and more difficult than changes in firm-level governance mechanisms, my result suggests that improvement in firm-specific governance mechanisms can be effective to reduce private benefits of control. My results complement finding in prior studies (Johnson et al. 2000; La Porta et al. 1998; Lang et al. 2004) that firms in countries with weak legal protection substitute with strong firm-level internal governance mechanisms to attract investors. My results also extend the finding in Leuz et al. (2003) that firms located in countries with weaker investor protection have higher earnings management. An important question is what factors may constrain managerial opportunism when country-level investor protection is weak? Because my sample consists of countries with generally weak investor protection, I shed light on this question by documenting that firm-level governance structures matter in improving earnings quality in countries with weak investor protection.

The rest of the paper proceeds as follows. Section 42.2 develops the hypotheses and places my paper in the context of related research. Section 42.3 describes the sample and method. Section 42.4 presents my results. I conclude the paper in Sect. 42.5.

42.2 Prior Research and Hypotheses Development

42.2.1 Equity Ownership of Outside Directors

Recent research examines the compensation structure of outside directors, who play an important monitoring role over management's actions. The central theme in this body of research is that incentive compensation leading to share ownership improves the outside directors' incentives to monitor. Mehran (1995) finds firm performance is positively associated with the proportion of directors' equity-based compensation. Perry (2000) finds that the likelihood of CEO turnover following poor performance increases when directors receive higher equity-based compensation. Shivdasani (1993) finds that probability of a hostile takeover is negatively associated with the percentage of shares owned by outside directors in target firms. He interprets this finding as suggesting that board monitoring may substitute for monitoring from the market of corporate control. Hermalin and Weisbach (1988) and Gillette et al. (2003) develop models where incentive compensation for directors increases their monitoring efforts and effectiveness. Ryan and Wiggins (2004) find that directors in firms with entrenched CEOs receive a significantly smaller proportion of compensation in the form of equity-based awards. Their result suggests that entrenched CEOs use their position to influence directors' compensation, which results in contracts that provide directors with weaker incentives to monitor management.

Internal monitoring is improved through aligning shareholders' and directors' incentives. If higher equity-based compensation contracts provide outside directors with stronger incentives to act in the interests of shareholders, I predict that managerial opportunism over the financial reporting process is reduced when outside directors have higher equity-based compensation. My first hypothesis is:

H1 Earnings quality is positively associated with the outside directors' equity ownership.

42.2.2 Board Independence

There is considerable literature on the role of outside directors in reducing agency problems between managers and shareholders. Fama and Jensen (1983) argue that outside directors have strong incentives to be effective monitors in order to maintain their reputational capital. Prior studies support the notion that board effectiveness in protecting shareholders' wealth is positively associated with the proportion of outside directors on the board (Weisbach 1988; Rosenstein and Wyatt 1990). In the United States, Klein (2002) finds that firms with high proportion of outside directors on the board have lower discretionary accruals. Using a sample of listed firms in the United Kingdom, Peasnell et al. (2005) document that the greater the board independence, the lower the propensity of managers making income-increasing discretionary accruals to avoid reporting losses and earnings reductions. Using US firms subjected to SEC enforcement action for alleged earnings manipulation, Dechow et al. (1996) and Beasley (1996) find that the probability of financial reporting fraud is negatively associated with the proportion of outside directors on the board.

In contrast, in emerging markets, conventional wisdom suggests that the agency conflicts between controlling owners and the minority shareholders may be difficult to mitigate through conventional corporate control mechanisms such as boards of directors (La Porta et al. 1998; Claessens et al. 2000; Fan and Wong 2005; Lee 2007; Lee et al. 2009). However, since the Asian economic crisis in 1997, many countries in Asia took steps to improve their corporate governance environment such as implementing country-specific code of corporate governance. For example, the Stock Exchange of Thailand Code of Best Practice for Directors of Listed Companies was implemented in 1998, the Code of Proper Practices for Directors for the Philippines was implemented in 2000, the Malaysian Code of Corporate Governance was implemented in 2000, and the Singapore Code of Corporate Governance was implemented in 2001. Among the key provisions of the code of corporate governance in these countries is the recommendation to have sufficient independent directors on the board to improve monitoring of management. For example, the 2001 Code of Corporate Governance for Singapore stated that there should be a strong and independent element on the board, which is able to exercise objective judgment on corporate affairs independently from management.

Although compliance with the code of corporate governance is not legally mandatory, listed companies are required to explain deviations from the recommendations of the code of corporate governance.[1]

I posit that in the waves of corporate governance reform in emerging markets in the early 2000s and the guidelines of country-specific code of corporate governance in emphasizing the importance of board independence, there is heightened awareness among outside directors on their increased monitoring responsibilities. To the extent that outside directors in listed firms in Asia perform a corporate governance role, I predict that:

H2 Earnings quality is positively associated with the proportion of outside directors on the board.

42.2.3 Equity-Based Compensation of Outside Directors and Control Divergence

The preceding discussion suggests that higher equity-based incentive compensation for outside directors improves their monitoring efforts. Greater monitoring from outside directors reduces managerial discretion over the financial reporting process (Lee et al. 2008). The benefits of more effective monitoring arising from higher equity ownership are likely to be concentrated in firms with high agency problems arising from the separation of control rights from cash flow rights. Thus, I predict that:

H3 The negative association between separation of control rights from cash flow rights and earnings quality is less pronounced in firms with high equity ownership by outside directors.

42.2.4 Board Independence and Control Divergence

Outside directors play an important corporate governance role in resolving agency problems between managers and shareholders. Following prior studies (Beasley 1996; Dechow et al. 1996; Klein 2002; Lee et al. 2007), higher proportion of outside directors on the board is associated with higher constraints on management discretion over the financial reporting process. I extend this notion to posit that the greater monitoring efforts from a high proportion of outside directors on the

[1]To illustrate, the Singapore Exchange Listing Rules require "listed companies to describe in the annual reports their corporate governance practices with specific reference to the principles of the Code, as well as disclose and explain any deviation from any guideline of the Code. Companies are also encouraged to make a positive confirmation at the start of the corporate governance section of the annual report that they have adhered to the principles and guidelines of the Code, or specify each area of non-compliance. Many of these guidelines are recommendations for companies to disclose their corporate governance arrangements."

board are likely to mitigate the negative effects of the separation of control rights from cash flow rights on earnings quality. Thus, I predict that:

H4 The negative association between separation of control rights from cash flow rights and earnings quality is mitigated by the proportion of outside directors.

42.3 Data

I begin with the Worldscope database to identify listed firms in five Asian countries comprising Indonesia, Malaysia, the Philippines, Singapore, and Thailand during the period 2004–2008. I exclude financial institutions because of their unique financial structure and regulatory requirements. I eliminate observations with extreme values of control variables such as return-on-assets and leverage. I obtain stock price data from the Datastream database. I obtain annual reports for the period 2004–2008 from the Global Report database and company websites. The sample consists of 617 firms for 2,875 firm-year observations during the period 2004–2008 in five Asian countries comprising Indonesia, Malaysia, the Philippines, Singapore, and Thailand.

I collect data on the board characteristics such as board size, the number of independent directors, and equity ownership of directors from the annual report. I also examine the annual report to trace the ultimate owners of the firms. The procedure of identifying ultimate owners is similar to the one used in La Porta et al. (1999).[2] In this study, I measure earnings quality with three financial reporting metrics: (i) discretionary accruals, (ii) mapping of accruals to cash flow, and (iii) informativeness of reported earnings.

Appendix 1 contains detailed description on the econometric method.

42.4 Results

42.4.1 Descriptive Statistics

Table 42.1 presents the descriptive statistics. Mean absolute discretionary accrual as a proportion of lagged assets is 0.062. Mean equity ownership of outside directors (computed as common stock and stock options held by outside directors divided by

[2]In summary, an ultimate owner is defined as the shareholder who has the determining voting rights of the company and who is not controlled by anyone else. If a company does not have an ultimate owner, it is classified as widely held. To economize on the data collection task, the ultimate owner's voting right level is set at 50 % and not traced any further once that level exceeds 50 %. Although a company can have more than one ultimate owner, we focus on the largest ultimate owner. We also identify the cash flow rights of the ultimate owners. To facilitate the measurement of the separation of cash flow and voting rights, the maximum cash flow rights level associated with any ultimate owner is also set at 50 %. However, there is no minimum cutoff level for cash flow rights.

Table 42.1 Descriptive statistics. The sample consists of 617 firms for 2,875 firm-year observations during the period 2004–2008 in five Asian countries comprising Indonesia, Malaysia, the Philippines, Singapore, and Thailand

	Mean	25th percentile	Median	75th percentile	Standard deviation
DISCAC	0.062	0.009	0.038	0.085	0.053
AQ	0.068	0.029	0.035	0.063	0.037
EBC (%)	2.041	0.837	1.752	2.663	1.035
OUTDIR	0.489	0.206	0.385	0.520	0.217
VOTE	1.198	1.000	1.175	1.326	0.638
BOARDSIZE	7	5	8	10	3
CEODUAL	0.405	0	0	1	–
LNASSET	11.722	9.867	11.993	13.078	2.115
MB	1.851	0.582	1.272	2.195	0.833
LEV	0.261	0.093	0.211	0.335	0.106
ROA	0.086	0.027	0.0705	0.109	0.071

DISCAC = absolute value of discretionary accruals estimated based on the modified Jones model
AQ = accrual quality measured by Dechow and Dichev's (2002) measure of mapping of accruals to past, present, and future cash from operations
DIROWN = common stock and stock options held by outside directors divided by number of ordinary shares outstanding in the firm
OUTDIR = proportion of outside directors on the board
VOTECASH = voting rights divided by cash flow rights of the largest controlling shareholder
CEODUAL = a dummy variable that equals 1 if the CEO is chairman of board and 0 otherwise
BOARDSIZE = number of directors on the board
LNASSET = natural logarithm of total assets
MB = market value of equity divided by book value of equity
LEV = long-term debt divided by total assets
ROA = net profit after tax divided by total assets

number of ordinary shares outstanding in the firm) is 2.04 %. The mean board size and proportion of outside directors on the board are 7 and 0.489, respectively. CEO chairs the board in 40 % of the firms. Consistent with Fan and Wong's (2002) study of East Asian economies, the firms in my sample also have high divergence of control rights from cash flow rights (mean VOTE = 1.198).

42.4.2 Discretionary Accruals

Table 42.2 presents the estimates of regressions of unsigned discretionary accruals on equity-based compensation, proportion of outside directors on the board, and the separation of control right from cash flow right. Following Gow et al. (2010), I employ the two-way clustering method where the standard errors are clustered by both firm and year in my regressions. In column (1), I document a negative association between the absolute value of discretionary accruals and the equity ownership of outside directors. Results also indicate that firms with higher proportion of outside directors on the board have lower discretionary accruals. I find that

Table 42.2 Regressions of unsigned discretionary accruals. The sample consists of 617 firms for 2,875 firm-year observations during the period 2004–2008 in five Asian countries comprising Indonesia, Malaysia, the Philippines, Singapore, and Thailand. The dependent variable is absolute discretionary accruals computed based on the modified Jones model. All variables are defined in Table 42.1. The t-statistics (in parentheses) are adjusted based on standard errors clustered by firm and year (Petersen 2009). The symbols *, **, and *** denote statistical significance at the 10 %, 5 %, and 1 % levels (two-tailed), respectively

	Predicted sign	1	2
EBC	–	−0.2513 (−3.07)***	−0.2142 (−2.86)***
OUTDIR	–	−0.1862 (−2.41)**	−0.1053 (−2.19)**
VOTE	+	0.8173 (2.85)***	0.9254 (2.94)***
VOTE *EBC	–		−0.4160 (−2.73)***
VOTE * OUTDIR	–		−0.1732 (−2.29)**
BOARDSIZE	+/–	0.0359 (1.57)	0.0817 (1.42)
CEODUAL	+	0.4192 (1.61)	0.2069 (1.45)
LNASSET	–	−0.5311 (−8.93)***	−0.5028 (−8.01)***
MB	+	0.1052 (4.25)***	0.2103 (3.02)***
LEV	+	1.2186 (5.38)***	1.1185 (5.19)***
ROA	+/–	−3.877 (−4.83)***	−4.108 (−4.72)***
Adjusted R^2		12.5 %	14.1 %

earnings management (as proxied by absolute discretionary accruals) increases as the separation between control rights and cash flow rights of controlling shareholders increases. This result is consistent with the finding in Haw et al. (2004). In column (2), I test whether the positive association between discretionary accruals and the separation between control rights and cash flow rights of controlling shareholder is mitigated by the equity ownership of outside directors. The coefficient on the interaction term between the separation of control rights from cash flow rights and the equity ownership of outside directors (VOTE* DIROWN) is negative and significant at the 1 % level, supporting the hypothesis that in firms with high separation of control rights from cash flow rights, earnings management is reduced when outside directors have higher equity ownership. This finding suggests that greater equity-based compensation increases the monitoring effectiveness of outside directors over the financial reporting process in firms with agency conflicts arising from their control rights from cash flow rights. In addition, the coefficient on the interaction term between the separation of control rights from cash flow rights and the proportion of outside directors (VOTE*OUTDIR) is negative and significant at the 5 % level, supporting the hypothesis that in firms with high separation of control rights from cash flow rights, earnings management is reduced in firms with high proportion of outside directors on the board. This result is consistent with the monitoring role of independent directors to improve the credibility of accounting information in firms with agency problems arising from their concentrated corporate ownership structure.

Table 42.3 Regressions of signed discretionary accruals. The sample consists of 617 firms for 2,875 firm-year observations during the period 2004–2008 in five Asian countries comprising Indonesia, Malaysia, the Philippines, Singapore, and Thailand. In column (1), the sample consists of firms with income-increasing discretionary accruals, and the dependent variable is positive discretionary accruals. In column (2), the sample consists of firms with income-decreasing discretionary accruals, and the dependent variable is negative discretionary accruals. All variables are defined in Table 42.1. All regressions contain dummy control variables for country, year, and industry. The t-statistics (in parentheses) are adjusted based on standard errors clustered by firm and year (Petersen 2009). The symbols *, **, and *** denote statistical significance at the 10 %, 5 %, and 1 % levels (two-tailed), respectively

	(1) Positive DISCAC	(2) Negative DISCAC
EBC	−0.1865 (−2.21)**	−0.2017 (−2.09)**
OUTDIR	−0.1172 (−2.08)**	−0.1302 (−2.11)**
VOTE	0.8103 (3.25)***	0.6735 (2.23)**
VOTE *EBC	−0.3952 (−2.49)***	−0.3064 (−2.05)**
VOTE * OUTDIR	−0.1732 (−2.13)**	−0.1105 (−2.01)**
BOARDSIZE	0.0533 (1.27)	0.0681 (1.09)
CEODUAL	0.1860 (1.32)	0.1562 (1.26)
LNASSET	−0.7590 (−5.22)***	−0.4463 (−6.12)***
MB	0.2019 (2.08)**	0.1085 (1.93)**
LEV	0.9781 (3.20)***	0.7701 (2.10)**
ROA	−3.087 (−4.13)***	−4.253 (−3.62)***
Adjusted R2	13.8 %	12.4 %

I partition the sample into two groups based on the sign of the firms' discretionary accruals. Table 42.3 column (1) presents the results using the subsample of firms with income-increasing discretionary accruals. Results indicate firms with higher equity ownership by outside directors, higher board independence, and lower divergence of control rights from cash flow rights, have lower income-increasing discretionary accruals. More importantly, I find that outside directors' equity ownership and proportion of outside directors mitigate the propensity of firms with high separation of control rights from cash flow rights to make higher income-increasing discretionary accruals. Table 42.3 column (2) presents the results using the subsample of firms with income-decreasing discretionary accruals. I find firms with higher equity ownership, higher board independence, and lower divergence of control rights from cash flow rights, have lower income-decreasing discretionary accruals. Furthermore, I find that equity ownership by outside directors and proportion of outside directors mitigate the propensity of firms with high separation of control rights from cash flow rights to make higher income-decreasing discretionary accruals.

In summary, when outside directors have equity ownership and when board independence is high, firms have *both* lower income-increasing and income-decreasing discretionary accruals, apparently mitigating earnings management

Table 42.4 Regressions of accrual quality. The sample consists of 617 firms for 2,875 firm-year observations during the period 2004–2008 in five Asian countries comprising Indonesia, Malaysia, the Philippines, Singapore, and Thailand. The dependent variable is AQ measured by Dechow and Dichev's (2002) measure of mapping of accruals to past, present, and future cash from operations with higher values of AQ denoting better accrual quality. All variables are defined in Table 42.1. All regressions contain dummy control variables for country, year, and industry. The t-statistics (in parentheses) are adjusted based on standard errors clustered by firm and year (Petersen 2009). The symbols *, **, and *** denote statistical significance at the 10 %, 5 %, and 1 % levels (two-tailed), respectively

	Predicted sign	1	2
EBC	+	0.3725 (3.11)***	0.2133 (3.26)***
OUTDIR	+	0.2049 (2.15)**	0.1557 (2.86)***
VOTE	–	−0.5108 (−3.74)***	−0.4352 (−3.05)***
VOTE *EBC	+		0.1751 (2.80)***
VOTE * OUTDIR	+		0.1163 (2.09)**
BOARDSIZE	+/–	0.1003 (1.29)	0.0642 (1.17)
CEODUAL	+/–	−0.1890 (−0.83)	−0.2173 (−1.56)
LNASSET	+	3.2513 (5.11)***	2.8764 (4.82)***
OPERCYCLE	–	−3.2941 (−4.75)***	−3.0185 (−5.01)***
NETPPE	+	1.1802 (3.35)***	0.9926 (3.72)***
STDCFO	–	−1.2981 (−2.83)***	−1.8344 (−3.32)***
STDSALE	–	−1.0203 (−2.77)***	−0.7845 (−2.09)**
NEGEARN	–	−0.8306 (−2.02)**	−1.0345 (−1.77)*
Adjusted R2		9.2 %	10.5 %

both on the upside and downside. For firms with greater separation of control rights from cash flow rights of controlling shareholders, those with high equity ownership by outside directors and those with high proportion of outside directors have lower income-increasing and lower income-decreasing discretionary accruals.

42.4.3 Accrual Quality

Table 42.4 presents regressions of accrual quality on corporate ownership structure and board characteristics. Following Gow et al. (2010), I employ the two-way clustering method where the standard errors are clustered by both firm and year in my regressions. In column (1), the coefficient EBC is positive and significant, suggesting that that firms whose directors receive higher equity ownership have higher accrual quality. Firms with high proportion of outside directors have higher accrual quality. The coefficient on VOTE is negative and significant, indicating the firms with high misalignment between control rights and cash flow rights have lower accrual quality. In column (2), the interaction term VOTE*DIROWN is positive and significant at the 1 % level. This finding suggests that firms with

Table 42.5 Regressions of returns on earnings. The sample consists of 617 firms for 2,875 firm-year observations during the period 2004–2008 in five Asian countries comprising Indonesia, Malaysia, the Philippines, Singapore, and Thailand. The dependent variable (RET) is 12-month cumulative raw return ending 3 months after the fiscal year-end. All regressions contain dummy control variables for country, year, and industry. The t-statistics (in parentheses) are adjusted based on standard errors clustered by firm and year (Petersen 2009). The symbols *, **, and *** denote statistical significance at the 10 %, 5 %, and 1 % levels (two-tailed), respectively

	Predicted sign	1	2
EARN	+	1.1735 (3.85)***	1.2811 (3.62)***
EARN *EBC	+	0.3122 (2.87)***	0.2983 (2.80)***
EARN * OUTDIR	+	0.1094 (2.13)**	0.1105 (2.08)**
EARN * VOTE	−	−0.6817 (−3.72)***	−0.7019 (−3.50)***
EARN * VOTE *EBC	+		0.2602 (2.83)***
EARN * VOTE * OUTDIR	+		0.1925 (2.15)**
EARN * BOARDSIZE	+/−	−0.0836 (−1.50)	−0.0801 (−1.22)
EARN * CEODUAL	+/−	−0.0405 (−1.42)	−0.0215 (−1.30)
EARN * LNASSET	+	0.2011 (2.89)***	0.3122 (3.07)***
EARN * MB	+	0.1573 (1.80)*	0.1806 (1.81)*
EARN * LEV	−	−0.7814 (−2.03)**	−0.6175 (−2.84)***
N		3,172	3,172
Adjusted R2		11.8 %	13.3 %

higher equity ownership by outside directors have a less pronounced negative association between accrual quality and the separation of control rights from cash flow rights of controlling shareholders. I then test whether board independence attenuates the negative association between accrual quality and the separation of control rights from cash flow rights of controlling shareholders. The interaction term VOTE*OUTDIR is positive and significant at the 5 % level. Hence, for firms with high separation of control rights from cash flow rights of controlling shareholder, those with higher proportion of outside directors have higher accrual quality. Collectively, my results suggest that stronger directors' equity ownership and higher board independence are associated with better financial reporting outcome, especially in firms with high expected agency costs arising from misalignment of control rights and cash flow rights.

42.4.4 Earnings Informativeness

Table 42.5 presents the regression results on earnings informativeness. The coefficient EARN* DIROWN is positive and significant, indicating the greater the equity ownership by outside directors, the higher informativeness of reported earnings. The coefficient EARN*OUTDIR is positive and significant, implying that firms with higher proportion of outside directors have higher informativeness of reported earnings. Consistent with prior studies (Fan and Wong 2002), the coefficient EARN*VOTE is negative and significant, indicating that the separation of control rights from cash flow rights of controlling shareholder reduces the informativeness of reported earnings.

In column (2), I examine the interaction between effectiveness of board monitoring and the divergence of control rights from cash flow rights in affecting earnings informativeness. The interaction term EARN*VOTE* DIROWN is positive and significant at the 1 % level. In firms with high misalignment between control rights from cash flow rights, the informativeness of earnings is higher when outside directors have higher equity ownership. The interaction term EARN*VOTE*OUTDIR is positive and significant at the 5 % level. The negative association between earnings informativeness and the separation between control rights from cash flow rights controlling shareholder is less pronounced in firms with higher proportion of outside directors. In other words, in firms with high misalignment between control rights from cash flow rights, the informativeness of earnings is higher in firms with higher proportion of outside directors.

42.4.5 Robustness Tests

As a sensitivity analysis, I repeat all my tests at the economy level. The economy-by-economy results indicate that earnings quality is positively associated with equity ownership by outside directors and board independence and negatively associated with the separation of cash flow rights from control rights. More importantly, the mitigating effects of equity ownership and board independence on the association between separation of cash flow rights from control rights and earnings quality are not concentrated in any given economy. Year-by-year regressions yield qualitatively similar results, suggesting my inferences are not time-period specific.

As a robustness test, I follow Haw et al. (2004) to include legal institutions that protect minority shareholder rights (proxied by legal tradition, minority shareholder rights, efficiency of judicial system, or disclosure system) and extralegal institutions (proxied by the effectiveness of competition law, diffusion of the press, and tax compliance) in my tests. I continue to document firm-specific internal governance mechanisms, namely, outside directors' equity ownership and board independence, still matter in constraining management opportunism over the financial reporting process, especially in firms with high expected agency costs arising from the divergence between control rights and cash flow rights. Thus, my results suggest that there is an incremental role for firm-specific internal governance mechanisms, beyond country-level institutions, in improving the quality of financial information by mitigating insiders' entrenchment.

42.5 Conclusion

Publicly reported accounting information, which measures a firm's financial position and performance, can be used as important input information in various corporate governance mechanisms such as managerial incentive plans. Whether and how reported accounting information is used in the governance of a firm depends on the quality and credibility of such information. I provide evidence that board of directors plays an important corporate governance role in improving the quality and credibility of accounting information in firms with high agency conflicts arising from their concentrated ownership structure.

I examine the relation among outside directors' equity ownership, board independence, separation of control rights from cash flow rights of controlling shareholder, and earnings quality. I measure earnings quality with three financial reporting metrics: (i) discretionary accruals, (ii) mapping of accruals to cash flow, and (iii) informativeness of reported earnings. I find that earnings quality is positively associated with outside directors' equity ownership and the proportion of outside directors on the board. I document that firms with higher agency problems arising from the separation of control rights from cash flow rights of controlling shareholders have lower earnings quality. The negative association between separation of control rights from cash flow rights and earnings quality is less pronounced in firms with higher equity ownership by outside directors. This finding suggests that equity ownership that aligns outside directors' and shareholders' interest is associated with more effective monitoring of managerial discretion on reported earnings. In addition, the low earnings quality induced by the separation of control rights from cash flow rights is mitigated by the proportion of outside directors on the board. Overall, my results suggest that directors' equity ownership and board independence are associated with better financial reporting outcomes, especially in firms with high expected agency costs arising from misalignment of control rights and cash flow rights.

Appendix 1: Discretionary Accruals

My first proxy for earnings quality is discretionary accruals. A substantial stream of prior studies uses absolute discretionary accruals as a proxy for earnings management (Ashbaugh et al. 2003; Warfield et al. 1995; Klein 2002; Kothari et al. 2005). Absolute discretionary accruals reflect corporate insiders' propensity to inflate reported income to conceal private benefits of control and to understate income in good performance years to create reserves for poor performance in the future. Accruals are estimated by taking the difference between net income and cash flow from operations. I employ the modified cross-sectional Jones (1991) model to decompose total accruals into non-discretionary accruals and discretionary accruals. Specifically, I estimate the following model for each country in each year at the one-digit SIC industry:

$$\mathrm{ACC} = \gamma_1(1/\mathrm{LAG1ASSET}) + \gamma_2(\mathrm{CHGSALE} - \mathrm{CHGREC}) + \gamma_3(\mathrm{PPE}) \quad (42.1)$$

where:
ACC = total accruals, which are calculated as net income minus operating cash flows scaled by beginning-of-year total assets.
LAG1ASSET = total assets at beginning of the fiscal year.
CHGSALE = sales change, which is net sales in year t less net sales in year $t-1$, scaled by beginning-of-year-t total assets.
CHGREC = change in accounts receivables scaled by beginning-of-year-t total assets.
PPE = gross property, plant, and equipment in year t scaled by beginning-of-year-t total assets.

I use the residuals from the annual cross-sectional country-industry regression model in (A1) as the modified Jones model discretionary accruals.

I use the following regression model to test the association between discretionary accruals and board structure:

$$\begin{aligned}\mathrm{DISCAC} = {} & \beta_0 + \beta_1\mathrm{EBC} + \beta_2\mathrm{OUTDIR} + \beta_3\mathrm{VOTE} + \beta_4\mathrm{VOTE} * \mathrm{EBC} \\ & + \beta_5\mathrm{VOTE} * \mathrm{OUTDIR} + \beta_6\mathrm{BOARDSIZE} + \beta_7\mathrm{CEODUAL} \\ & + \beta_8\mathrm{LNASSET} + \beta_9\mathrm{MB} + \beta_{10}\mathrm{LEV} + \beta_{11}\mathrm{ROA} \\ & + \text{Year controls} + \text{Country Controls} \end{aligned} \quad (42.2)$$

where:
DISCAC = absolute value of discretionary accruals estimated based on the modified Jones model (see Eq. 42.1).
DIROWN = common stock and stock options held by outside directors divided by number of ordinary shares outstanding in the firm.
OUTDIR = proportion of outside directors on the board.
VOTE = control rights divided by cash flow rights of the largest controlling shareholder.
BOARDSIZE = number of directors on the board.
CEODUAL = a dummy variable that equals 1 if the CEO is chairman of board and 0 otherwise.
LNASSET = natural logarithm of total assets.
MB = market value of equity divided by book value of equity.
LEV = long-term debt divided by total assets.
ROA = net profit after tax divided by total assets.
Country Controls = a set of country dummy variables.
Year Controls = a set of year dummy variables.

If high equity ownership for outside directors improves the board monitoring of managerial discretion over the financial accounting process, I predict the coefficient β_1 to be negative. Similarly, a negative coefficient for β_2 suggests that board independence curtails managerial opportunism on financial reporting. A positive coefficient β_3 indicates greater separation of control rights from cash flow rights of the largest controlling shareholder induces greater earnings management. I predict the positive association between absolute discretionary accruals and the separation of control rights from cash flow rights to be less pronounced in firms with high equity ownership by outside directors. Thus, I expect coefficient β_4 to be negative. Furthermore, I predict the positive association between absolute discretionary accruals and the separation of control rights from cash flow rights to be less pronounced in firms high proportion of outside directors on the board. Thus, I expect coefficient β_5 to be negative.

Other board characteristics include the total number of directors (BOARDSIZE) and CEO-chairman duality (CEODUAL). The evidence is mixed on whether board size and CEO duality impairs board effectiveness. Thus, ex ante, there is no prediction on the sign on both variables. The model controls for the effects of firm size, growth opportunities, and leverage on discretionary accruals. Large firms have greater external monitoring, have more stable operations and stronger control structures, and hence report smaller abnormal accruals (Dechow and Dichev 2002). Firm size (LNASSET) is measured based on book value of total assets. Because discretionary accruals are higher for firms with higher growth opportunities, I employ the market-to-book equity (MB) ratio to control for the effect of growth opportunities on discretionary accruals (Kothari et al. 2005). I also include financial leverage (LEV), defined as long-term debt divided by total assets, to control for the managerial discretion over the financial accounting process to mitigate constraints of accounting-based debt covenants (Smith and Watts 1992). To control for the effect of firm performance on discretionary accruals, I include firm profitability (ROA), defined as net income divided by total assets. Finally, I include country dummy variables to capture country-specific factors that may affect the development of capital markets and financial accounting quality. I include dummy variables for years and industries to control for time effect and industry effects, respectively.

Appendix 2: Accruals Quality

My second proxy for earnings quality is accruals quality. Dechow and Dichev (2002) propose a measure of earnings quality that captures the mapping of current accruals into last-period, current-period, and next-period cash flows. Francis et al. (2005) find that this measure (which they term accrual quality) is associated with measures of cost of equity capital. My measure of accrual quality is based on Dechow and Dichev's (2002) model relating current accruals to last-period, current-period, and next-period cash flows:

$$\frac{TCA_{j,t}}{Assets} = \gamma_{0,j} + \gamma_{1j}\frac{CFO_{j,t-1}}{Assets} + \gamma_{2j}\frac{CFO_{j,t}}{Assets} + \gamma_{2j}\frac{CFO_{j,t+1}}{Assets} + e_{j,t} \tag{42.3}$$

where:

$TCA_{j,t}$ = firm j's total current accruals in year t = $\Delta CA_{j,t} - \Delta CL_{j,t} - \Delta CASH_{j,t} + \Delta STD_{j,t}$

Assets = firm j's average total assets in year t − 1 and year t

$CFO_{j,t}$ = cash flow from operations in year t is calculated as net income less total accruals (TA) where:

$TA_{j,t} = \Delta CA_{j,t} - \Delta CL_{j,t} - \Delta CASH_{j,t} + \Delta STD_{j,t} - DEPN_{j,t}$ where

$\Delta CA_{j,t}$ = firm j's change in current assets between year t − 1 and year t

$\Delta CL_{j,t}$ = firm j's change in current liabilities between year t − 1 and year t

$\Delta CASH_{j,t}$ = firm j's change in cash between year t − 1 and year t

$\Delta STD_{j,t}$ = firm j's change in debt in current liabilities between year t − 1 and year t

$DEPN_{j,t}$ = firm j's change in depreciation and amortization expense in year t

I estimate Eq. 42.3 for each one-digit SIC industry for each country-year combination. These estimations yield firm- and year-specific residuals, e_{jt}, which form the basis for the accrual quality metric. AQ is the standard deviation of firm j's estimated residuals multiplied by −1. Hence, large values of AQ correspond to high accrual quality.

I employ the following model to test the association between accrual quality and board characteristics:

$$\begin{aligned} AQ = {} & \beta_0 + \beta_1 EBC + \beta_2 OUTDIR + \beta_3 VOTE + \beta_4 VOTE * DIROWN \\ & + \beta_5 VOTE * OUTDIR + \beta_6 CEODUAL + \beta_7 BOARDSIZE \\ & + \beta_8 LNASSET + \beta_9 OPERCYCLE + \beta_{10} NETPPE + \beta_{11} STDSALE \\ & + \beta_{12} STDCFO + \beta_{13} NEGEARN + \text{Country controls} \\ & + \text{Industry Controls} + \text{Year Controls}. \end{aligned} \tag{42.4}$$

where:

AQ = the standard deviation of firm j's residuals from a regression of current accruals on lagged, current, and future cash flows from operations. I multiply the variable by −1 so that higher AQ measure denotes higher accrual quality.

OPERCYCLE = log of the sum of the firm's days accounts receivable and days inventory.

NETPPE = ratio of the net book value of PP&E to total assets.

STDCFO = standard deviation of the firm's rolling 5-year cash flows from operations.

STDSALE = standard deviation of the firm's rolling 5-year sales revenue.

NEGEARN = the firm's proportion of losses over the prior 5 years.

All other variables are previously defined.

If high equity ownership for outside directors improves the board monitoring of managerial discretion over the financial accounting process, I predict coefficient β_1 to be positive. Similarly, a positive coefficient for β_2 suggests that higher board independence is associated with higher accrual quality. If greater agency costs arise from the higher separation of control rights from cash flow rights of the largest controlling shareholder, coefficient β_3 should be negative. I predict that the negative association between accrual quality and the separation of control rights from cash flow rights is mitigated in firms with high equity ownership by outside directors. Thus, I expect coefficient β_4 to be positive. Furthermore, the negative effect of the separation of control rights from cash flow on rights accrual quality should be attenuated in firms with high proportion of outside directors on the board. Thus, I expect coefficient β_5 to be positive.

In Eq. 42.4, the control variables include innate determinants of accrual quality. Briefly, Dechow and Dichev (2002) find that accrual quality is positively associated with firm size and negatively associated with cash flow variability, sales variability, operating cycle, and incidence of losses. Firm size is measured by the natural logarithm of total assets (LNASSET). Operating cycle (OPERCYCLE) is the log of the sum of the firm's days accounts receivable and days inventory. Capital intensity, NETPPE, is proxied by the ratio of the net book value of PP&E to total assets. Cash flow variability (STDCFO) is the standard deviation of the firm's rolling 5-year cash flows from operations. Sales variability (STDSALE) is the standard deviation of the firm's rolling 5-year sales revenue. Incidence of negative earnings realizations, NEGEARN, is measured as the firm's proportion of losses over the prior 5 years.

Appendix 3: Earnings Informativeness

My third proxy of earnings quality is earnings informativeness, measured by the earnings response coefficients (Warfield et al. 1995; Fan and Wong 2002; Francis et al. 2005). The following model is adopted to investigate the relation between earnings informativeness and equity-based compensation, board independence, and separation of control rights from cash flow rights:

$$\begin{aligned}\mathrm{RET} = {} & \beta_0 + \beta_1\mathrm{EARN} + \beta_2\mathrm{EARN} * \mathrm{DIROWN} + \beta_3\mathrm{EARN} * \mathrm{OUTDIR} \\ & + \beta_4\mathrm{EARN} * \mathrm{VOTE} + \beta_5\mathrm{EARN} * \mathrm{VOTE} * \mathrm{EBC} \\ & + \beta_6\mathrm{EARN} * \mathrm{VOTE} * \mathrm{OUTDIR} + \beta_7\mathrm{EARN} * \mathrm{BOARDSIZE} \\ & + \beta_8 * \mathrm{EARN} * \mathrm{CEODUAL} + \beta_9\mathrm{EARN} * \mathrm{LNASSET} \\ & + \beta_{10}\mathrm{EARN} * \mathrm{MB} + \beta_{11}\mathrm{EARN} * \mathrm{LEV} + \beta_{12}\mathrm{EARN} * \mathrm{ROA} \\ & + \text{Year controls} + \text{Country Controls} + \text{Industry Controls} + e\end{aligned} \tag{42.5}$$

where:

RET = 12-month cumulative raw return ending 3 months after the fiscal year-end.

EARN = net income for year t, scaled by the market value of equity at the end of $t - 1$.

All other variables are as previously defined.

The estimated coefficient on β_1 reflects the earnings response coefficient. A positive estimate on β_2 will be consistent with the notion that equity ownership for outside directors is associated with more informative earnings. A positive estimate on β_3 indicates that the greater proportion of outside directors on the board, the greater the informativeness of earnings. From Fan and Wong (2002), I expect coefficient β_4 to be negative, indicating that the reported earnings are less informative when the ultimate shareholder's control rights exceed his cash flow rights. If high equity ownership for outside directors improves their monitoring of management, the negative effects of the divergence of control rights from cash flow rights should be mitigated in firms with high equity-based compensation. I expect coefficient β_5 to be positive. If monitoring intensity is positively associated with the proportion of outside directors on the board, the reduced informativeness of reported earnings in firm with high divergence of control rights from cash flow rights should be mitigated in firms with higher board independence. I expect coefficient β_6 to be positive.

Appendix 4: Adjusting for Standard Errors in Panel Data

Gow et al. (2010) examine several approaches to address issues of cross-sectional and time-series dependence in accounting research. They identified a number of common approaches: Fama-MacBeth, Newey-West, the Z2 statistic, and standard errors clustered by firm, industry, or time. Gow et al. (2010) review each of these approaches and discuss the circumstances in which they produce valid inferences.

This section is drawn heavily from Gow et al. (2010). Correcting for cross-sectional and time-series dependence in accounting research. The Accounting Review 85(2), 483–512. Reader should refer to the paper for details.

(i) **OLS and White Standard Errors**

OLS standard errors assume that errors are both homoskedastic and uncorrelated across observations. While White (1980) standard errors are consistent in the presence of heteroskedasticity, both OLS and White produce misspecified test statistics when either forms of dependence is present.

(ii) **Newey-West**

Newey and West (1987) generalize the White (1980) approach to yield a covariance matrix estimator that is robust to both heteroskedasticity and serial correlation. Gow et al. (2010) find that the Newey-West procedure produces slightly biased estimates of standard errors when time-series dependence alone is present. However, Gow et al. (2010) find that, in the presence of both cross-sectional and time-series dependence, Newey and West method produces misspecified test statistics with even moderate levels of cross-sectional dependence.

(iii) **Fama-MacBeth**

The Fama-MacBeth approach (Fama and MacBeth 1973) is designed to address concerns about cross-sectional correlation. The Fama-MacBeth approach (FM-t) involves estimating T cross-sectional regressions (one for each period) and basing inferences on a t-statistic calculated as

$$t = \frac{\overline{\beta}}{se(\beta)}, \text{where } \overline{\beta} = \frac{1}{T}\sum_{t=1}^{T}\hat{\beta}_t \tag{42.6}$$

and $se(\beta)$ is the standard error of the coefficients based on their empirical distribution. When there is no cross-regression (time-series) dependence, this approach yields consistent estimates of the standard error of the coefficients as T goes to infinity.

Two common variants of the Fama-MacBeth approach appear in the accounting literature. The first variant, FM-i, involves estimating firm- or portfolio-specific time-series regressions with inferences based on the cross-sectional distribution of coefficients. This modification of the Fama-MacBeth approach is appropriate if there is time-series dependence but not cross-sectional dependence. However, FM-i is frequently used when cross-sectional dependence is likely, such as when returns are the dependent variable.

The second common variant of the FM-t approach, FM-NW, is intended to correct for serial correlation in addition to cross-sectional correlation. FM-NW modifies FM-t by applying a Newey-West adjustment in an attempt to correct for serial correlation.

Gow et al. (2010) suggest two reasons to believe that FM-NW may not correct for serial correlation. First, FM-NW involves applying Newey-West to a limited number of observations, a setting in which Newey-West is known to perform poorly. Second, FM-NW applies Newey-West to a time-series of coefficients, whereas the dependence is in the underlying data.

(iv) **Z2 Statistic**

The Z2-t (Z2-i) statistic is calculated using t-statistics from separate cross-sectional (time-series) regressions for each time period (cross-sectional unit) and is given by the expression:

$$Z2 = \frac{\overline{t}}{se(t)}, \text{where } \overline{t} = \frac{1}{T}\sum_{t=1}^{T}\hat{t}_t, \tag{42.7}$$

$se(t)$ is the standard error of the t-statistics based on their empirical distribution, and T is the number of time periods (cross-sectional units) in the sample.

Gow et al. (2010) suggest Z2 may suffer from cross-regression dependence in the same way as the Fama-MacBeth approach does.

(v) **One-Way Cluster-Robust Standard Errors**

A number of studies in our survey use cluster-robust standard errors, with clustering either along a cross-sectional dimension (e.g., analyst, firm, industry, or country) or along a time-series dimension (e.g., year); we refer to the former as CL-i and the latter as CL-t. Cluster-robust standard errors (also referred to as Huber-White or Rogers standard errors) were proposed by White (1980) as a generalization of the heteroskedasticity-robust standard errors of White (1980). With observations grouped into G clusters of Ng observations, for g in $\{1,\ldots,G\}$, the covariance matrix is estimated using the following expression:

$$\hat{V}(\hat{B}) = \left(X'X\right)^{-1}\hat{B}\left(X'X\right)^{-1}, \hat{B} = \sum_{g=1}^{G} X_g' u_g u_g' X_g, \tag{42.8}$$

where Xg is the $Ng \times K$ matrix of regressors, and ug is the Ng-vector of residuals for cluster g.

While one-way cluster-robust standard errors allow for correlation of unknown form within cluster, it is assumed that errors are uncorrelated across clusters. For example, clustering by time (firm) allows observations to be cross-sectionally (serially) correlated but assumes independence over time (across firms). While some studies consider both CL-i and CL-t separately, separate consideration of CL-t and CL-i does not correct for both cross-sectional and time-series dependence. Gow et al. (2010) find that t-statistics for CL-t are inflated in the presence of time-series dependence and t-statistics for CL-i are inflated in the presence of cross-sectional dependence. Thus, when both forms of dependence are present, both CL-t and CL-i produce overstated t-statistics.

(vi) **Two-Way Cluster-Robust Standard Errors**

An extension of cluster-robust standard errors is to allow for clustering along more than one dimension. In contrast to one-way clustering, two-way clustering (CL-2) allows for *both* time-series *and* cross-sectional dependence. For example, two-way clustering by firm and year allows for within-firm (time-series) dependence *and* within-year (cross-sectional) dependence (e.g., the observation for firm j in year t can be correlated with that for firm j in year $t + 1$ and that for firm k in year t). To estimate two-way cluster-robust standard errors, the expression in (A4) is evaluated using clusters along each dimension (e.g., clustered by industry and clustered by year) to yield $V1$ and $V2$. Then the same expression is calculated using the "intersection" clusters (in the example, observations within an industry-year) to yield VI. The two-way cluster-robust estimator V is calculated as $V = V1 + V2 - VI$. Standard econometric software packages (e.g., Stata and SAS) contain routines for calculating one-way cluster-robust standard errors, making it relatively straightforward to implement two-way cluster-robust standard errors. Gow et al. (2010) find that in the presence of both cross-sectional and time-series dependence, the two-way clustering method (by year and by firm) which allows for both cross-sectional and time-series dependence produces well-specified test statistics. Johnston and DiNardo (1997) and Greene (2000) are two econometric textbooks that contain a detailed discussion of the econometrics issues relating to panel data.

Acknowledgment I appreciate the research funding provided by Institute of Certified Public Accountants of Singapore. I would like to dedicate this paper to my late father, Yew-Ming Lee.

References

Ashbaugh, H., LaFond, R., & Mayhew, B. W. (2003). Do nonaudit services compromise auditor independence? Further evidence. *The Accounting Review, 78*(3), 611–639.

Beasley, M. S. (1996). An empirical analysis of the relation between the board of director composition and financial statement fraud. *The Accounting Review, 71*, 443–465.

Claessens, S., Djankov, S., & Lang, L. H. P. (2000). The separation of ownership and control in East Asian corporations. *Journal of Financial Economics, 58*, 81–112.

Claessens, S., Djankov, S., Fan, J. P. H., & Lang, L. H. P. (2002). Disentangling the incentive and entrenchment effects of large shareholdings. *Journal of Finance, 57*(2), 2741–2771.

Dechow, P. M., & Dichev, I. (2002). The quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review, 77*, 35–59.

Dechow, P. M., Sloan, R. G., & Sweeney, A. P. (1996). Causes and consequences of earnings manipulation: An analysis of firms subject to enforcement actions by the SEC. *Contemporary Accounting Research, 13*, 1–36.

Fama, E., & Jensen, M. (1983). Agency problems and residual claims. *Journal of Law and Economics, 26*, 327–349.

Fama, E., & MacBeth, J. (1973). Risk, return, and equilibrium: Empirical tests. *The Journal of Political Economy, 81*(3), 607–636.

Fan, J., & Wong, T. J. (2002). Corporate ownership structure and the informativeness of accounting earnings in East Asia. *Journal of Accounting & Economics, 33*, 401–425.

Fan, J., & Wong, T. J. (2005). Do external auditors perform a corporate governance role in emerging markets? *Journal of Accounting Research, 43*(1), 35–72.

Francis, J., Lafond, R., Olsson, P., & Schipper, K. (2005). The market pricing of accrual quality. *Journal of Accounting & Economics, 39*, 295–327.

Gillette, A., Noe, T., & Rebello, M. (2003). Corporate board composition, protocols and voting behavior: Experimental evidence. *Journal of Finance, 58*, 1997–2032.

Gow, I. D., Ormazabal, G., & Taylor, D. J. (2010). Correcting for cross-sectional and time series dependence in accounting research. *The Accounting Review, 85*(2), 483–512.

Greene, W. (2000). *Econometrics analysis*. Upper Saddle River: Prentice-Hall.

Haw, I., Hu, B., Hwang, L., & Wu, W. (2004). Ultimate ownership, income management, and legal and extra-legal institutions. *Journal of Accounting Research, 42*, 423–462.

Hermalin, B., & Weisbach, M. (1988). The determinants of board composition. *Rand Journal of Economics, 19*, 589–606.

Johnson, S., Boone, P., Breach, A., & Friedman, E. (2000). Corporate governance in the Asian financial crisis. *Journal of Financial Economics, 58*, 141–186.

Johnston, J., & DiNardo, J. (1997). *Econometrics method*. New York: Mc-Graw Hill.

Klein, A. (2002). Audit committee, board of director characteristics, and earnings management. *Journal of Accounting and Economics, 33*(3), 375–400.

Kothari, S. P., Leone, A. J., & Wasley, C. E. (2005). Performance matched discretionary accruals. *Journal of Accounting and Economics, 39*, 163–197.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., & Vishny, R. (1998). Law and finance. *Journal of Political Economy, 106*, 1113–1155.

La Porta, R., Lopez-De-Silanes, F., & Shleifer, A. (1999). Corporate ownership around the world. *Journal of Finance, 54*, 471–518.

Lang, M., Lins, K., & Miller, D. (2004). Concentrated control, analyst following, and valuation: Do analysts matter most when investors are protected least? *Journal of Accounting Research, 42*(3), 589–623.

Lee, K. W. (2007). Corporate voluntary disclosure and the separation of cash flow rights from control rights. *Review of Quantitative Finance and Accounting, 28*, 393–416.
Lee, K. W., Lev, B., & Yeo, H. H. (2007). Organizational structure and earnings management. *Journal of Accounting, Auditing and Finance, 22*(2), 293–391.
Lee, K. W., Lev, B., & Yeo, H. H. (2008). Executive pay dispersion, corporate governance and firm performance. *Review of Quantitative Finance and Accounting, 30*, 315–338.
Lee, C. F., Lee, K. W., & Yeo, H. H. (2009). Investor protection and convertible debt design. *Journal of Banking and Finance, 33*(6), 985–995.
Leuz, C., Nanda, D., & Wysocki, P. (2003). Earnings management and investor protection: An international comparison. *Journal of Financial Economics, 69*, 505–527.
Mehran, H. (1995). Executive compensation structure, ownership, and firm performance. *Journal of Financial Economics, 38*(2), 163–184.
Morck, R., Shleifer, A., & Vishny, R. W. (1988). Management ownership and market valuation: An empirical analysis. *Journal of Financial Economics, 20*, 293–315.
Newey, W., & West, D. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica, 55*(3), 703–708.
Peasnell, K., Pope, P., & Young, S. (2005). Board monitoring and earnings management: Do outside directors influence abnormal accruals? *Journal of Business Finance & Accounting, 32*(7/8), 1311–1345.
Perry, T. (2000). *Incentive compensation for outside directors and CEO turnover*. Working paper, Arizona State University.
Petersen, M. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies, 22*(1), 435–480.
Rosenstein, S., & Wyatt, J. (1990). Outside directors, board independence, and shareholder wealth. *Journal of Financial Economics, 26*, 175–192.
Ryan, H., & Wiggins, R. (2004). Who is whose pocket? Director compensation, board independence, barriers to effective monitoring. *Journal of Financial Economics, 73*, 497–524.
Shivdasani, A. (1993). Board composition, ownership structure, and hostile takeovers. *Journal of Accounting & Economics, 16*, 167–188.
Shleifer, A., & Vishny, R. (1997). A survey of corporate governance. *Journal of Finance, 52*, 737–783.
Smith, C., & Watts, R. (1992). The investment opportunity set and corporate financing, dividend and compensation policies. *Journal of Financial Economics, 32*, 263–292.
Warfield, T., Wild, J. J., & Wild, K. (1995). Managerial ownership, accounting choices, and informativeness of earnings. *Journal of Accounting and Economics, 20*, 61–91.
Weisbach, M. (1988). Outside directors and CEO turnover. *Journal of Financial Economics, 20*, 431–460.
White, H. (1980). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica, 48*(4), 817–838.

Rationality and Heterogeneity of Survey Forecasts of the Yen-Dollar Exchange Rate: A Reexamination

43

Richard Cohen, Carl S. Bonham, and Shigeyuki Abe

Contents

R. Cohen
University of Hawaii Economic Research Organization and Economics, University of Hawaii at Manoa, Honolulu, HI, USA
e-mail: afrc2@cbpp.uaa.alaska.edu

C.S. Bonham (✉)
College of Business and Public Policy, University of Alaska Anchorage, Anchorage, AK, USA
e-mail: bonham@hawaii.edu

S. Abe
Faculty of Policy Studies, Doshisha University, Kyoto, Japan
e-mail: sabe@mail.doshisha.ac.jp

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_43,

Abstract

This chapter examines the rationality and diversity of industry-level forecasts of the yen-dollar exchange rate collected by the Japan Center for International Finance. In several ways we update and extend the seminal work by Ito (1990, *American Economic Review 80*, 434–449). We compare three specifications for testing rationality: the "conventional" bivariate regression, the univariate regression of a forecast error on a constant and other information set variables, and an error correction model (ECM). We find that the bivariate specification, while producing consistent estimates, suffers from two defects: first, the conventional restrictions are sufficient but not necessary for unbiasedness; second, the test has low power. However, before we can apply the univariate specification, we must conduct pretests for the stationarity of the forecast error. We find a unit root in the 6-month horizon forecast error for all groups, thereby rejecting unbiasedness and weak efficiency at the pretest stage. For the other two horizons, we find much evidence in favor of unbiasedness but not weak efficiency. Our ECM rejects unbiasedness for all forecasters at all horizons. We conjecture that these results, too, occur because the restrictions test sufficiency, not necessity.

We extend the analysis of industry-level forecasts to a SUR-type structure using an innovative GMM technique (Bonham and Cohen 2001, *Journal of Business & Economic Statistics, 19*, 278–291) that allows for forecaster cross-correlation due to the existence of common shocks and/or herd effects. Our GMM tests of micro-homogeneity uniformly reject the hypothesis that forecasters exhibit similar rationality characteristics.

Keywords

Rational expectations • Unbiasedness • Weak efficiency • Micro-homogeneity • Heterogeneity • Exchange rate • Survey forecasts • Aggregation bias • GMM • SUR

43.1 Introduction

This chapter examines the rationality of industry-level survey forecasts of the yen-dollar exchange rate collected by the Japan Center for International Finance (JCIF). Tests of rationality take on additional significance when performed on asset market prices, since rational expectations is a necessary condition for market efficiency. In the foreign exchange market, tests of forward rate unbiasedness simultaneously test a zero risk premium in the exchange rate; hence this joint hypothesis is also called the risk-neutral efficient market hypothesis (RNEMH). The practical significance of such a hypothesis is that if the forward rate is indeed an unbiased predictor of the future spot rate, then exchange risk can be costlessly hedged in the forward market. However, the RNEMH has been rejected nearly universally. Since the risk premium is unobservable, insight into the reason for the rejection of the RNEMH can be gained by separately testing for rationality using survey data on expectations. Because forecasters cannot be assumed to have identical information sets, we must use individual survey forecasts to avoid the aggregation bias inherent in the use of mean or median forecasts.

We use data from the same source as Ito (1990), the seminal study recognizing the importance of using individual data to test rationality hypotheses about the exchange rate. To achieve stationarity of the realizations and forecasts (which each have a unit root), Ito (1990) followed the conventional specification at the time of subtracting the current realization from each. These variables are then referred to as being in "change" form. To test unbiasedness he regressed the future rate of depreciation on the forecasted return and tested the joint restrictions that the intercept equalled zero and the slope coefficient equalled one. At the industry level he found approximately twice as many rejections (at the 1 % level) at the longest horizon (6 months) than at the two shorter horizons (1 and 3 months).

We extend Ito's analysis in two principal respects: the specification of unbiasedness tests and inference in tests for micro-homogeneity of forecasters. One problem with the change specification of unbiasedness tests is that, since there is much more variation in the change in the realization than in the forecast, there is a tendency to under-reject the part of the joint hypothesis that the coefficient on the forecast equals one. This is precisely what we would expect in tests of variables which are near random walks.

Second, and more fundamentally, Ito's (1990) bivariate (joint) regression test of unbiasedness is actually a test of sufficiency, not necessity as well as sufficiency. Following Holden and Peel (1990), the necessary and sufficient condition for unbiasedness is a mean zero forecast error. This is tested in a univariate regression by imposing a coefficient of unity on the forecast and testing the restriction that the intercept equals zero. This critique applies whether or not the forecast and realization are integrated in levels. However, when the realization and forecast are integrated in levels, we must conduct a pretest to determine whether the forecast error is stationary. If the forecast and realization are both integrated and cointegrated, then a necessary and sufficient condition for unbiasedness is that intercept and slope in the cointegrating regression (using levels of the realization and forecast) are zero and one, respectively. We test this hypothesis using Liu and Maddala's (1992) method of imposing the (0, 1) vector, then testing the "restricted" cointegrating residual for stationarity.[1,2]

Third, we use the result from Engle and Granger (1987) that cointegrated variables have an error correction representation. First, we employ the specification and unbiasedness restrictions originally proposed by Hakkio and Rush (1989). However, the unbiasedness tests using the ECM specification produce more rejections over industry groups and horizons than the univariate or bivariate specifications.

[1]If in addition the residuals from the cointegrating regression are white noise, this supports a type of weak efficiency.

[2]Pretesting the forecast error for stationarity is a common practice in testing the RNMEH, but the only study we know of that applies this practice to survey forecasts of exchange rates is Osterberg (2000), and he does not test for a zero intercept in the cointegrating regression.

We conjecture that one possible explanation for this apparent anomaly is that, similar to the joint restrictions in the bivariate test, the ECM restrictions test sufficient conditions for unbiasedness, while the univariate restriction only tests a necessary and sufficient condition. Thus, the ECM has a tendency to over-reject. We then respecify the ECM, so that only the necessary and sufficient conditions are tested. We compare our results to those obtained using the sufficient conditions represented by the joint restrictions as well as the necessary and sufficient condition represented by the univariate restriction.

The second direction in which we extend Ito's (1990) analysis has to do with testing for differences among forecasters' ability to produce rational predictions.[3] We recognize, as does Ito, that differences among forecasters over time indicate that at least some individuals form biased forecasts. (The converse does not necessarily hold, since a failure to reject micro-homogeneity could conceivably be due to the same degree of irrationality of each individual in the panel.) Ito's heterogeneity test is a single-equation test of deviations of individual forecasts from the mean forecast, where the latter may or may not be unbiased. In contrast, we test for differences in individual forecast performance using a micro-homogeneity test, i.e., testing for equal coefficients across the system of individual univariate rationality equations.

In our tests for micro-homogeneity, we expect cross-forecaster error correlation due to the possibility of common macro shocks and/or herd effects in expectations. To this end, we incorporate two innovations not previously used by investigators studying survey data on exchange rate expectations. First, in our micro-homogeneity tests, we use a GMM system with a variance-covariance matrix that allows for cross-sectional as well as moving average and heteroscedastic errors. Here we follow the widely used practice of modeling the individual regression residuals as an MA process of order h-1, where h is the number of periods in the forecast horizon. However, no other researchers have actually tested whether an MA process of this length is required to model the cross-sectional behavior of rational forecast errors. Thus, second, to investigate the nature of the actual MA processes, we use Pesaran's (2004) CD test to examine the statistical significance of the cross-sectional dependence of forecast errors, both contemporaneous and lagged.

The organization of the rest of the chapter is as follows: in Sect. 43.2 we review some fundamental issues in testing rationality in the foreign exchange market. In Sects. 43.3 and 43.4, we conduct various rationality tests on the JCIF data. Section 43.5 contains our micro-homogeneity tests. Section 43.6 summarizes and discusses areas for future research.

[3]Market microstructure theories assume that there is a minimum amount of forecaster (as well as cross-sectional forecast) diversity. Also, theories of exchange rate determination that depend upon the interaction between chartists (or noise traders) and fundamentalists by definition require a certain structure of forecaster heterogeneity.

43.2 Background: Testing Rationality in the Foreign Exchange Market

The Rational Expectations Hypothesis (REH) assumes that economic agents know the true data-generating process (DGP) for the forecast variable. This implies that the market's subjective probability distribution of the variable is identical to the objective probability distribution, conditional on a given information set, Φ_t. Equating first moments of the market, $E_m(s_{t+h}|\Phi_t)$, and objective, $E(s_{t+h}|\Phi_t)$, distributions,

$$E_m(s_{t+h}|\Phi_t) = E(s_{t+h}|\Phi_t), \tag{43.1}$$

where the right-hand side can be shortened to $E_t(s_{t+h})$.

It follows that the REH implies that forecast errors have both unconditional and conditional means equal to zero. A forecast is unbiased if its forecast error has an unconditional mean of zero. A forecast is efficient if its error has a conditional mean of zero. The condition that forecast errors be serially uncorrelated is a subset of the efficiency condition where the conditioning information set consists of past values of the realization and current as well as past values of the forecast.[4]

In this chapter we focus on testing whether forecasters can form rational expectations of future depreciation. If not, then at least part of the explanation for the failure of the RNEMH is due to the failure of the REH. There are two related interest parity conditions. Covered interest parity, an arbitrage condition, holds if $f_{t,h} - s_t = i_t - i_t^*$, i.e., the forward premium is equal to the interest differential between domestic and foreign risk-free assets. Uncovered interest parity holds if $s_{t+h} - s_t^e = i_t - i_t^*$. Because uncovered interest parity assumes both unbiased expectations and risk neutrality, some authors view it as equivalent to the RNEMH (see Phillips and Maynard 2001).

The ability to decompose deviations from UIP into time-varying risk premium and systematic forecast error components also has implications for policymakers. Consider first the possibility of a violation of the risk neutrality hypothesis. According to the portfolio balance model, if a statistically significant time-varying risk premium component is found, this means that $i_t - i_t^*$ is time-varying, which in turn implies that foreign and domestic bonds are not perfect substitutes; changes in relative quantities (which are reflected in changes in current account balances) will affect the interest rate differential. In this way, sterilized official intervention can have significant effects on exchange rates. Second, consider the possibility of a violation of the REH. If a statistically significant expectational error of the destabilizing (e.g., "bandwagon") type is found, and policymakers are more rational than speculators, a policy of "leaning against the wind" could have a stabilizing

[4]It is important to note that the result from one type of rationality test does not have implications for the results from any other types of rationality tests. In this chapter we test for unbiasedness and weak efficiency, leaving the more stringent tests of efficiency with respect to publicly available information for future analysis.

effect on exchange rate movements. (See Cavaglia et al. 1994.) More generally, monetary models of the exchange rate (in which the UIP condition is embedded), which assume model-consistent (i.e., rational) expectations with risk neutrality, generally have not performed well empirically, especially in out-of-sample forecasting. (See, e.g., Bryant 1995.) One would like to be able to attribute the model failure to some combination of a failure of the structural assumptions (including risk neutrality) and a failure of the expectational assumption.

43.2.1 Why Test Rational Expectations with Disaggregated Survey Forecast Data?

Beginning with Frankel and Froot (1987) and Froot and Frankel (1989), much of the literature examining exchange rate rationality in general, and the decomposition of deviations from the RNEMH in particular, has employed the representative agent assumption to justify using the mean or median survey forecast as a proxy for the market's expectation. In both studies, Frankel and Froot found significant evidence of irrationality. Subsequent research has found mixed results. Liu and Maddala (1992, p. 366) articulate the mainstream justification for using aggregated forecasts in tests of the REH. "Although . . .data on individuals are important to throw light on how expectations are formed at the individual level, to analyze issues relating to market efficiency, one has to resort to aggregates." In fact, Muth's (1961, p. 316) original definition of rational expectations seemed to allow for the possibility that rationality could be applied to an aggregate (e.g., mean or median) forecast. '. . . [E]xpectations of firms (or, more generally, the subjective probability distribution of outcomes) tend to be distributed, *for the same information set*, about the predictions of the theory (or the "objective" probability distribution of outcomes).' (Emphasis added.)

However, if individual forecasters have different information sets, Muth's definition does not apply. To take the simplest example, the (current) mean forecast is not in any forecaster's information set, since all individuals' forecasts must be made before a mean can be calculated. Thus, current mean forecasts contain private information (see MacDonald 1992) and therefore cannot be tested for rationality.[5]

Using the mean forecast may also result in inconsistent parameter estimates. Figlewski and Wachtel (1983) were the first to show that, in the traditional bivariate

[5]A large theoretical literature relaxes Muth's assumption that all information relevant for forming a rational forecast is publicly available. Instead, this literature examines how heterogeneous individual expectations are mapped into an aggregate market expectation, and whether the latter leads to market efficiency. (See, e.g., Figlewski 1978, 1982, 1984; Kirman 1992; Haltiwanger and Waldman 1989.) Our paper focuses on individual rationality but allows for the possibility of synergism by incorporating not only heteroscedasticity and autocorrelation consistent standard errors in individual rationality tests but also cross-forecaster correlation in tests of micro-homogeneity. The extreme informational requirement of the REH led Pesaran and Weale (2006) to propose a weaker form of the REH that is based on the (weighted) average expectation using only publicly available (i.e., common) information.

unbiasedness equation, the presence of private information variables in the mean forecast error sets up a correlation with the mean forecast. This inconsistency occurs even if all individual forecasts are rational. In addition, Keane and Runkle (1990) pointed out that, when some forecasters are irrational, using the mean forecast may lead to false acceptance of the unbiasedness hypothesis, in the unlikely event that offsetting individual biases allow parameters to be consistently estimated. See also Bonham and Cohen (2001), who argue that, in the case of cointegrated targets and predictions, inconsistency of estimates in rationality tests using the mean forecast can be avoided if corresponding coefficients in the individual rationality tests pass a test for micro-homogeneity.[6] Nevertheless, until the 1990s, few researchers tested for the rationality of individual forecasts, even when those data were available.

43.2.2 Rational Reasons for the Failure of the Rational Expectations Hypothesis Using Disaggregated Data

Other than a failure to process available information efficiently, there are numerous explanations for a rejection of the REH. One set of reasons relates to measurement error in the individual forecast. Researchers have long recognized that forecasts of economic variables collected from public opinion surveys should be less informed than those sampled from industry participants. However, industry participants, while relatively knowledgeable, may not be properly motivated to devote the time and resources necessary to elicit their best responses. The opposite is also possible.[7] Having devoted substantial resources to produce a forecast of the price of a widely traded asset, such as foreign exchange, forecasters may be reluctant to reveal their true forecast before they have had a chance to trade for their own account.[8]

Second, some forecasters may not have the symmetric quadratic loss function embodied in typical measures of forecast accuracy, e.g., minimum mean squared error. (See Zellner 1986; Stockman 1987; Batchelor and Peel 1998.) In this case, the optimal forecast may not be the MSE. In one scenario, related to

[6]The extent to which private information influences forecasts is more controversial in the foreign exchange market than in the equity or bond markets. While Chionis and MacDonald (1997) maintain that there is little or no private information in the foreign exchange market, Lyons (2002) argues that order flow explains much of the variation in prices. To the extent that one agrees with the market microstructure emphasis on the importance of the private information embodied in dealer order flow, the Figlewski-Wachtel critique remains valid in the returns regression.

[7]Elliott and Ito (1999) show that, although a random walk forecast frequently outperforms the JCIF survey forecasts using an MSE criterion, survey forecasts generally outperform the random walk, based on an excess profits criterion. This supports the contention that JCIF forecasters are properly motivated to produce their best forecasts.

[8]To mitigate the confidentiality problem in this case, the survey typically withholds individual forecasts until the realization is known or (as with the JCIF) masks the individual forecast by only reporting some aggregate forecast (at the industry and total level) to the public.

the incentive aspect of the measurement error problem, forecasters may have strategic incentives involving product differentiation.[9]

In addition to strategic behavior, another scenario in which forecasters may deviate from the symmetric quadratic loss function is simply to maximize trading profits. This requires predicting the direction of change, regardless of MSE.[10]

Third, despite their best efforts, forecasters may find it difficult to distinguish between a temporary and permanent shift in the DGP. This difficulty underlies at least three theories of rational forecast errors: the peso problem, learning about past regime changes, and bubbles.

Below we conduct tests for structural change in estimated unbiasedness coefficients. When unbiasedness cannot be rejected, the structural change test may show certain subperiods in which unbiasedness did not hold. In the obverse case, when unbiasedness can be rejected, the structural change test may show certain subperiods in which unbiasedness cannot be rejected. Either situation would lend some support to the theories attributing bias to the difficulty of distinguishing temporary from permanent shifts.

43.3 Description of Data

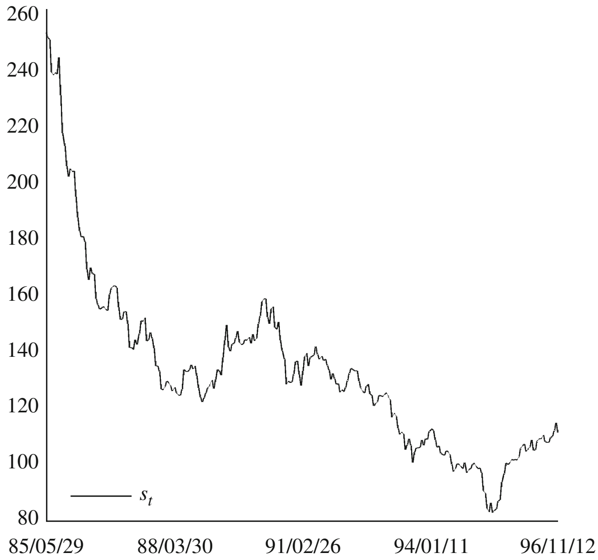
Every 2 weeks, the JCIF in Tokyo conducts telephone surveys of yen/dollar exchange rate expectations from 44 firms. The forecasts are for the future spot rate at horizons of 1 month, 3 months, and 6 months. Our data cover the period May 1985 to March 1996. This data set has very few missing observations, making it close to a true panel. For reporting purposes, the JCIF currently groups individual firms into four industry categories: (1) banks and brokers, (2) insurance and trading companies, (3) exporters, and (4) life insurance companies and importers. On the day after the survey, the JCIF announces overall and industry average forecasts. (For further details concerning the JCIF database, see the descriptions in Ito (1990, 1994), Bryant (1995), and Elliott and Ito (1999).)

Figure 43.1 shows that, over the sample period (one of flexible exchange rates and no capital controls), the yen appreciated dramatically relative to the dollar, from a spot rate of approximately 270 yen/dollar in May 1985 to approximately 90 yen/dollar in March 1996. The path of appreciation was not steady, however. In the first 2 years of the survey alone, the yen appreciated to about 140 per dollar.

[9]Laster et al. (1999) called this practice "rational bias." Prominent references in this growing literature include Lamont (2002), Ehrbeck and Waldmann (1996), and Batchelor and Dua (1990a, b, 1992). Because we have access only to forecasts at the industry average level, we cannot test the strategic incentive hypotheses.

[10]See Elliott and Ito (1999), Boothe and Glassman (1987), LeBaron (2000), Leitch and Tanner (1991), Lai (1990), Goldberg and Frydman (1996), and Pilbeam (1995). This type of loss function may appear to be relevant only for relatively liquid assets such as foreign exchange, but not for macroeconomic flows. However, the directional goal is also used in models to predict business cycle turning points. Also, trends in financial engineering may lead to the creation of derivative contracts in macroeconomic variables, e.g., CPI futures.

Fig. 43.1 Yen-dollar exchange rate versus time



The initial rapid appreciation of the yen is generally attributed to the Plaza meeting in September 1985, in which the Group of Five countries decided to let the dollar depreciate, relative to the other currencies. At the Louvre meeting in February 1987, the Group of Seven agreed to stabilize exchange rates by establishing soft target zones. These meetings may well be interpreted as unanticipated regime changes, since, as we will see below, forecasters generally underestimated the rapid appreciation following the Plaza meeting, then overestimated the value of the yen following the Louvre meeting. Thus, forecasts during these periods may have been subject to peso and learning problems. The period of stabilization lasted until about 1990, when yen appreciation resumed and continued through the end of the sample period.

43.4 Empirical Tests of Rationality

Early studies of the unbiasedness aspect of rationality regressed the level of the realization on the level of the forecast, testing the joint hypothesis that the intercept equalled zero and the slope equalled one.[11] However, since many macroeconomic variables have unit roots, and realizations and forecasts typically share a common stochastic trend, a rational forecast will be integrated and cointegrated with the target series. (See Granger 1991, pp. 69–70.) According to the modern theory of regressions

[11]The efficiency aspect of rationality is sometimes tested by including additional variables in the forecaster's information set, with corresponding hypotheses of zero coefficients on these variables. See, e.g., Keane and Runkle (1990) for a more recent study using the level specification and Bonham and Cohen (1995) for a critique of Keane and Runkle's integration accounting.

with integrated processes (see, *inter alia* Banerjee et al. 1993), conventional OLS estimation and inference produce a slope coefficient that is biased toward one and, therefore, a test statistic that is biased toward accepting the null of unbiasedness. The second generation studies of unbiasedness addressed this inference problem by subtracting the current realization from the forecast as well as the future realization, transforming the levels regression into a "changes" regression. In this specification of stationary variables, unbiasedness was still tested using the same (0, 1) joint hypothesis as in the levels regression. (Ito (1990) is an example of this methodology.) However, an implication of Engle and Granger (1987) is that the levels regression is now interpreted as a cointegrating regression, with conventional t-statistics following nonstandard distributions which depend on nuisance parameters. After establishing that the realization and forecast are integrated and cointegrated, we perform two types of rationality tests. The first is a "restricted cointegration" test due to Liu and Maddala (1992). This is a cointegration test imposing the (0, 1) restriction on the levels regression.

It is significant that, if realization and forecast are cointegrated, Liu and Maddala's (1992) technique is equivalent to regressing a stationary forecast error on a constant and then testing whether the coefficient equals zero (to test unbiasedness) and/or whether the residuals are white noise (to test a type of weak efficiency). Pretests for unit roots in the realization, forecast, and forecast error are required for at least three reasons. First, univariate tests of unbiasedness are invalid if the forecast error is not stationary. Second, following Holden and Peel (1990), we show below (in Sect. 43.4.1.1) that nonrejection of the joint test in the bivariate regression is sufficient but not necessary for unbiasedness, since the joint test is also an implicit test of weak efficiency with respect to the lagged forecast error. A zero intercept in the (correctly specified) univariate test is a necessary as well as sufficient condition for unbiasedness. Third, the Engle and Granger (1987) representation theorem proves that a cointegrating regression such as the levels joint regression (Eq. 43.2 below) has an error correction form that includes both differenced variables and an error correction term in levels. Under the joint null, the error correction term is the forecast error. While the change form of the bivariate regression, is not, strictly speaking, misspecified (since the regressor subtracts s_t, not s^e_{t-1}, from s^e_t), the ECM specification may produce a better fit to the data and, therefore, a more powerful test of the unbiasedness restrictions. We conduct such tests using a form of the ECM due to Hakkio and Rush (1989).

43.4.1 Joint Tests of Unbiasedness and Weak Efficiency

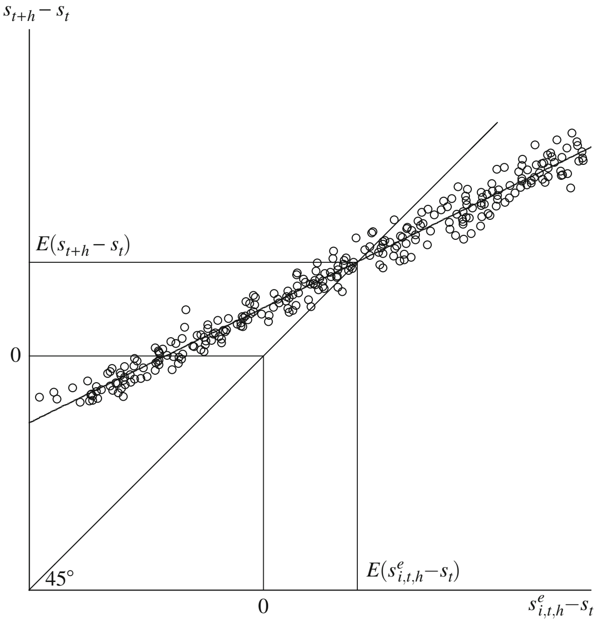
43.4.1.1 The Lack of Necessity Critique

Many, perhaps most, empirical tests of the "unbiasedness" of survey forecasts are conducted using the bivariate regression equation

$$s_{t+h} - s_t = \alpha_{i,h} + \beta_{i,h}\left(s^e_{i,t,h} - s_t\right) + \varepsilon_{i,t,h}. \tag{43.2}$$

It is typical for researchers to interpret their nonrejection of the joint null $(\alpha_{i,h}, \beta_{i,h}) = (0, 1)$ as a necessary condition for unbiasedness. However,

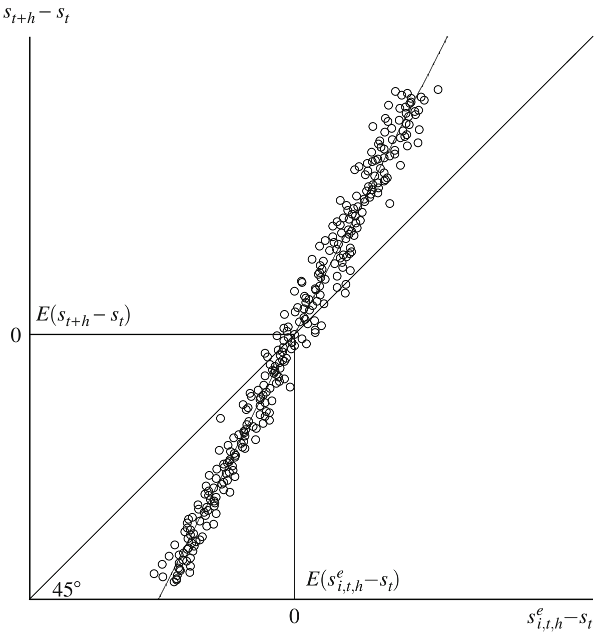
Fig. 43.2 Unbiasedness with $\alpha \neq 0$, $\beta \neq 1$



Holden and Peel (1990) show that this result is a sufficient, though not a necessary, condition for unbiasedness. The intuition for the lack of necessity comes from interpreting the right-hand side of the bivariate unbiasedness regression as a linear combination of two potentially unbiased forecasts: a constant equal to the unconditional mean forecast plus a variable forecast, i.e., $s_{t+h} - s_t = (1 - \beta_{i,h}) \times E(s^e_{i,t,h} - s_t) + \beta_{i,h}(s^e_{i,t,h} - s_t) + \varepsilon_{i,t,h}$. Then the intercept is $\alpha_{i,h} = (1 - \beta_{i,h}) \times E(s^e_{i,t,h} - s_t)$. The necessary and sufficient condition for unbiasedness is that the unconditional mean of the subjective expectation $E[s^e_{i,t,h} - s_t]$ equal the unconditional mean for the objective expectation $E[s_{t+h} - s_t]$. However, this equality can be satisfied without $\alpha_{i,h}$ being equal to zero, i.e., $\beta_{i,h} = 1$.

Figure 43.2 shows that an infinite number of $\alpha_{i,h}$, $\beta_{i,h}$ estimates are consistent with unbiasedness. The only constraint is that the regression line intersects the 45° ray from the origin where the sample mean of the forecast and target are equal. Note that, in the case of differenced variables, this can occur at the origin, so that $\alpha_{i,h} = 0$, but $\beta_{i,h}$ is unrestricted (see Fig. 43.3). It is easy to see why unbiasedness holds: in Figs. 43.2 and 43.3 the sum of all horizontal deviations from the 45° line to the regression line, i.e., forecast errors, equal zero. However, when $\alpha_{i,h} \neq 0$, and $\alpha_{i,h} \neq (1 - \beta_{i,h}) \times E(s^e_{i,t,h} - s_t)$, there is bias regardless of the value of $\beta_{i,h}$. See Fig. 43.4, where the bias, $E(s_{t+h} - s^e_{i,t,h})$, implies systematic underforecasts.

Fig. 43.3 Unbiasedness with $\alpha = 0$, $\beta \neq 1$



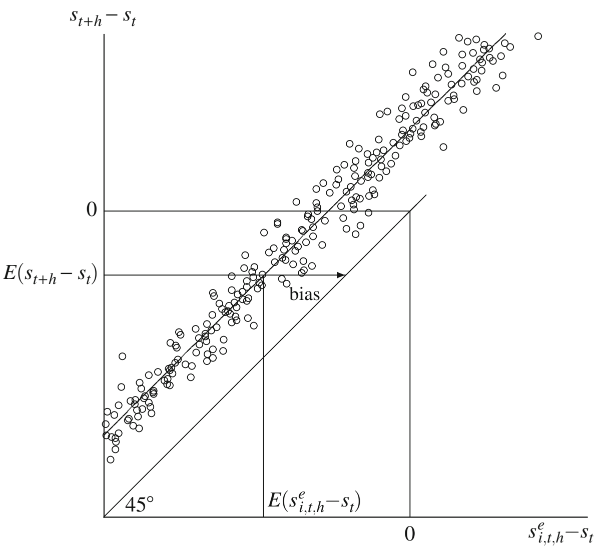


Fig. 43.4 Bias with $\alpha \neq 0$, $\beta = 1$

To investigate the rationality implications of different values for $\alpha_{i,h}$ and $\beta_{i,h}$, we follow Clements and Hendry (1998) and rewrite the forecast error in the bivariate regression framework of Eq. 43.2 as

$$\eta_{i,t,h} = s_{t+h} - s^{e}_{i,t,h} = \alpha_{i,h} + \left(\beta_{i,h} - 1\right)\left(s^{e}_{i,t,h} - s_t\right) + \varepsilon_{i,t,h} \tag{43.3}$$

A special case of weak efficiency occurs when the forecast and forecast error are uncorrelated, i.e.,

$$E\left[\eta_{i,t,h}\left(s^{e}_{i,t,h} - s_t\right)\right] = 0 = \alpha_{i,h}E\left(s^{e}_{i,t,h} - s_t\right) + \left(\beta_{i,h} - 1\right)E\left(s^{e}_{i,t,h} - s_t\right)^2 + E\left[\varepsilon_{i,t,h}\left(s^{e}_{i,t,h} - s_t\right)\right] \tag{43.4}$$

Thus, satisfaction of the joint hypothesis $(\alpha_{i,h}, \beta_{i,h}) = (0,1)$ is also sufficient for weak efficiency with respect to the current forecast. However, it should be noted that Eq. 43.4 may still hold even if the joint hypothesis is rejected. Thus, satisfaction of the joint hypothesis represents sufficient conditions for both unbiasedness and this type of weak efficiency, but necessary conditions for neither.

If $\beta_{i,h} = 1$, then, whether or not $\alpha_{i,h} = 0$, the variance of the forecast error equals the variance of the bivariate regression residual, since then $\text{var}(\eta_{i,t,h}) = (\beta_{i,h} - 1)^2\text{var}(s^e_{i,t,h} - s_t) + \text{var}(\varepsilon_{i,t,h}) + 2(\beta_{i,h} - 1)\text{cov}[(s^e_{i,t,h} - s_t),\varepsilon_{i,t,h}] = \text{var}(\varepsilon_{i,t,h})$. Figure 43.4 illustrates this point. Mincer and Zarnowitz (1969) required only that $\beta_{i,h} = 1$ in their definition of forecast efficiency. If in addition to $\beta_{i,h} = 1$, $\alpha_{i,h} = 0$, then the mean square forecast error also equals the variance of the forecast. Mincer and Zarnowitz emphasized that as long as the loss function is symmetric, as is the case with a minimum mean square error criterion, satisfaction of the joint hypothesis implies optimality of forecasts.

43.4.1.2 Empirical Results of Joint Tests

Since Hansen and Hodrick (1980), researchers have recognized that, when data are sampled more frequently than the forecast horizon (h), forecast errors may follow an h-1 period moving average process. The typical procedure has been to use a variance-covariance matrix which allows for generalized serial correlation. Throughout this chapter, we use the Newey-West (1987) procedure, with the number of lagged residuals set to h-1. To ensure a positive semi-definite VCV matrix, we use a Bartlett window (see Hamilton 1994, pp. 281–84).

In Tables 43.1, 43.2, and 43.3 we report results for the joint unbiasedness tests. We reject the joint hypothesis $(\alpha_{i,h}, \beta_{i,h}) = (0, 1)$ at the 5 % significance level for all groups except banks and brokers at the 1-month horizon (indicating the possible role of inefficiency with respect to the current forecast), but only for the exporters at the 3- and 6-month horizons.

Now consider the results of the separate tests of the joint hypothesis. The significance of the $\alpha_{i,h}$'s in the joint regressions (Eq. 43.2) generally deteriorates

Table 43.1 Joint unbiasedness tests (1-month forecasts)

Individual regressions				
$s_{t+h} - s_t = \alpha_{i,h} + \beta_{i,h}(s^e_{i,t,h} - s_t) + \varepsilon_{i,t,h}$ for $h = 2$				(43.2)
Degrees of freedom = 260				
	i = 1	i = 2	i = 3	i = 4
	Banks and brokers	Insurance and trading companies	Export industries	Life insurance and import companies
$\alpha_{i,h}$	−0.003	−0.004	−0.007	−0.006
t (NW)	−1.123	−1.428	−2.732	−1.903
p-value	0.262	0.153	0.006	0.057
$\beta_{i,h}$	0.437	0.289	−0.318	0.008
t (NW)	1.674	1.382	−1.237	−0.038
R^2	0.014	0.008	0.007	0.000
$H_0 : \beta_{i,h} = 1$, for $i = 1,2,3,4$				
χ^2	4.666	4.666	4.666	4.666
p-value	0.031	0.001	0.000	0.000
Unbiasedness tests: $H_0 : \alpha_{i,h} = 0,\ \beta_{i,h} = 1$, for $i = 1,2,3,4$				
χ^2(NW)	4.696	11.682	29.546	19.561
p-value	0.096	0.003	0.000	0.000
MH tests $H_0 : \alpha_{i,h} = \alpha_{j,h},\ \beta_{i,h} = \beta_{j,h}$ for all $i, j \neq i$				
χ^2(GMM)	9.689			
p-value	0.138			

See Appendix 1 for structure of GMM variance-covariance matrix

with horizon. There are only two rejections at the 5 % level for each of the two shorter horizons. However, the $\alpha_{i,h}$'s are all rejected at the 6.7 % significance level for the 6-month horizon. The test results for the $\beta_{i,h}$'s follow the opposite pattern with respect to horizon. The null that $\beta_{i,h} = 1$ is rejected for all groups at the 1-month horizon, but only for the exporters at the 3-month horizon. There are no rejections at the 6-month horizon. Thus, it appears that the pattern of rejection of the joint hypothesis is predominantly influenced by tests of whether the slope coefficient equals one. In particular, tests of the joint hypothesis at the 1-month horizon are rejected due to failure of this type of weak efficiency, not simple unbiasedness.

For this reason, Mincer and Zarnowitz (1969) and Holden and Peel (1990) suggest that, if one begins by testing the joint hypothesis, rejections in this first stage should be followed by tests of the simple unbiasedness hypothesis in a second stage. Only if unbiasedness is rejected in this second stage should one conclude that forecasts are biased. For reasons described below (in Sect. 43.4.2), our treatment eliminates the first stage, so that unbiasedness and weak efficiency are separately assessed using the forecast error as the dependent variable.

Finding greater efficiency at the longer horizon is unusual, because forecasting difficulty is usually thought to increase with horizon. However, the longer horizon result may not be as conclusive as the $\beta_{i,h}$ statistics suggest. For all tests at all horizons, in only one case can the null hypothesis that $\beta_{i,h}$ equals *zero* not be rejected. Thus, for the longer two horizons (with just the one exception for exporters at the

Table 43.2 Joint unbiasedness tests (3-month forecasts)

Individual regressions				
$s_{t+h} - s_t = \alpha_{i,h} + \beta_{i,h}(s^e_{i,t,h} - s_t) + \varepsilon_{i,t,h}$ for $h = 6$				(43.2)
Degrees of freedom = 256				
	i = 1	i = 2	i = 3	i = 4
	Banks and brokers	Insurance and trading companies	Export industries	Life insurance and import companies
$\alpha_{i,h}$	−0.013	−0.011	−0.017	−0.014
t (NW)	−1.537	−1.362	−2.060	−1.517
p-value	0.124	0.173	0.039	0.129
$\beta_{i,h}$	0.521	0.611	0.082	0.484
t (NW)	1.268	1.868	0.215	1.231
R^2	0.018	0.026	0.001	0.016
$H_0 : \beta_{i,h} = 1$, for $i = 1, 2, 3, 4$				
χ^2	1.362	1.415	5.822	1.728
p-value	0.243	0.234	0.016	0.189
Unbiasedness tests: $H_0 : \alpha_{i,h} = 0$, $\beta_{i,h} = 1$, for $i = 1, 2, 3, 4$				
χ^2(NW)	2.946	2.691	11.156	3.023
p-value	0.229	0.260	0.004	0.221
MH tests $H_0 : \alpha_{i,h} = \alpha_{j,h}$, $\beta_{i,h} = \beta_{j,h}$ for all $i, j \neq i$				
χ^2(GMM)	5.783			
p-value	0.448			

See Appendix 1 for structure of GMM variance-covariance matrix

3-month horizon), hypothesis testing cannot distinguish between the null hypotheses that $\beta_{i,h}$ equals one or zero. Therefore, we cannot conclude that weak efficiency with respect to the current forecast holds while unbiasedness may not. The failure to precisely estimate the slope coefficient also produces R^2s that are below 0.05 in all regressions.[12] The conclusion is that testing only the joint hypothesis has the potential to obscure the difference in performance between the unbiasedness and weak efficiency tests. This conclusion is reinforced by an examination of Figs. 43.5, 43.6, and 43.7, the scatterplots, and regression lines for the bivariate regressions.[13] All three scatterplots have a strong vertical orientation. With this type of data, it is easy to find the vertical midpoint and test whether it is different from zero. Thus, (one-parameter) tests of simple unbiasedness are feasible. However, it is difficult to fit a precisely

[12]As we report in Sect. 43.5, this lack of power is at least consistent with the failure to reject micro-homogeneity at all three horizons.

[13]Note that, for illustrative purposes only, we compute the expectational variable as the four-group average percentage change in the forecast. However, recall that, despite the failure to reject micro-homogeneity at any horizon, the Figlewski-Wachtel critique implies that these parameter estimates are inconsistent in the presence of private information. (See the last paragraph in this subsection.)

Table 43.3 Joint unbiasedness tests (6-month forecasts)

Individual regressions				
$s_{t+h} - s_t = \alpha_{i,h} + \beta_{i,h}(s^e_{i,t,h} - s_t) + \varepsilon_{i,t,h}$ for $h = 12$				(43.2)
Degrees of freedom = 256				
	i = 1	i = 2	i = 3	i = 4
	Banks and brokers	Insurance and trading companies	Export industries	Life insurance and import companies
$\alpha_{i,h}$	−0.032	−0.032	−0.039	−0.034
t (NW)	−1.879	−1.831	−2.099	−1.957
p-value	0.060	0.067	0.036	0.050
$\beta_{i,h}$	0.413	0.822	0.460	0.399
t (NW)	0.761	1.529	0.911	−0.168
R^2	0.01	0.044	0.021	0.012
$H_0 : \beta_{i,h} = 1$, for $i = 1, 2, 3, 4$				
χ^2	1.166	0.110	1.147	1.564
p-value	0.280	0.740	0.284	0.211
Unbiasedness tests: $H_0 : \alpha_{i,h} = 0, \beta_{i,h} = 1$, for $i = 1, 2, 3, 4$				
χ^2(NW)	4.332	3.5	7.899	5.006
p-value	0.115	0.174	0.019	0.082
MH tests $H_0 : \alpha_{i,h} = \alpha_{j,h}, \beta_{i,h} = \beta_{j,h}$ for all $i, j \neq i$				
χ^2(GMM)	7.071			
p-value	0.314			

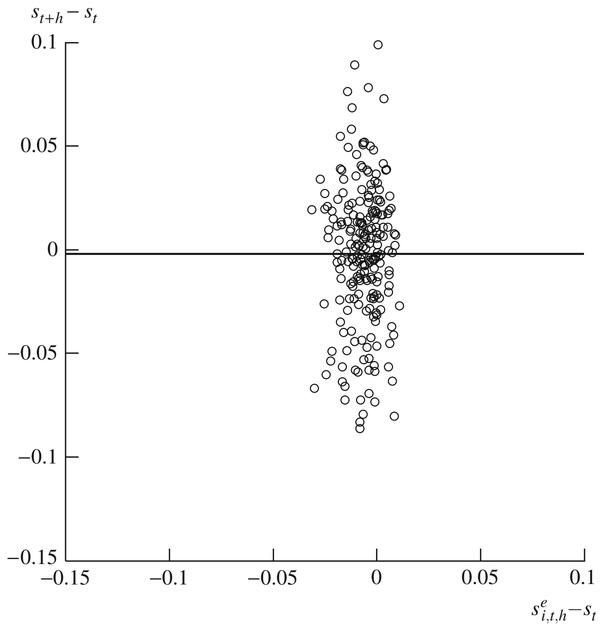
See Appendix 1 for structure of GMM variance-covariance matrix

estimated regression line to this scatter, because the small variation in the forecast variable inflates the standard error of the slope coefficient. This explains why the $\beta_{i,h}$'s are so imprecisely estimated that the null hypotheses that $\beta_{i,h} = 1$ and 0 are simultaneously not rejected. This also explains why the R^2s are so low. Thus, examination of the scatterplots also reveal why bivariate regressions are potentially misleading about weak efficiency as well as simple unbiasedness. Therefore, in contrast to both Mincer and Zarnowitz (1969) and Holden and Peel (1990), we prefer to separate tests for unbiasedness from tests for (all types of) weak efficiency at the initial stage. This obviates the need for a joint test. In the next section, we conduct such tests, making use of cointegration between forecast and realization where it exists.[14]

More fundamentally, the relatively vertical scatter of the regression observations around the origin is consistent with an approximately unbiased forecast of a random

[14]However, in the general case of biased and/or inefficient forecasts, Mincer and Zarnowitz (1969, p. 11) also viewed the bivariate regression 'as a method of correcting the forecasts . . . to improve [their] accuracy . . . Theil (1966, p. 33) called it the "optimal linear correction."' That is, the correction would involve (1) subtracting $\alpha_{i,h}$ and then (2) multiplying by $1/\beta_{i,h}$. Graphically, this is a translation of the regression line followed by a rotation, until the regression line coincides with the 45° line.

Fig. 43.5 Actual versus expected depreciation, 1-month-ahead forecast



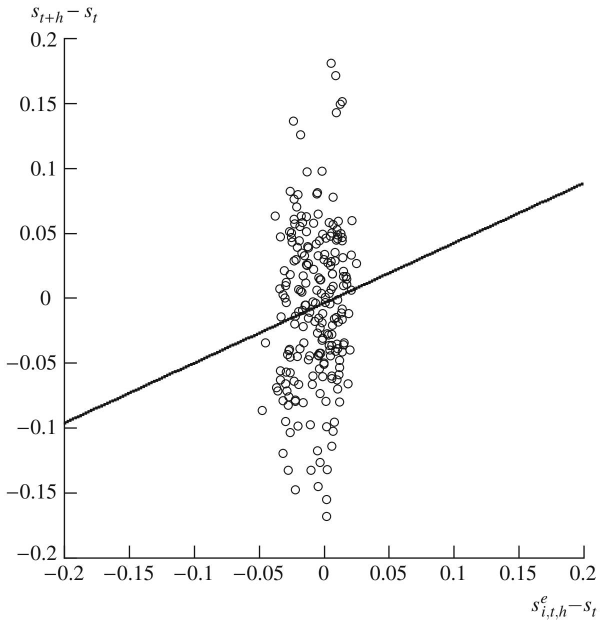
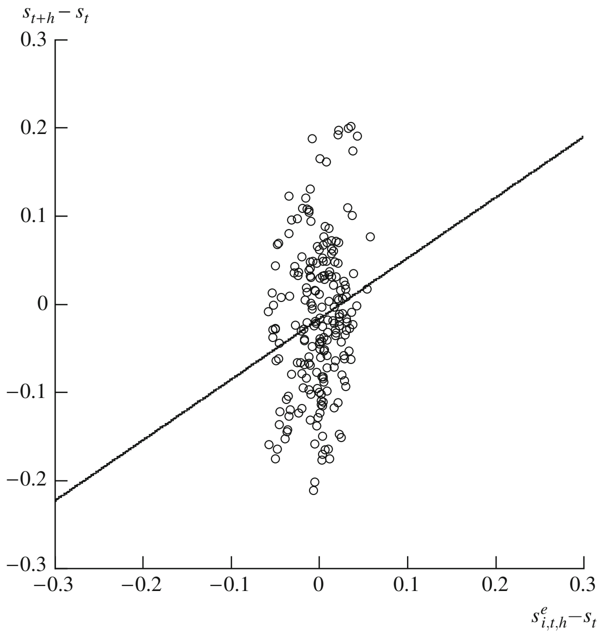


Fig. 43.6 Actual versus expected depreciation, 3-month-ahead forecast

Fig. 43.7 Actual versus expected depreciation: 6-month-ahead forecast



walk-in exchange rate levels.[15] In Figs. 43.8, 43.9, and 43.10, we observe a corresponding time series pattern of variation between the forecasts and realizations in return form. As Bryant lamented in reporting corresponding regressions using a shorter sample from the JCIF, "the regression. . .is. . .not one to send home proudly to grandmother" (Bryant 1995, p. 51). He drew the conclusion that "analysts should have little confidence in a model specification [e.g., uncovered interest parity] setting [the average forecast] exactly equal to the next-period value of the model. . .[M]odel-consistent expectations. . .presume a type of forward-looking behavior [e.g., weak efficiency] that is not consistent with survey data on expectations" (Bryant 1995, p. 40).

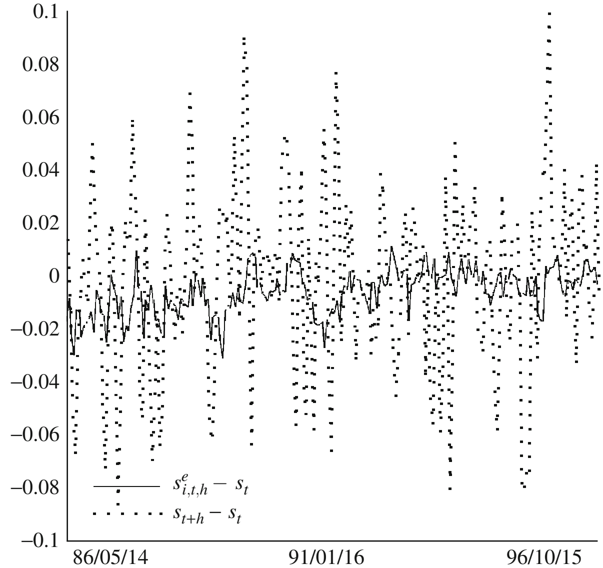
43.4.2 Pretests for Rationality: The Stationarity of the Forecast Error

To test the null hypothesis of a unit root, we estimate the augmented Dickey-Fuller (1979) (ADF) regression

$$\Delta y_{t+1} = \alpha + \beta y_t + \gamma t + \sum_{k=1}^{p} \theta_k \Delta y_{t+1-k} + \varepsilon_{t+1} \tag{43.5}$$

[15]Other researchers (e.g., Bryant 1995) have found similar vertical scatters for regressions where the independent variable, e.g., the forward premium/discount $f_{t,h} - s_t$, the "exchange risk premium" $f_{t,h} - s_{t+h}$, or the difference between domestic and foreign interest rates $(i - i^*)$, exhibits little variation.

Fig. 43.8 Actual and expected depreciation, 1-month-ahead forecast



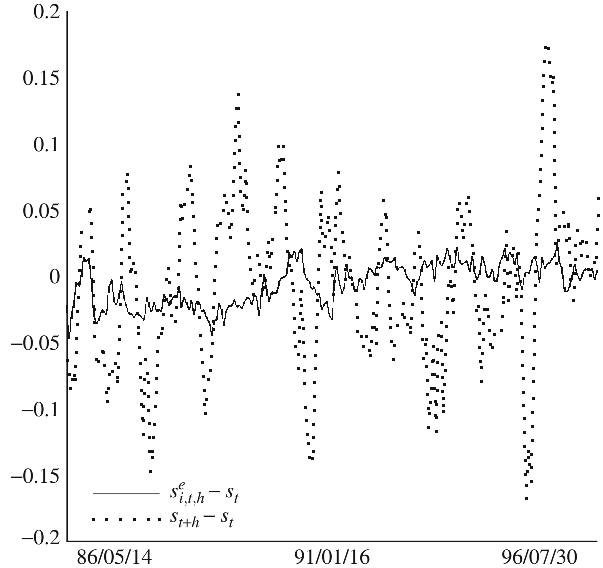


Fig. 43.9 Actual and expected depreciation, 3-month-ahead forecast

where y is the level and first difference of the spot exchange rate, the level and first difference of each group forecast, the residual from the (unrestricted) cointegrating regression, and the forecast error (i.e., the residual from the "restricted" cointegrating equation). The number of lagged differences to include in Eq. 43.5

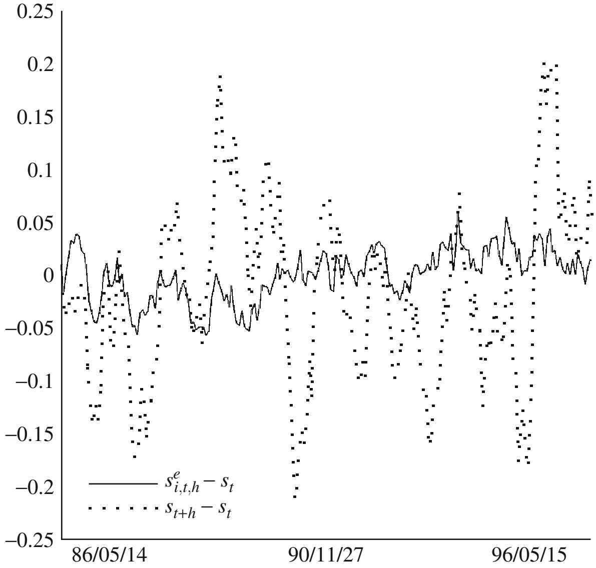


Fig. 43.10 Actual and expected depreciation, 6-month-ahead forecast

is chosen by adding lags until a Lagrange multiplier test fails to reject the null hypothesis of no serial correlation (up to lag 12). We test the null hypothesis of a unit root (i.e., $\beta = 0$) with the ADF t and z tests. We also test the joint null hypothesis of a unit root and no linear trend (i.e., $\beta = 0$ and $\gamma = 0$).

As can be seen in Tables 43.4, 43.5, and 43.6, we fail to reject the null of a unit root in the log of the spot rate in two of the three unit root tests (the exception being the joint null), but we reject the unit root in the h th difference for all three horizons. We conclude that the log of the spot rate is integrated of order one. Similarly, we conclude that the log of the forecast of each spot rate is integrated of order one. Thus, we can conduct cointegration tests on the spot rate and each corresponding forecast. The null of a unit root in the (unrestricted) residual in the "cointegrating regression" is rejected at the 10 % level or less for all groups and horizons except group three (exporters) at the 6-month horizon. Thus, we can immediately reject unbiasedness for the latter group and horizon. Next, since a stationary forecast error is a necessary condition for unbiasedness, we test for unbiasedness (as well as) and weak efficiency in levels using Liu and Maddala's (1992) method of "restricted cointegration." This specification imposes the joint restriction $\alpha_{i,h} = 0$, $\beta_{i,h} = 1$ on the bivariate regression

$$s_{t+h} = \alpha_{i,h} + \beta_{i,h} s^e_{i,t,h} + \varepsilon_{i,t,h} \tag{43.6}$$

and tests whether the residual (the forecast error) is nonstationary. In a bivariate regression, any cointegrating vector is unique. Therefore, if we find that the forecast

Table 43.4 Unit root tests (1-month forecasts $h = 2$)

$\Delta y_t = \alpha + \beta y_{t-1} + \gamma t + \sum_{k=1}^{p} \theta_t \Delta y_{t-k} + \varepsilon_t$ for $h = 2$				(43.5)
	Lags	ADF t test	ADF z test	Joint test
Log of spot rate ($n = 276$)	0	−2.828	−9.660	6.820**
Hth difference log spot rate	12	−4.306***	−167.473***	9.311***
Group 1 Banks and brokers				
Log of forecast	0	−2.274*	−5.072	6.327**
Hth difference log forecast	0	−7.752***	−96.639***	30.085***
Forecast error				
Restricted CI eq.	1	−11.325***	−249.389***	64.676***
Unrestricted CI eq.	1	65.581***		
Group 2 Insurance and trading companies				
Log of forecast	0	−2.735*	−5.149	6.705***
Hth difference log forecast	0	−7.895***	−94.986***	31.252***
Forecast error				
Restricted CI eq.	1	−11.624***	−270.302***	68.053***
Unrestricted CI eq.	1	−11.750***		
Group 3 Export industries				
Log of forecast	1	−2.372	−4.806	5.045**
Hth difference log forecast	0	−8.346***	−111.632***	34.889***
Forecast error				
Restricted CI eq.	1	−10.324***	−211.475***	53.757***
Unrestricted CI eq.	1	−10.392***		
Group 4 Life insurance and import companies				
Log of forecast	0	−2.726*	−5.009	6.438**
Hth difference log forecast	1	−5.216***	−52.911***	3.630***
Forecast error				
Restricted CI eq.	1	−10.977***	−231.837***	60.820***
Unrestricted CI eq.	1	−10.979***		

*Rejection at 10 % level
**Rejection at 5 % level
***Rejection at 1 % level

errors are stationary, then the joint restriction is not rejected, and (0, 1) must be the unique cointegrating vector.[16] The advantage of the one-step restricted cointegration is that if the joint hypothesis is true, then tests which impose this cointegrating vector have greater power than those which estimate a cointegrating vector. See, e.g., Maynard and Phillips (2001).

Note that the Holden and Peel (1990) critique does not apply in the I(1) case, because the intercept cannot be an unbiased forecast of a nonstationary variable. Thus, the cointegrating regression line of the level realization on the level forecast

[16]It is also possible to estimate the cointegrating parameters and jointly test whether they are zero and one. A variety of methods, such as those due to Saikkonen (1991) or Phillips and Hansen (1990), exist that allow for inference in cointegrated bivariate regressions.

Table 43.5 Unit root tests (3-month forecasts $h = 6$)

$\Delta y_t = \alpha + \beta y_{t-1} + \gamma t + \sum_{k=1}^{p} \theta_t \Delta y_{t-k} + \varepsilon_t$ for $h = 6$				(43.5)
	Lags	ADF *t* test	ADF *z* test	Joint test
Log of spot rate ($n = 276$)	0	−2.828	−9.660	6.820**
Hth difference log spot rate	2	−4.760***	−49.769***	11.351***
Group 1 Banks and brokers				
Log of forecast	0	−2.840*	−4.852	7.610***
Hth difference log forecast	0	−5.092***	−48.707***	12.990***
Forecast error				
Restricted CI eq.	6	−3.022**	−29.429***	4.673**
Unrestricted CI eq.	6	−3.343*		
Group 2 Insurance and trading companies				
Log of forecast	0	−2.778*	−4.533	8.858***
Hth difference log forecast	0	−6.514***	−71.931***	21.588***
Forecast error				
Restricted CI eq.	6	−3.068**	−31.038***	4.956**
Unrestricted CI eq.	2	−4.539***		
Group 3 Export industries				
Log of forecast	0	−3.105**	−4.549	9.090***
Hth difference log forecast	1	−4.677***	−41.524***	10.944***
Forecast error				
Restricted CI eq.	6	−3.317**	−31.207***	5.659**
Unrestricted CI eq.	5	−5.115***		
Group 4 Life insurance and import companies				
Log of forecast	0	−2.863*	−4.400	8.161***
Hth difference log forecast	1	−4.324***	−39.870***	9.352***
Forecast error				
Restricted CI eq.	5	−4.825***	−118.586***	11.679***
Unrestricted CI eq.	4	−5.123***		

*Rejection at 10 % level
**Rejection at 5 % level
***Rejection at 1 % level

must have both $\alpha = 0$ and $\beta = 1$ for unbiasedness to hold. This differs from Fig. 43.3, the scatterplot in differences, where $\alpha_{i,h} = 0$ but $\beta_{i,h} \neq 1$. Intuitively, the reason for the difference in results is that the scatterplot in levels must lie in the first quadrant, i.e., no negative values of the forecast or realization.

At the 1-month horizon, the null of a unit root in the residual of the restricted cointegrating regression (i.e., the forecast error) is rejected at the 1 % level for all groups. We find nearly identical results at the 3-month horizon; the null of a unit root in the forecast error is rejected at the 5 % level for all groups. Thus, for these regressions we can conduct rationality tests by regressing the forecast error on a constant (hypothesized equal to zero for unbiasedness) and other information set variables (whose coefficients are hypothesized equal to zero for efficiency). (Recall just above that we failed to reject the null of a unit root in the unrestricted residual

Table 43.6 Unit root tests (6-month forecasts $h = 12$)

$\Delta y_t = \alpha + \beta y_{t-1} + \gamma t + \sum_{k=1}^{p} \theta_t \Delta y_{t-k} + \varepsilon_t$ for $h = 12$				(43.5)
	Lags	ADF *t* test	ADF *z* test	Joint test
Log of spot rate ($n = 276$)	0	−2.828	−9.660	6.820**
Hth difference log spot rate	17	−3.189**	−26.210***	5.500**
Group 1 Banks and brokers				
Log of forecast	0	−2.947**	−4.254	9.131***
Hth difference log forecast	0	−4.772***	−44.018***	11.389***
Forecast error				
Restricted CI eq.	1	−2.373	−13.577*	2.947
Unrestricted CI eq.	7	−3.285**		
Group 2 Insurance and trading companies				
Log of forecast	0	−2.933**	−4.004	9.531***
Hth difference log forecast	0	−6.007***	−64.923***	18.044***
Forecast error				
Restricted CI eq.	1	−2.114	−11.464*	2.399
Unrestricted CI eq.	1	−2.684*		
Group 3 Export industries				
Log of forecast	0	−3.246**	−4.059	10.704***
Hth difference log forecast	12	−4.961***	−44.532***	12.331***
Forecast error				
Restricted CI eq.	12	−1.515	−5.601	1.466
Unrestricted CI eq.	0	−2.931		
Group 4 Life insurance and import companies				
Log of forecast	0	−3.133**	−4.196	9.549***
Hth difference log forecast	0	−4.795***	−44.062***	11.537***
Forecast error				
Restricted CI eq.	2	−2.508	−14.535**	3.148
Unrestricted CI eq.	1	−2.851*		

*Rejection at 10 % level
**Rejection at 5 % level
***Rejection at 1 % level

for the 6-month forecasts of exporters.) Now, in the case of the restricted residual, the other three groups failed to reject a unit root at the 10 % level in two out of three of the unit root tests.[17] (See Figs. 43.11, 43.12, and 43.13.) Thus, in contrast to the results for the two shorter horizons, at the 6-month horizon, the evidence is clearly in favor of a unit root in the forecast error for all four groups. Therefore, we reject the null of simple unbiasedness because a forecast error with a unit root cannot be mean zero. In fact, given our finding of a unit root in the forecast errors, rationality tests regressing the forecast error on a constant and/or other information set variables would be invalid.

[17] As expected, exporters failed to reject at the 10 % level in all three tests.

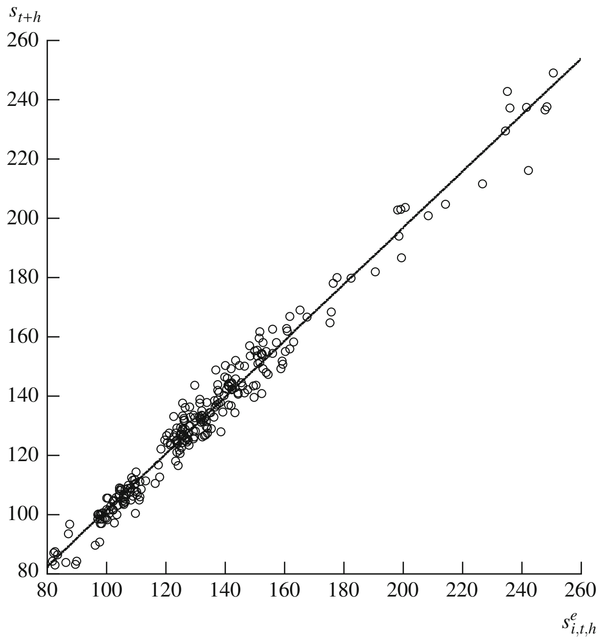


Fig. 43.11 Actual versus expected exchange rate: 1-month-ahead forecast

43.4.3 Univariate Tests for Unbiasedness

The unbiasedness equation is specified as

$$\eta_{i,t,h} = s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \varepsilon_{i,t,h}, \tag{43.7}$$

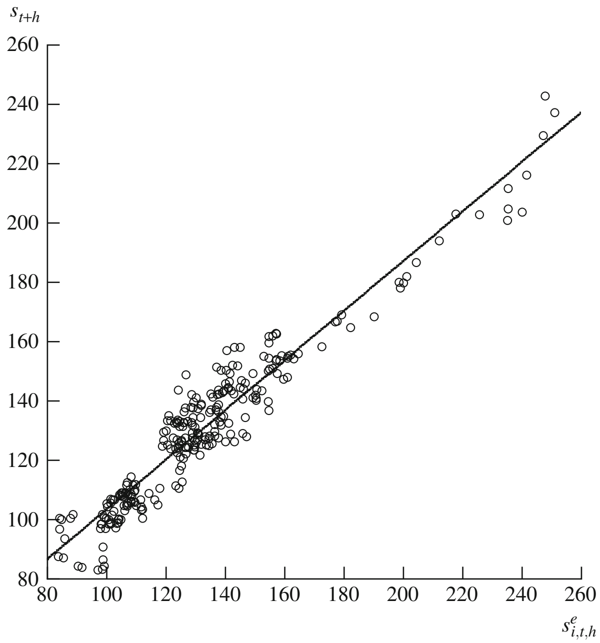
where $\eta_{i,t,h}$ is the forecast error of individual i, for an h-period-ahead forecast made at time t. The results are reported in Tables 43.7 and 43.8. For the 1-month horizon, unbiasedness cannot be rejected at conventional significance levels for any group. For the 3-month horizon, unbiasedness is rejected only for exporters (at a p-value of 0.03). As we saw in the previous subsection, rationality is rejected for all groups at the 6-month horizon, due to nonstationary forecast errors.[18]

In these unbiasedness tests, as well as all others, it is possible that coefficient estimates for the entire sample are not stable over subsamples. The lower panels of Tables 43.7 and 43.8 contain results of the test for equality of intercepts in four equal subperiods, each consisting of approximately 75 biweekly forecasts:

$$\eta_{i,t,h} = s_{t+h} - s^e_{i,t,h} = \alpha_{i,h,1} + \alpha_{i,h,2} + \alpha_{i,h,3} + \alpha_{i,h,4} + \varepsilon_{i,t,h}. \tag{43.8}$$

[18]The direction of the bias for exporters is negative; that is, they systematically underestimate the value of the yen, relative to the dollar. Ito (1990) found the same tendency using only the first two years of survey data (1985–1987). He characterized this depreciation bias as a type of "wishful thinking" on the part of exporters.

Fig. 43.12 Actual versus expected exchange rate: 3-month-ahead forecast



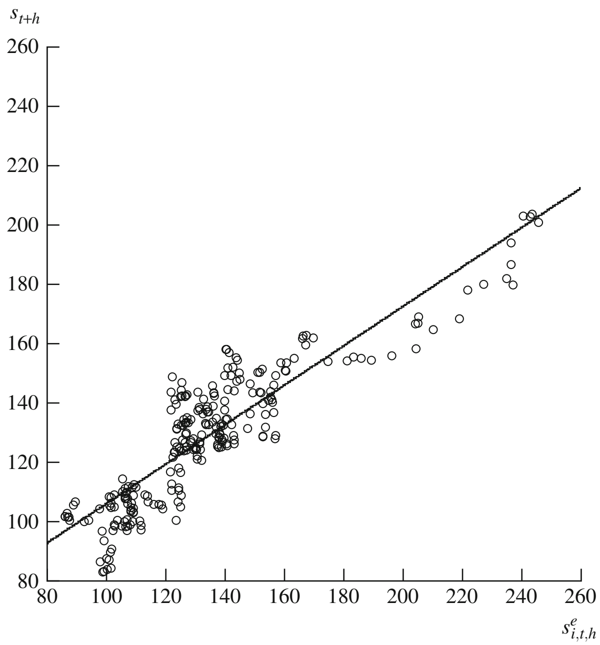


Fig. 43.13 Actual versus expected exchange rate: 6-month-ahead forecast

Table 43.7 Simple unbiasedness tests on individuals (1-month forecasts)

$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \varepsilon_{i,t,h}$				(43.7)
h = 2, degrees of freedom = 261				
	i = 1	i = 2	i = 3	i = 4
	Banks and brokers	Insurance and trading companies	Export industries	Life insurance and import companies
$\alpha_{i,h}$	0.000	0.001	−0.002	0.002
t (NW)	0.115	0.524	−0.809	0.720
p-value	0.909	0.600	0.418	0.472
MH tests $H_0 : \alpha_{i,h} = \alpha_{j,h}$, for all $i, j \neq i$				
χ^2(GMM)	41.643		*p*-value	0.000
$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h,1} + \alpha_{i,h,2} + \alpha_{i,h,3} + \alpha_{i,h,4} + \varepsilon_{i,t,h}$				(43.8)
	i = 1	i = 2	i = 3	i = 4
1985:05:29–1988.03:16	Banks and brokers	Insurance and trading companies	Export industries	Life insurance and import companies
$\alpha_{i,h,1}$	−0.007	−0.005	−0.013	−0.003
p-value	0.213	0.353	0.017	0.573
1988:03:30–1991:01:16				
$\alpha_{i,h,2}$	0.009	0.008	0.010	0.010
p-value	0.134	0.164	0.090	0.114
1991:01:29–1993:11:16				
$\alpha_{i,h,3}$	−0.002	−0.001	−0.004	−0.001
p-value	0.598	0.810	0.374	0.757
1993:11:30–1996:10:15				
$\alpha_{i,h,4}$	0.002	0.004	−0.001	0.003
p-value	0.675	0.433	0.814	0.519
Structural break tests $H_0 : \alpha_{i,h,1} = \alpha_{i,h,2} = \cdots = \alpha_{i,h,4}$				
χ^2	4.245	3.267	8.425	2.946
p-value	0.236	0.352	0.038	0.400

See Appendix 1 for structure of GMM variance-covariance matrix

For both 1- and 3-month horizons, all four forecaster groups undervalued the yen in the first and third subperiods. This is understandable, as both these subperiods were characterized by overall yen appreciation. (See Fig. 43.1.) Evidently, forecasters underestimated the degree of appreciation. Exporters were the only group to undervalue the yen in the last subperiod as well, although that was not one of overall yen appreciation. This is another perspective on the "wishful thinking" of exporters.[19]

The main difference between the two horizons is in the significance of the test for structural breaks. For the 1-month horizon, the estimates of the individual break dummies generally do not reach statistical significance, and the test for their

[19]Ito (1994) conducted a similar analysis for the aggregate of all forecasters, but without an explicit test for structural breaks.

Table 43.8 Simple unbiasedness tests on individuals (3-month forecasts)

$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \varepsilon_{i,t,h}$				(43.7)
h = 6, degrees of freedom = 257				
	i = 1	i = 2	i = 3	i = 4
	Banks and brokers	Insurance and trading companies	Export industries	Life insurance and import companies
$\alpha_{i,h}$	−0.01	−0.008	−0.019	−0.01
t (NW)	−1.151	−0.929	−2.165	−1.121
p-value	0.25	0.353	0.03	0.262
MH tests $H_0 : \alpha_{i,h} = \alpha_{j,h}$, for all $i, j \neq i$				
χ^2(GMM)	40.16		*p*-value	0.000

$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h,1} + \alpha_{i,h,2} + \alpha_{i,h,3} + \alpha_{i,h,4} + \varepsilon_{i,t,h}$				(43.8)
	i = 1	i = 2	i = 3	i = 4
1985:05:29–1988.02:24	Banks and brokers	Insurance and trading companies	Export industries	Life insurance and import companies
$\alpha_{i,h,1}$	−0.040	−0.039	−0.057	−0.039
p-value	0.005	0.007	0.000	0.008
1988:03:16–1990:12:11				
$\alpha_{i,h,2}$	0.023	0.025	0.022	0.021
p-value	0.169	0.111	0.187	0.229
1990:12:25–1993:09:28				
$\alpha_{i,h,3}$	−0.020	−0.018	−0.029	−0.020
p-value	0.064	0.125	0.023	0.094
1993:10:12–1996:0730				
$\alpha_{i,h,4}$	0.000	0.001	−0.013	0.001
p-value	0.994	0.941	0.466	0.970
Structural break test $H_0 : \alpha_{i,h,1} = \alpha_{i,h,2} = \cdots = \alpha_{i,h,4}$				
χ^2	9.319	9.925	13.291	7.987
p-value	0.025	0.019	0.004	0.046

See Appendix 1 for structure of GMM variance-covariance matrix

equality rejects only for the exporters. Thus, the exporters' bias was not constant throughout the sample. In contrast, for the 3-month horizon, the test for no structural breaks is rejected at the 5 % level for all groups, even though unbiasedness itself is rejected for the full sample only for exporters. Even setting aside the bias and variability of exporters' forecasts, our structural break tests allow us to conclude that there is considerably more variation around roughly zero mean forecast errors at the longer horizon. This probably reflects the additional uncertainty inherent in longer-term forecasts.[20]

[20]This is consistent with the finding of nonstationary forecast errors for all groups at the 6-month horizon.

43.4.4 Unbiasedness Tests Using Error Correction Models

As mentioned at the beginning of the previous subsection, the Error Correction Model provides an alternate specification for representing the relationship between cointegrated variables:

$$\begin{aligned} s_{t+h} - s_t = {} & \alpha_{i,h}\left(s_t - \gamma_{i,h} s^e_{i,t-h,h}\right) + \beta_{i,h}\left(s^e_{i,t,h} - s^e_{i,t-h,h}\right) \\ & + \delta_i(\text{lags of } s_{t+h} - s_t) + \eta_i\left(\text{lags of } s^e_{i,t,h} - s^e_{i,t-h,h}\right) + \varepsilon_{i,t,h} \end{aligned} \tag{43.9}$$

According to this specification of the ECM, the change in the spot rate is a function of the change in the forecast, interpreted as a short-run effect, and the current forecast error, interpreted as a long-run adjustment to past disequilibria. $\alpha_{i,h}$, the coefficient of the error correction term, represents the fraction of the forecast error observed at t-h that is corrected by time t. A negative coefficient indicates a stabilizing adjustment of expectations. This formulation of the ECM has the advantage that the misspecification (due to omitted variable bias) of the regression of the differenced future spot rate on the differenced current forecast can be gauged by the statistical significance of the error correction term.[21]

The regressors include the smallest number of lagged dependent variables required such that we do not reject the hypothesis that the residuals are white noise. We impose $\gamma_{i,h} = 1$ when "restricted" cointegration of s_{t+h} and $s^e_{i,t,h}$ is not rejected. Recall that 1- and 3-month forecast errors were found to be stationary, so it was for these two horizons that estimation of the simple unbiasedness equation was possible. Although it would be valid to estimate the ECM at the 6-month horizon using the (unrestricted) stationary cointegrating residual (i.e., for all groups but exporters), we elect not to, because the nonstationarity of the forecast error itself implies a failure of the unbiasedness restrictions.[22]

Then, as first asserted by Hakkio and Rush (1989), the unbiasedness restriction is represented by the joint hypothesis that $-\alpha_{i,h} = \beta_{i,h} = 1$ and all δ and η coefficients equal zero.[23] (The hypothesized coefficient on the error correction term of -1

[21]Zacharatos and Sutcliffe (2002) note that the inclusion of the contemporaneous spot forecast (in their paper, the forward rate) as a regressor assumes that the latter is weakly exogenous; that is, deviations from unbiasedness are corrected only by movements in the realized spot rate. These authors prefer a bivariate ECM specification, in which the change in the future spot rate and the change in the contemporaneous forecast are functions of an error correction term and lags of the dependent variables. However, Zivot (2000) points out that if the spot rate and forecast are contemporaneously correlated, then our single-equation specification does not make any assumptions about the weak exogeneity of the forecast.

[22]Our empirical specification of the ECM also includes an intercept. This will help us to determine whether there are structural breaks in the ECM.

[23]Since we include an intercept, we also test the restriction that the intercept equals zero – both individually and as part of the joint unbiasedness hypothesis.

reflects the unbiasedness requirement that the entire forecast error is corrected within the forecast horizon h.) We also test unbiasedness without including lagged dependent variables but incorporating robust standard errors which allow for generalized serial correlation and heteroscedasticity. This allows comparison with the univariate and bivariate unbiasedness equations.

First, we compare the ECM results to the joint unbiasedness restrictions in the change regressions, using robust standard errors in both cases. Although the estimated coefficient of the error correction term is generally negative, indicating a stable error correction mechanism,[24] the coefficient does not reach a 5 % significance level in any of the regressions. Thus, there is little evidence that the error correction term plays a significant role in the long-run dynamics of exchange rate changes. The ECM test results are nearly identical to the joint unbiasedness test results in Table 43.9. In both specifications, unbiasedness is rejected for three of four groups at the 1-month horizon and not rejected for three of four groups at the 3-month horizon. However, even though the EC term is not (individually) significant in the ECMs, it does provide explanatory power, relative to the joint unbiasedness specification. The R^2s in the ECM, while never more than 0.044, still are greater than in the joint unbiasedness specification, typically by factors of three to five (Tables 43.10, 43.11, and 43.12).

Second, we compare the ECM results to the univariate simple unbiasedness regressions, again using robust standard errors in both cases. The ECM unbiasedness restrictions are rejected at a 5 % level more often than in the simple unbiasedness tests. Whereas the only rejection of simple unbiasedness at the shorter two horizons is for exporters at the 3-month horizon, the ECM restrictions are rejected for three out of four groups at the 1-month horizon as well as for exporters at the 3-month horizon.

While it is uncontroversial that, for testing unbiasedness, the ECM is preferred to the conventional bivariate specification in returns, it is not at all clear that the ECM is preferred to the simple univariate test of unbiasedness. Can the more decisive rejections of unbiasedness using the ECM versus the simple univariate specification be reconciled?[25]

One way to proceed is to determine whether the unbiasedness restrictions imposed on the ECM are necessary as well as sufficient, as is the case for the simple unbiasedness test, or just sufficient, as is the case for the bivariate unbiasedness test. Thus, it is possible that the stronger rejections of unbiasedness in the ECM specification are due to the implicit test of weak efficiency with respect to the current forecast. That is, the Holden and Peel (1990) critique applies to the Hakkio and Rush (1989) test in Eq. 43.9, as well as the joint unbiasedness test in the returns regression. Setting $\beta_{i,h}$, the coefficient of the contemporaneous differenced forecast, equal to one produces an ECM in which the dependent variable is the forecast error:

[24]The only exception is for exporters at the 1-month horizon.

[25]The standard errors in the univariate regression are about the same as those for the ECM. (By definition, of course, the R^2s for the univariate regression equal zero.)

Table 43.9 Error correction models (1-month forecasts)

Group 1 Banks and brokers				
$s_{t+h} - s_t = c_i + \alpha_i(s_t - \gamma_i s^e_{i,t-h,h}) + \beta_i(s^e_{i,t,h} - s^e_{i,t-h,h}) + \varepsilon_{i,h}$				(43.9)
With robust standard errors ($R^2 = 0.0195$)	Coeff	$\chi^2(n)$	n	p-value
Constant	−0.002	0.377	1	0.539
$\alpha_{i,h} = 0$	−0.465	2.884	1	0.089
$\alpha_{i,h} = -1$	−0.465	3.813	1	0.051
$\beta_{i,h} = 1$	0.491	3.847	1	0.050
Constant = 0 and $\alpha_{i,h} = -1$ and $\beta_{i,h} = 1$		3.910	3	0.271
With whitened residuals ($R^2 = 0.605$)	Coeff	$F(n,df)$	n	p-value
Constant	−0.001	0.238	1	0.627
$\alpha_{i,h} = -1$	−0.025	14.696	1	0.000
$\beta_{i,h} = 1$	0.453	4.895	1	0.028
Constant = 0 and $\alpha_{i,h} = -1$ and $\beta_{i,h} = 1$		6.790	3	0.000
Constant = 0 and $\alpha_{i,h} = -1$ and $\beta_{i,h} = 1$ and all lags of realizations and forecasts = 0		51.115	12	0.000
$s_{t+h} - s^e_{i,t,h} = d_i + (1 + \alpha_{i,h})(s_t - s^e_{i,t-h,h})$				(43.10)
$\beta_{i,h} = 1$ imposed, robust standard errors	Coeff	$\chi^2(n)$	n	p-value
Constant	0.002	0.293	1	0.582
$\alpha_{i,h} = -1$	−0.991	0.015	1	0.903
Constant = 0 and $\alpha_{i,h} = -1$		0.294	2	0.863
Group 2 Insurance and trading companies				
$s_{t+h} - s_t = c_i + \alpha_i(s_t - \gamma_i s^e_{i,t-h,h}) + \beta_i(s^e_{i,t,h} - s^e_{i,t-h,h}) + \varepsilon_{i,h}$				(43.9)
With robust standard errors ($R^2 = 0.009$)	Coeff	$\chi^2(n)$	n	p-value
Constant	−0.003	1.168	1	0.280
$\alpha_{i,h} = 0$	−0.262	1.111	1	0.292
$\alpha_{i,h} = -1$	−0.262	8.807	1	0.003
$\beta_{i,h} = 1$	0.278	10.703	1	0.001
Constant = 0 and $\alpha_{i,h} = -1$ and $\beta_{i,h} = 1$		10.965	3	0.012
With whitened residuals ($R^2 = 0.596$)	Coeff	$F(n,df)$	n	p-value
Constant	−0.002	0.563	1	0.454
$\alpha_{i,h} = -1$	−0.036	13.614	1	0.000
$\beta_{i,h} = 1$	0.207	12.639	1	0.001
Constant = 0 and $\alpha_{i,h} = -1$ and $\beta_{i,h} = 1$		6.792	3	0.000
Constant = 0 and $\alpha_{i,h} = -1$ and $\beta_{i,h} = 1$ and all lags of realizations and forecasts = 0		54.557	11	0.000
$s_{t+h} - s^e_{i,t,h} = d_i + (1 + \alpha_{i,h})(s_t - s^e_{i,t-h,h})$				(43.10)
$\beta_{i,h} = 1$ imposed, robust standard errors	Coeff	$\chi^2(n)$	n	p-value
Constant	0.003	0.787	1	0.3751
$\alpha_{i,h} = -1$	−1.026	0.113	1	0.7368
Constant = 0 and $\alpha_{i,h} = -1$		1.052	2	0.591

Table 43.10 Error correction models (1-month forecasts)

Group 3 Export industries				
$s_{t+h} - s_t = c_i + \alpha_i(s_t - \gamma_i s^e_{i,t-h,h}) + \beta_i(s^e_{i,t,h} - s^e_{i,t-h,h}) + \varepsilon_{i,h}$				(43.9)
With robust standard errors ($R^2 = 0.009$)	Coeff	$\chi^2(n)$	n	p-value
Constant	−0.006	4.202	1	0.040
$\alpha_{i,h} = 0$	0.305	1.335	1	0.248
$\alpha_{i,h} = -1$	0.305	24.402	1	0.000
$\beta_{i,h} = 1$	−0.256	20.516	1	0.000
Constant = 0 and $\alpha_{i,h} = -1$ and $\beta_{i,h} = 1$		27.207	3	0.000
With whitened residuals ($R^2 = 0.602$)	Coeff	$F(n, df)$	n	p-value
Constant	−0.002	0.632	1	0.428
$\alpha_{i,h} = -1$	0.107	18.043	1	0.000
$\beta_{i,h} = 1$	−0.055	17.987	1	0.000
Constant = 0 and $\alpha_{i,h} = -1$ and $\beta_{i,h} = 1$		8.455	3	0.000
Constant = 0 and $\alpha_{i,h} = -1$ and $\beta_{i,h} = 1$ and all lags of realizations and forecasts = 0		55.054	10	0.000
$s_{t+h} - s^e_{i,t,h} = d_i + (1 + \alpha_{i,h})(s_t - s^e_{i,t-h,h})$				(43.10)
$\beta_{i,h} = 1$ imposed, robust standard errors	Coeff	$\chi^2(n)$	n	p-value
Constant	−0.001	0.082	1	0.7749
$\alpha_{i,h} = -1$	−0.887	2.321	1	0.1277
Constant = 0 and $\alpha_{i,h} = -1$		2.578	2	0.276
Group 4 Life Insurance and import companies				
$s_{t+h} - s_t = c_i + \alpha_i(s_t - \gamma_i s^e_{i,t-h,h}) + \beta_i(s^e_{i,t,h} - s^e_{i,t-h,h}) + \varepsilon_{i,h}$				(43.9)
With robust standard errors ($R^2 = 0.003$)	Coeff	$\chi^2(n)$	n	p-value
Constant	−0.004	1.734	1	0.188
$\alpha_{i,h} = 0$	−0.066	0.083	1	0.773
$\alpha_{i,h} = -1$	−0.066	16.501	1	0.000
$\beta_{i,h} = 1$	0.112	16.086	1	0.000
Constant = 0 and $\alpha_{i,h} = -1$ and $\beta_{i,h} = 1$		17.071	3	0.001
With whitened residuals ($R^2 = 0.607$)	Coeff	$F(n, df)$	n	p-value
Constant	−0.002	0.392	1	0.532
$\alpha_{i,h} = -1$	−0.026	20.268	1	0.000
$\beta_{i,h} = 1$	0.226	12.020	1	0.001
Constant = 0 and $\alpha_{i,h} = -1$ and $\beta_{i,h} = 1$		10.794	3	0.000
Constant = 0 and $\alpha_{i,h} = -1$ and $\beta_{i,h} = 1$ and all lags of realizations and forecasts = 0		57.702	11	0.000
$s_{t+h} - s^e_{i,t,h} = d_i + (1 + \alpha_{i,h})(s_t - s^e_{i,t-h,h})$				(43.10)
$\beta_{i,h} = 1$ imposed, robust standard errors	Coeff	$\chi^2(n)$	n	p-value
Constant	0.003	1.254	1	0.2629
$\alpha_{i,h} = -1$	−0.949	0.481	1	0.4879
Constant = 0 and $\alpha_{i,h} = -1$		1.628	2	0.443

Table 43.11 Error correction models (3-month forecasts)

Group 1 Banks and brokers				
$s_{t+h} - s_t = c_i + \alpha_i(s_t - \gamma_i s^e_{i,t-h,h}) + \beta_i(s^e_{i,t,h} - s^e_{i,t-h,h}) + \varepsilon_{i,h}$				(43.9)
Tests with robust standard errors ($R^2 = 0.036$)	Coeff	$\chi^2(n)$	n	p-value
Constant	−0.010	1.306	1	0.253
$\alpha_{i,h} = 0$	−0.377	0.590	1	0.443
$\alpha_{i,h} = -1$	−0.377	1.604	1	0.205
$\beta_{i,h} = 1$	0.501	1.268	1	0.260
Constant $= 0$ and $\alpha_{i,h} = -1$ and $\beta_{i,h} = 1$		2.348	3	0.503
Tests with whitened residuals ($R^2 = 0.863$)	Coeff	$F(n, df)$	n	p-value
Constant	−0.008	4.173	1	0.044
$\alpha_{i,h} = -1$	0.233	56.755	1	0.000
$\beta_{i,h} = 1$	−0.178	51.113	1	0.000
Constant $= 0$ and $\alpha_{i,h} = -1$ and $\beta_{i,h} = 1$		28.974	3	0.000
Constant $= 0$ and $\alpha_{i,h} = -1$ and $\beta_{i,h} = 1$ and all lags of realizations and forecasts $= 0$		121.851	14	0.000
$s_{t+h} - s^e_{i,t,h} = d_i + (1 + \alpha_{i,h})(s_t - s^e_{i,t-h,h})$				(43.10)
$\beta_{i,h} = 1$ imposed, robust standard errors	Coeff	$\chi^2(n)$	n	p-value
Constant	−0.006	0.556	1	0.456
$\alpha_{i,h} = -1$	−0.889	0.896	1	0.344
Constant $= 0$ and $\alpha_{i,h} = -1$		1.330	2	0.514
Group 2 Insurance and trading companies				
$s_{t+h} - s_t = c_i + \alpha_i(s_t - \gamma_i s^e_{i,t-h,h}) + \beta_i(s^e_{i,t,h} - s^e_{i,t-h,h}) + \varepsilon_{i,h}$				(43.9)
Tests with robust standard errors ($R^2 = 0.044$)	Coeff	$\chi^2(n)$	n	p-value
Constant	−0.008	0.874	1	0.350
$\alpha_{i,h} = 0$	−0.556	2.061	1	0.151
$\alpha_{i,h} = -1$	−0.556	1.310	1	0.252
$\beta_{i,h} = 1$	0.663	0.965	1	0.326
Constant $= 0$ and $\alpha_{i,h} = -1$ and $\beta_{i,h} = 1$		1.833	3	0.608
Tests with whitened residuals ($R^2 = 0.844$)	Coeff	$F(n, df)$	n	p-value
Constant	−0.005	1.400	1	0.239
$\alpha_{i,h} = -1$	0.080	31.425	1	0.000
$\beta_{i,h} = 1$	−0.167	40.346	1	0.000
Constant $= 0$ and $\alpha_{i,h} = -1$ and $\beta_{i,h} = 1$		23.551	3	0.000
Constant $= 0$ and $\alpha_{i,h} = -1$ and $\beta_{i,h} = 1$ and all lags of realizations and forecasts $= 0$		148.338	10	0.000
$s_{t+h} - s^e_{i,t,h} = d_i + (1 + \alpha_{i,h})(s_t - s^e_{i,t-h,h})$				(43.10)
$\beta_{i,h} = 1$ imposed, robust standard errors	Coeff	$\chi^2(n)$	n	p-value
Constant	−0.005	0.291	1	0.589
$\alpha_{i,h} = -1$	−0.897	0.773	1	0.379
Constant $= 0$ and $\alpha_{i,h} = -1$		0.945	2	0.623

Table 43.12 Error correction models (3-month forecasts)

Group 3 Export industries				
$s_{t+h} - s_{i,t,h} = c_i + \alpha_i(s_t - \gamma_i s^e_{i,t-h,h}) + \beta_i(s^e_{i,t,h} - s^e_{i,t-h,h}) + \varepsilon_{i,h}$				
Tests with robust standard errors ($R^2 = 0.026$)	Coeff	$\chi^2(n)$	n	p-value
Constant	−0.013	2.303	0	0.129
$\alpha_{i,h} = 0$	−0.253	0.393	1	0.531
$\alpha_{i,h} = -1$	−0.253	3.422	1	0.064
$\beta_{i,h} = 1$	0.411	2.102	1	0.147
Constant = 0 and $\alpha_{i,h} = -1$ and $\beta_{i,h} = 1$		7.663	3	0.054
Tests with whitened residuals ($R^2 = 0.856$)	Coeff	$F(n, df)$	n	p-value
Constant	−0.003	0.840	1	0.361
$\alpha_{i,h} = -1$	−0.006	29.512	1	0.000
$\beta_{i,h} = 1$	−0.205	40.582	1	0.000
Constant = 0 and $\alpha_{i,h} = -1$ and $\beta_{i,h} = 1$		16.290	3	0.000
Constant = 0 and $\alpha_{i,h} = -1$ and $\beta_{i,h} = 1$ and all lags of realizations and forecasts = 0		182.912	10	0.000
$s_{t+h} - s^e_{i,t,h} = d_i + (1 + \alpha_{i,h})(s_t - s^e_{i,t-h,h})$				(43.10)
$\beta_{i,h} = 1$ imposed, robust standard errors	Coeff	$\chi^2(n)$	n	p-value
Constant	−0.012	1.971	1	0.160
$\alpha_{i,h} = -1$	−0.775	3.791	1	0.052
Constant = 0 and $\alpha_{i,h} = -1$		6.337	2	0.042
Group 4 Life insurance and import companies				
$s_{t+h} - s_{i,t,h} = c_i + \alpha_i(s_t - \gamma_i s^e_{i,t-h,h}) + \beta_i(s^e_{i,t,h} - s^e_{i,t-h,h}) + \varepsilon_{i,h}$				
Tests with robust standard errors ($R^2 = 0.038$)	Coeff	$\chi^2(n)$	n	p-value
Constant	−0.009	0.993	1	0.319
$\alpha_{i,h} = 0$	−0.478	1.250	1	0.264
$\alpha_{i,h} = -1$	−0.478	1.488	1	0.223
$\beta_{i,h} = 1$	0.604	0.919	1	0.338
Constant = 0 and $\alpha_{i,h} = -1$ and $\beta_{i,h} = 1$		2.451	3	0.484
Tests with whitened residuals ($R^2 = 0.845$)	Coeff	$F(n, df)$	n	p-value
Constant	−0.003	0.510	1	0.477
$\alpha_{i,h} = -1$	0.050	32.000	1	0.000
$\beta_{i,h} = 1$	0.062	32.469	1	0.000
Constant = 0 and $\alpha_{i,h} = -1$ and $\beta_{i,h} = 1$		21.673	3	0.000
Constant = 0 and $\alpha_{i,h} = -1$ and $\beta_{i,h} = 1$ and all lags of realizations and forecasts = 0		169.286	9	0.000
$s_{t+h} - s^e_{i,t,h} = d_i + (1 + \alpha_{i,h})(s_t - s^e_{i,t-h,h})$				(43.10)
$\beta_{i,h} = 1$ imposed, robust standard errors	Coeff	$\chi^2(n)$	n	p-value
Constant	−0.006	0.455	1	0.500
$\alpha_{i,h} = -1$	−0.865	1.405	1	0.236
Constant = 0 and $\alpha_{i,h} = -1$		1.726	2	0.422

$$s_t - s^e_{i,t,h} = \left(1 + \alpha_{i,h}\right)\left(s_t - s^e_{i,t-h,h}\right) \tag{43.10}$$

Thus, in the ECM the necessary and sufficient condition for unbiasedness is that $\alpha_{i,h}$ equals -1.[26] Table 43.9 contains tests of this conjecture. Here the joint hypothesis that the intercept equals zero and $\alpha_{i,h}$ equals minus one produces exactly the same results as in the simple unbiasedness tests.[27] It is interesting that, even when we can decouple the test for weak efficiency with respect to the current forecast from the unbiasedness test, the test of unbiasedness using this ECM specification still requires weak efficiency with respect to the current forecast error.[28]

43.4.5 Explicit Tests of Weak Efficiency

The literature on rational expectations exhibits even less consensus as to the definition of efficiency than it does for unbiasedness. In general, an efficient forecast incorporates all available information – private as well as public. It follows that there should be no relationship between forecast error and any information variables known to the forecaster at the time of the forecast. Weak efficiency commonly denotes the orthogonality of the forecast error with respect to functions of the target and prediction. For example, there is no contemporaneous relationship between forecast and forecast error which could be exploited to reduce the error. Strong efficiency denotes orthogonality with respect to the remaining variables in the information set. Below we perform two types of weak efficiency tests. In the first type, we regress each group's forecast error on three sets of weak efficiency variables[29]:

[26]Since we estimate the restricted ECM with an intercept, unbiasedness also requires the intercept to be equal to zero.

[27]Since the intercept in Eq. 43.10 is not significant in any regression, the simple hypothesis that $\alpha_{i,h}$ equals one also fares the same as the simple unbiasedness tests.

[28]For purposes of comparison with both the bivariate joint and simple unbiasedness restrictions, we have used the ECM results using the robust standard errors. In all cases testing the ECM restrictions using F-statistics based on whitened residuals produces rejections of all restrictions, simple and joint, except a zero intercept. Hakkio and Rush (1989) found similarly strong rejections of Eq. 43.9, where the forecast was the forward rate.

[29]Notice that the first two sets of weak efficiency variables include the *mean* forecast, rather than the individual group forecast. Our intention is to allow a given group to incorporate information from other groups' forecasts via the *prior* mean forecast. This requires an extra lag in the information set variables, relative to a contemporaneously available variable such as the realized exchange rate depreciation.

1. Single and cumulative lags of the mean forecast error (lagged one period):

$$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \sum_{k=h+1}^{h+7} \beta_{i,t+h-k}\left(s_{t+h-k} - s^e_{m,t+h-k,h}\right) + \varepsilon_{i,t,h} \tag{43.11}$$

2. Single and cumulative lags of mean expected depreciation (lagged one period):

$$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \sum_{k=h+1}^{h+7} \beta_{i,t+h-k}\left(s^e_{m,t+h-k,h} - s_{t-k}\right) + \varepsilon_{i,t,h} \tag{43.12}$$

3. Single and cumulative lags of actual depreciation:

$$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \sum_{k=h}^{h+6} \beta_{i,t+h-k}\left(s_{t+h-k} - s_{t-k}\right) + \varepsilon_{i,t,h} \tag{43.13}$$

For each group and forecast horizon, we regress the forecast error on the most recent seven lags of the information set variable, both singly and cumulatively. We use a Wald test of the null hypothesis $\alpha_{i,h} = \beta_{i,t+h-k} = 0$ and report chi-square test statistics, with degrees of freedom equal to the number of regressors excluding the intercept. If we were to perform only simple regressions (i.e., on each lag individually), estimates of coefficients and tests of significance could be biased toward rejection due to the omission of relevant variables. If we were to perform only multivariate regressions, tests for joint significance could be biased toward nonrejection due to the inclusion of irrelevant variables. It is also possible that joint tests are significant but individual tests are not. This will be the case when the linear combination of (relatively uncorrelated) regressors spans the space of the dependent variable, but individual regressors do not.

In the only reported efficiency tests on JCIF data, Ito (1990) separately regressed the forecast error (average, group, and individual firm) on a single lagged forecast error, lagged forward premium, and lagged actual change. He found that, for the 51 biweekly forecasts between May 1985 and June 1987, rejections increased from a relative few at the 1- or 3-month horizons to virtual unanimity at the 6-month horizon. When he added a second lagged term for actual depreciation, rejections increased “dramatically” for all horizons.

The second type of weak efficiency tests uses the Breusch (1978)-Godfrey (1978) LM test for the null of no serial correlation of order k = h or greater, up to order k = h + 6, in the residuals of the forecast error regression, Eq. 43.11. Specifically, we estimate

$$\hat{\varepsilon}_{i,t,h} = \alpha_{i,h} + \sum_{k=1}^{h-1} \beta_{i,k}\left(s_{t+h-k} - s^e_{i,t-k,h}\right) + \sum_{l=h}^{h+6} \phi_{i,l}\hat{\varepsilon}_{i,t-l,h} + \eta_{i,t,h} \tag{43.14}$$

and test the null hypothesis $H_0 : \beta_{i,h} = \ldots = \beta_{i,h+6} = 0$ for $h = 2,6$.[30,31] Results for all efficiency tests for the 1- and 3-month horizons are presented in Tables 43.13, 43.14, 43.15, 43.16, 43.17, 43.18, 43.19, and 43.20. (Recall that the nonstationarity of the forecast errors at the 6-month horizon is an implicit rejection of weak efficiency.) For each group, horizon, and variable, there are seven individual tests, i.e., on a single lag, and six joint tests, i.e., on multiple lags. These 13 tests are multiplied by four groups times two horizons times three weak efficiency variables for a total of 312 efficiency tests.

Using approximately nine more years of data than Ito (1990), we find many rejections. In some cases, nearly all single lag tests are rejected, yet few, if any, joint tests are rejected. (See, e.g., expected depreciation at the 3-month horizon.) In other cases, nearly all joint tests are rejected, but few individual tests. (See, e.g., actual depreciation at the 3-month horizon.) Remarkably, all but one LM test for serial correlation at a specified lag produces a rejection at less than a 10 % level, with most at less than a 5 % level. Thus, it appears that the generality of the alternative hypothesis in the LM test permits it to reject at a much greater rate than the conventional weak efficiency tests, in which the variance-covariance matrix incorporates the Newey-West-Bartlett correction for heteroscedasticity and serial correlation. Finally, unlike Ito (1990), we find no strong pattern between horizon length and number of rejections.

43.5 Micro-homogeneity Tests

In addition to testing the rationality hypotheses at the individual level, we are interested in the degree of heterogeneity of coefficients across forecasters. Demonstrating that individual forecasters differ systematically in their forecasts (and forecast-generating processes) has implications for the market microstructure research program. As Frankel and Froot (1990, p. 182) noted, "the tremendous volume of foreign exchange trading is another piece of evidence that reinforces the idea of heterogeneous expectations, since it takes differences among market participants to explain why they trade."

Micro-homogeneity should have implications for rationality as well. Intuitively, if all forecasters pass rationality tests, then their corresponding regression coefficients should be equal. However, the converse is not necessarily true: if all forecasters have equal regression coefficients, they will not satisfy rationality conditions if they are all biased or inefficient to the same degree with respect to

[30]This is a general test, not only because it allows for an alternative hypothesis of higher-order serial correlation of specified order but also because it allows for serial correlation to be generated by AR, MA, or ARMA processes.

[31]We use the F-statistic because the χ^2 test statistics tend to over-reject, while the F-tests have more appropriate significance levels (see Kiviet 1987).

Table 43.13 Weak efficiency tests (1-month forecasts)

$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \sum_{p=h}^{h+6} \beta_{i,t+h-p}(s_{t+h-p} - s^e_{m,t+h-p,h}) + \varepsilon_{i,t,h}$ for $h = 2$ (43.11)

	i = 1		i = 2		i = 3		i = 4	
	Banks and brokers		Insurance and trading companies		Export industries		Life insurance and import companies	
Lags	χ^2	p-value	χ^2	p-value	χ^2	p-value	χ^2	p-value
Single								
2	0.132	0.895	0.139	0.709	1.871	0.171	0.334	0.563
3	−0.914	0.361	1.971	0.160	0.027	0.869	0.186	0.667
4	0.160	0.689	0.006	0.938	1.634	0.201	0.714	0.398
5	0.450	0.502	0.050	0.823	1.749	0.186	1.180	0.277
6	0.046	0.831	0.104	0.747	0.686	0.408	0.188	0.665
7	0.002	0.967	0.282	0.595	0.069	0.793	0.001	0.970
8	0.091	0.763	0.436	0.509	0.022	0.883	0.300	0.584
Cum.								
3	0.765	0.682	1.778	0.411	1.746	0.418	0.585	0.746
4	4.626	0.201	3.463	0.326	8.763	0.033	5.349	0.148
5	4.747	0.314	4.382	0.357	7.680	0.104	5.081	0.279
6	5.501	0.358	5.592	0.348	7.652	0.176	5.768	0.329
7	6.252	0.396	6.065	0.416	8.879	0.180	6.677	0.352
8	5.927	0.548	5.357	0.617	8.390	0.299	6.087	0.530

Selected micro-homogeneity tests

$H_0: \alpha_{i,h} = \alpha_{j,h}, \beta_{i,t+h-p} = \beta_{j,t+h-p}$ for all $i, j \neq i$

	$\chi^2(GMM)$	p-value	n
Single			
2	122.522	0.000	6
8	43.338	0.000	6
Cum.			
3	136.830	0.000	9
8	201.935	0.000	24

See Appendix 1 for structure of GMM VCV matrix incorporating Newey-West correction for serial correlation χ^2 statistics for mean forecast error regressions (p-value underneath)
Degrees of freedom (n) represent number of regressors, excluding intercept (n = 1 for single lag, n = max. lag −2 for cumulative lags)

the same variables. For the univariate unbiasedness regressions, the null of micro-homogeneity is given by H_0: $\alpha_{ih} = \alpha_{jh}$, for all $i, j \neq i$. Before testing for homogeneous intercepts in Eq. 43.7, we must specify the form for our GMM system variance-covariance matrix. Keane and Runkle (1990) first accounted for cross-sectional correlation (in price level forecasts) using a GMM estimator on pooled data. Bonham and Cohen (2001) tested the pooling specification by replacing Zellner's (1962) SUR variance-covariance matrix with a GMM counterpart that incorporates the Newey-West single-equation corrections (used in our individual equation tests above) plus allowances for corresponding cross-covariances, both contemporaneous and lagged. Bonham and Cohen (2001)

Table 43.14 Weak efficiency tests (1-month forecasts)

$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \sum_{p=h}^{h+6} \beta_{i,t+h-p}(s^e_{m,t+h-p,h} - s_{t-p}) + \varepsilon_{i,t,h}$ for $h = 2$ (43.12)

	i = 1		i = 2		i = 3		i = 4	
	Banks and brokers		Insurance and trading companies		Export industries		Life insurance and import companies	
Lags	χ^2	p-value	χ^2	p-value	χ^2	p-value	χ^2	p-value
Single								
2	−2.325	0.020	5.641	0.018	3.658	0.056	7.011	0.008
3	4.482	0.106	3.519	0.061	3.379	0.066	5.877	0.015
4	3.162	0.075	2.580	0.108	2.805	0.094	4.911	0.027
5	3.956	0.047	2.993	0.084	3.102	0.078	7.467	0.006
6	6.368	0.012	4.830	0.028	5.952	0.015	9.766	0.002
7	8.769	0.003	6.786	0.009	7.755	0.005	12.502	0.000
8	5.451	0.020	4.114	0.043	4.417	0.036	7.564	0.006
Cum.								
3	5.592	0.061	6.138	0.046	4.116	0.128	7.508	0.023
4	5.638	0.131	5.896	0.117	4.283	0.232	7.888	0.048
5	5.189	0.268	4.964	0.291	3.784	0.436	8.009	0.091
6	6.025	0.304	5.068	0.408	4.847	0.435	8.401	0.136
7	7.044	0.317	5.746	0.452	5.940	0.430	9.434	0.151
8	10.093	0.183	8.494	0.291	7.919	0.340	12.530	0.084

Selected micro-homogeneity tests

$H_0 : \alpha_{i,h} = \alpha_{j,h}, \beta_{i,t+h-p} = \beta_{j,t+h-p}$ for all $i, j \neq i$

	$\chi^2(GMM)$	p-value	n
Single			
2	40.462	0.000	6
8	30.739	0.000	6
Cum.			
3	42.047	0.000	6
8	46.124	0.004	24

See Appendix 1 for structure of GMM VCV matrix incorporating Newey-West correction for serial correlation χ^2 statistics for mean forecast error regressions (p-value underneath)
Degrees of freedom (n) represent number of regressors, excluding intercept (n = 1 for single lag, n = max. lag −2 for cumulative lags)

constructed a Wald statistic for testing the micro-homogeneity of individual forecaster regression coefficients in a system.[32]

Keane and Runkle (1990) provided some empirical support for their modeling of cross-sectional correlations, noting that the average covariance between a pair of

[32]Elliott and Ito (1999) used single-equation estimation that incorporated a White correction for heteroscedasticity and a Newey-West correction for serial correlation. (See the discussion below of Ito's tests of forecaster heterogeneity.)

Table 43.15 Weak efficiency tests (1-month forecasts)

$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \sum_{p=h}^{h+6} \beta_{i,t+h-p}(s_{t+h-p} - s_{t-p}) + \varepsilon_{i,t,h}$ for $h = 2$ (43.13)

	i = 1		i = 2		i = 3		i = 4	
	Banks and brokers		Insurance and trading companies		Export industries		Life insurance and import companies	
Lags	χ^2	p-value	χ^2	p-value	χ^2	p-value	χ^2	p-value
Single								
2	−0.328	0.743	0.639	0.424	1.249	0.264	0.023	0.879
3	1.621	0.203	3.060	0.080	0.000	0.993	0.550	0.458
4	0.002	0.964	0.335	0.562	0.819	0.366	0.146	0.702
5	0.086	0.770	0.042	0.837	1.001	0.317	0.344	0.557
6	0.165	0.685	0.916	0.339	0.029	0.864	0.095	0.758
7	0.850	0.357	1.861	0.172	0.329	0.566	1.152	0.283
8	0.597	0.440	1.088	0.297	0.317	0.574	1.280	0.258
Cum.								
3	1.978	0.372	3.169	0.205	1.940	0.379	1.132	0.568
4	3.304	0.347	3.501	0.321	5.567	0.135	3.318	0.345
5	3.781	0.436	4.248	0.373	5.806	0.214	3.598	0.463
6	3.651	0.601	4.646	0.461	5.756	0.331	3.819	0.576
7	4.493	0.610	5.609	0.468	6.608	0.359	5.040	0.539
8	5.619	0.585	6.907	0.439	7.907	0.341	6.521	0.480

Selected micro-homogeneity tests

$H_0 : \alpha_{i,h} = \alpha_{j,h},\ \beta_{i,t+h-p} = \beta_{j,t+h-p}$ for all $i, j \neq i$

	$\chi^2(GMM)$	p-value	n
Single			
2	150.698	0.000	6
8	45.652	0.000	6
Cum.			
3	161.950	0.000	9
8	214.970	0.000	24

See Appendix 1 for structure of GMM VCV matrix incorporating Newey-West correction for serial correlation
χ^2 statistics for mean forecast error regressions (p-value underneath)
Degrees of freedom (n) represent number of regressors, excluding intercept (n = 1 for single lag, n = max. lag −2 for cumulative lags)

forecasters is 58 % of the average forecast variance. In contrast, we use Pesaran's (2004) CD (cross-sectional dependence) test to check for lagged as well as contemporaneous correlations of forecast errors among pairs of forecasters:

$$CD = \sqrt{\frac{2T}{N(N-1)}} \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} \hat{\rho}_{ij}, \tag{43.15}$$

where T is the number of time periods, N = 4 is the number of individual forecasters, and $\hat{\rho}_{ij}$ is the sample correlation coefficient between forecasters i and j, $i \neq j$.

Table 43.16 LM test for serial correlation (1-month forecasts)

$H_0 : \beta_{i,h} = \ldots = \beta_{i,h+6} = 0$, for $h = 2$							
in $\hat{\varepsilon}_{i,t,h} = \alpha_{i,h} + \sum_{k=1}^{h-1} \beta_{i,k}\left(s_{t+h-k} - s^e_{i,t-k,h}\right) + \sum_{l=h}^{h+6} \phi_{i,l}\hat{\varepsilon}_{i,t-l,h} + \eta_{i,t,h}$,							
where ε is generated from							
$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \sum_{k=1}^{h-1} \beta_{i,k}(s_{t+h-k} - s^e_{i,t-k,h}) + \varepsilon_{i,t,h}$							
Cum. lags (k)	2	3	4	5	6	7	8
$n - k$	219	205	192	179	166	153	144
i = 1							
Banks and brokers							
$F(k, n-k)$	29.415	18.339	14.264	11.180	9.699	7.922	6.640
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000
i = 2							
Insurance and trading companies							
$F(k, n-k)$	30.952	19.506	15.372	11.661	9.695	8.120	7.050
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000
i = 3							
Export industries							
$F(k, n-k)$	32.387	20.691	16.053	12.951	10.628	9.418	7.520
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000
i = 4							
Life insurance and import companies							
$F(k, n-k)$	29.694	18.606	14.596	11.093	9.586	9.154	7.937
p-value	0.000	0.000	0.000	0.000	0.000	0.000	0.000

Under the null hypothesis of no cross-correlation, $CD \sim^{a} N(0, 1)$.[33] See Table 43.21 for CD test results. We tested for cross-correlation in forecast errors from lag zero up to lags four and eight for the 1 and 3-month forecast horizons, respectively. (The nonstationarity of the 6-month forecast error precludes using the CD test at that horizon.) At the 1-month horizon, cross-correlations from lags zero to four are each significant at the 5 % level. Since rational forecasts allow for (individual) serial correlation of forecast errors at lags of h-1 or less, and h = 2 for the 1-month horizon, the cross-correlations at lags two through four indicate violations of weak efficiency. Similarly, at the 3-month horizon, where h-1 = 5, there is significant cross-correlation at lag six.[34] However, it should be noted that, for many lags shorter than h, one cannot reject the null hypothesis that there are no cross-correlated forecast errors.

[33]Unlike Breusch and Pagan's (1980) LM test for cross-sectional dependence, Pesaran's (2004) CD test is robust to multiple breaks in slope coefficients and error variances, as long as the unconditional means of the variables are stationary and the residuals are symmetrically distributed.

[34]There are three instances of statistically significant negative test statistics for lags greater than h-1, none for lags less than or equal to h-1. Thus, some industries produce relatively high forecast errors several periods after others produce relative low forecast errors, and this information is not fully incorporated in some current forecasts.

Table 43.17 Weak efficiency tests (3-month forecasts)

$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \sum_{p=h}^{h+6} \beta_{i,t+h-p}(s_{t+h-p} - s^e_{m,t+h-p,h}) + \varepsilon_{i,t,h}$ for $h = 6$ (43.11)

	i = 1		i = 2		i = 3		i = 4	
	Banks and brokers		Insurance and trading companies		Export industries		Life insurance and import companies	
Lags	χ^2	p-value	χ^2	p-value	χ^2	p-value	χ^2	p-value
Single								
6	0.667	0.414	0.954	0.329	4.493	0.034	1.719	0.190
7	0.052	0.820	0.071	0.789	1.434	0.231	0.268	0.605
8	0.006	0.940	0.010	0.921	0.382	0.537	0.001	0.976
9	0.055	0.814	0.043	0.836	0.140	0.708	0.060	0.806
10	0.264	0.607	0.278	0.598	0.001	0.980	0.432	0.511
11	0.299	0.585	0.381	0.537	0.020	0.888	0.598	0.439
12	0.172	0.678	0.336	0.562	0.011	0.918	0.633	0.426
Cum.								
7	8.966	0.011	11.915	0.003	19.663	0.000	12.350	0.002
8	12.288	0.006	16.263	0.001	23.290	0.000	15.146	0.002
9	11.496	0.022	15.528	0.004	22.417	0.000	14.778	0.005
10	8.382	0.136	12.136	0.033	16.839	0.005	12.014	0.035
11	11.596	0.072	18.128	0.006	23.782	0.001	15.330	0.032
12	11.527	0.117	15.983	0.025	21.626	0.003	13.038	0.071

Selected micro-homogeneity tests

$H_0 : \alpha_{i,h} = \alpha_{j,h}, \beta_{i,t+h-p} = \beta_{j,t+h-p}$ for all $i, j \neq i$

	$\chi^2(GMM)$	p-value	n
Single			
6	188.738	0.000	6
12	63.364	0.000	6
Cum.			
7	217.574	0.000	9
12	229.567	0.000	24

See Appendix 1 for structure of GMM VCV matrix incorporating Newey-West correction for serial correlation
χ^2 statistics for mean forecast error regressions (p-value underneath)
Degrees of freedom (n) represent number of regressors, excluding intercept
(n = 1 for single lag, n = max. lag −2 for cumulative lags)

Nevertheless, in our micro-homogeneity tests, we follow Bonham and Cohen (2001), allowing for an MA(h-1) residual process, both individually and among pairs of forecast errors. (See the Appendix 1 for details.) By more accurately describing the panel's residual variance-covariance structure, we expect this systems approach to improve the consistency of our estimates. Consider first the four bivariate regressions in Tables 43.1, 43.2, and 43.3. Recall that we rejected the joint hypothesis ($\alpha_{i,h}$, $\beta_{i,h}$) = (0, 1) at the 5 % significance level for all groups at the 1-month horizon (indicating the possible role of inefficiency with respect to the

Table 43.18 Weak efficiency tests (3-month forecasts)

$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \sum_{p=h}^{h+6} \beta_{i,t+h-p}(s^e_{m,t+h-p,h} - s_{t-p}) + \varepsilon_{i,t,h}$ for $h = 6$								(43.12)
	i = 1		i = 2		i = 3		i = 4	
	Banks and brokers		Insurance and trading companies		Export industries		Life insurance and import companies	
Lags	χ^2	p-value	χ^2	p-value	χ^2	p-value	χ^2	p-value
Single								
6	3.457	0.063	2.947	0.086	3.470	0.062	3.681	0.055
7	4.241	0.039	3.834	0.050	4.390	0.036	4.370	0.037
8	5.748	0.017	5.177	0.023	5.410	0.020	6.053	0.014
9	6.073	0.014	5.843	0.016	5.968	0.015	6.474	0.011
10	8.128	0.004	7.868	0.005	7.845	0.005	8.521	0.004
11	8.511	0.004	8.004	0.005	8.308	0.004	8.429	0.004
12	6.275	0.012	6.691	0.010	6.635	0.010	6.079	0.014
Cum.								
7	4.717	0.095	4.985	0.083	4.954	0.084	4.928	0.085
8	5.733	0.125	5.209	0.157	5.045	0.168	6.736	0.081
9	5.195	0.268	5.411	0.248	5.112	0.276	6.053	0.195
10	7.333	0.197	9.245	0.100	9.456	0.092	7.872	0.163
11	8.539	0.201	6.658	0.354	7.488	0.278	7.955	0.241
12	8.758	0.271	6.747	0.456	7.796	0.351	8.698	0.275
Selected micro-homogeneity tests								
$H_0: \alpha_{i,h} = \alpha_{j,h}, \beta_{i,t+h-p} = \beta_{j,t+h-p}$ for all $i, j \neq i$								
	$\chi^2(GMM)$	p-value	n					
Single								
6	57.130	0.000	6					
12	58.230	0.000	6					
Cum.								
7	63.917	0.000	9					
12	126.560	0.000	24					

See Appendix 1 for structure of GMM VCV matrix incorporating Newey-West correction for serial correlation
χ^2 statistics for mean forecast error regressions (p-value underneath)
Degrees of freedom (n) represent number of regressors, excluding intercept
(n = 1 for single lag, n = max. lag −2 for cumulative lags)

current forecast), but only for the exporters at the 3- and 6-month horizons. However, there are no rejections of micro-homogeneity for any horizon.[35]

The micro-homogeneity test results are very different for both the 1- and 3-month systems of univariate unbiasedness regressions in Tables 43.7 and 43.8. (Recall that

[35]The nonrejection of micro-homogeneity in bivariate regressions does not, however, mean that one can avoid aggregation bias by using the mean forecast. Even if the bivariate regressions were correctly interpreted as joint tests of unbiasedness and weak efficiency with respect to the current forecast, and even if the regressions had sufficient power to reject a false null, the micro-homogeneity tests would be subject to additional econometric problems. According to the Figlewski-Wachtel (1983) critique, successfully passing a pretest for micro-homogeneity does not ensure that estimated coefficients from such consensus regressions will be consistent. See Sect. 43.2.1.

Table 43.19 Weak efficiency tests (3-month forecasts)

$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \sum_{p=h}^{h+6} \beta_{i,t+h-p}(s_{t+h-p} - s_{t-p}) + \varepsilon_{i,t,h}$ for $h = 6$								(43.13)
	i = 1		i = 2		i = 3		i = 4	
	Banks and brokers		Insurance and trading companies		Export industries		Life insurance and import companies	
Lags	χ^2	*p*-value	χ^2	*p*-value	χ^2	*p*-value	χ^2	*p*-value
Single								
6	0.268	0.604	0.450	0.502	3.657	0.056	1.065	0.302
7	0.055	0.814	0.037	0.848	0.599	0.439	0.003	0.957
8	0.331	0.565	0.305	0.581	0.029	0.864	0.230	0.632
9	0.513	0.474	0.482	0.488	0.022	0.883	0.577	0.448
10	1.038	0.308	1.077	0.299	0.318	0.573	1.344	0.246
11	1.335	0.248	1.532	0.216	0.563	0.453	1.872	0.171
12	1.184	0.276	1.620	0.203	0.616	0.433	1.979	0.159
Cum.								
7	6.766	0.034	8.767	0.012	15.683	0.000	10.052	0.007
8	8.752	0.033	11.784	0.008	18.330	0.000	11.162	0.011
9	8.654	0.070	11.588	0.021	18.929	0.001	11.309	0.023
10	9.421	0.093	12.890	0.024	19.146	0.002	12.275	0.031
11	9.972	0.126	13.137	0.041	19.597	0.003	13.003	0.043
12	8.581	0.284	11.823	0.107	17.670	0.014	11.431	0.121
Selected micro-homogeneity tests								
$H_0 : \alpha_{i,h} = \alpha_{j,h}, \beta_{i,t+h-p} = \beta_{j,t+h-p}$ for all $i, j \neq i$								
	$\chi^2(GMM)$	*p*-value	*n*					
Single								
6	151.889	0.000	6					
12	66.313	0.000	6					
Cum.								
7	164.216	0.000	9					
12	193.021	0.000	24					

See Appendix 1 for structure of GMM VCV matrix incorporating Newey-West correction for serial correlation
χ^2 statistics for mean forecast error regressions (*p*-value underneath)
Degrees of freedom (n) represent number of regressors, excluding intercept
(n = 1 for single lag, n = max. lag −2 for cumulative lags)

unbiasedness was rejected for all groups at the 6-month horizon due to the nonstationarity of the forecast error.) Despite having only one failure of unbiasedness at the 5 % level for the two shorter horizons, micro-homogeneity is rejected at a level of virtually zero for both horizons. The rejection of micro-homogeneity at the 1-month horizon occurs despite the failure to reject unbiasedness for any of the industry groups. We hypothesize that the consistent rejection of micro-homogeneity regardless of the results of individual unbiasedness tests is the result of sufficient variation in individual bias estimates as well as precision in these estimates. According to these tests, aggregation of individual forecasts into a mean forecast is invalid at all horizons.

Table 43.20 LM test for serial correlation (3-month forecasts)

$H_0 : \beta_{i,h} = \ldots = \beta_{i,h+6} = 0$, for $h = 6$							
in $\hat{\varepsilon}_{i,t,h} = \alpha_{i,h} + \sum_{k=1}^{h-1} \beta_{i,k}\left(s_{t+h-k} - s^e_{i,t-k,h}\right) + \sum_{l=h}^{h+6} \phi_{i,l}\hat{\varepsilon}_{i,t-l,h} + \eta_{i,t,h}$,							
where ε is generated from							
$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \sum_{k=1}^{h-1} \beta_{i,k}\left(s_{t+h-k} - s^e_{i,t-k,h}\right) + \varepsilon_{i,t,h}$							
Cum. lags (k)	6	7	8	9	10	11	12
$n - k$	126	117	108	99	94	89	84
i = 1							
Banks and brokers							
$F(k, n - k)$	3.452	2.856	3.023	2.951	2.599	2.652	2.921
p-value	0.003	0.009	0.004	0.004	0.008	0.006	0.002
i = 2							
Insurance and trading companies							
$F(k, n - k)$	3.499	2.850	3.408	2.907	2.492	2.584	2.341
p-value	0.003	0.009	0.002	0.004	0.011	0.007	0.012
i = 3							
Export industries							
$F(k, n - k)$	4.687	3.956	4.409	3.572	2.928	2.819	2.605
p-value	0.000	0.001	0.000	0.001	0.003	0.003	0.005
i = 4							
Life insurance and import companies							
$F(k, n - k)$	2.352	2.482	2.501	2.168	1.866	1.794	1.811
p-value	0.035	0.021	0.016	0.031	0.060	0.067	0.059

In addition to testing the weak efficiency hypothesis at the individual level, we are interested in the degree of heterogeneity of coefficients across forecasters. Here the null of micro-homogeneity is given by H_0: $\phi_{il} = \phi_{jl}$, for $l = h, \ldots h + 6$, for all i, $j \neq i$. As explained in the section on efficiency tests, there are 312 tests (not 468, due to a nonstationary forecast error for all four groups at the 6-month horizon)/four groups = 83 micro-homogeneity tests. The null hypothesis of equal coefficients is $H_0 : \alpha_{i,h} = \alpha_{j,h}, \beta_{i,t+h-k} = \beta_{j,t+h-k}$ for all $i, j \neq i$. As with the micro-homogeneity tests for unbiasedness, our GMM variance-covariance matrix accounts for serial correlation of order h-1 or less, generalized heteroscedasticity, and cross-sectional correlation or order h-1 or less. We report $\chi^2(n)$ statistics, where n is the number of coefficient restrictions, with corresponding p-values. Rather than perform all 83 micro-homogeneity tests, we choose a sample consisting of the shortest and longest lag for which there are corresponding individual and joint tests (i.e., for the k = h + 1st and k = h + 6th lag). Thus, there are four tests (two individual and two corresponding joint tests) times two horizons times three variables for a total of 24 tests. Every one of the micro-homogeneity tests are rejected at the 0 % level. As pointed out by Bryant (1995), a finding of micro-heterogeneity in unbiasedness and weak efficiency tests also casts doubt on the assumption of a rational representative agent commonly used in macroeconomic and asset-pricing models (Table 43.22).

Table 43.21 CD tests for cross-sectional

$s_{t+h} - s^e_{i,t,h} = \alpha_{i,h} + \varepsilon_{i,t,h}$ (43.7)

$CD = \sqrt{\frac{2T}{N(N-1)}}\sum_{i=1}^{N-1}\sum_{j=i+1}^{N}\hat{\rho}_{ij} \overset{a}{\sim} N(0,1)$ (43.15)

Lag length	CD	p-value
3-month horizon		
0	31.272	0.000
1	2.461	0.014
2	0.387	0.699
3	2.322	0.020
4	1.594	0.111
$h-1=5$	1.461	0.144
6	−5.887	0.000
7	0.340	0.734
8	1.456	0.145

$N = 24$, $T = 276$, $\hat{\rho}_{ij}$ is the sample correlation coefficient between forecasters i and j, $i \neq j$

43.5.1 Ito's Heterogeneity Tests

In Table 43.23, we replicate Ito's (1990) and Elliott and Ito's (1999) test for forecaster "heterogeneity." This specification regresses the deviation of the individual forecast from the cross-sectional average forecast on a constant. Algebraically, Ito's regression can be derived from the individual forecast error regression by subtracting the mean forecast error regression. Thus, because it simply replaces the forecast error with the individual deviation from the mean forecast, it does not suffer from aggregation bias (c.f. Figlewski and Wachtel (1983)) or pooling bias (c.f. Zarnowitz 1985) (Table 43.24).[36, 37]

$$s^e_{i,t,h} - s^e_{m,t,h} = \left(\alpha_{i,h} - \alpha_m\right) + \left(\varepsilon_{i,t,h} - \varepsilon_{m,t}\right) \quad (43.16)$$

As above, we use the Newey-West-Bartlett variance-covariance matrix.

One may view Ito's "heterogeneity" tests as complementary to our micro-homogeneity tests. On the one hand, one is not certain whether a single (or pair of?) individual rejection(s) of, say, the null hypothesis of a zero mean deviation in Ito's test would result in a rejection of micro-homogeneity overall. On the other hand, a rejection of micro-homogeneity does not tell us which groups are the most significant violators of the null hypothesis. It turns out that Ito's mean deviation test produces rejections at a level of 6 % or less for all groups at all horizons except for

[36]Recall that our group results are not entirely comparable to Ito's (1990), since our data set, unlike his, combines insurance companies and trading companies into one group and life insurance companies and import-oriented companies into another group.

[37]Chionis and MacDonald (1997) performed an Ito-type test on individual expectations data from Consensus Forecasts of London.

Table 43.22 Ito tests (1-month forecasts)

Individual regressions				
$s^e_{i,t,h} - s^e_{m,t,h} = (\alpha_{i,h} - \alpha_m) + (\varepsilon_{i,t,h} - \varepsilon_{m,t})$ for $h = 2$				
Degrees of freedom = 263				
	i = 1	i = 2	i = 3	i = 4
	Banks and brokers	Insurance and trading companies	Export industries	Life insurance and import companies
$\alpha_{i,h}$	0	−0.001	0.003	−0.002
t (NW)	0.173	−2.316	5.471	−3.965
p-value	0.863	0.021	0	0
MH tests $H_0 : \alpha_{i,h} = \alpha_{j,h}$, for all $i, j \neq I$				
χ^2(GMM)	40.946			
p-value	0			

Table 43.23 Ito tests (3-month forecasts)

Individual regressions				
$s^e_{i,t,h} - s^e_{m,t,h} = (\alpha_{i,h} - \alpha_m) + (\varepsilon_{i,t,h} - \varepsilon_{m,t})$ for $h = 6$				
Degrees of freedom = 263				
	i = 1	i = 2	i = 3	i = 4
	Banks and brokers	Insurance and trading companies	Export industries	Life insurance and import companies
$\alpha_{i,h}$	−0.002	−0.003	0.008	−0.002
t (NW)	−2.307	−3.986	5.903	−1.883
p-value	0.021	0	0	0.06
MH tests $H_0 : \alpha_{i,h} = \alpha_j$, for all $i, j \neq I$				
χ^2(GMM)	37.704			
p-value	0			

Table 43.24 Ito tests (6-month forecasts)

Individual regressions				
$s^e_{i,t,h} - s^e_{m,t,h} = (\alpha_{i,h} - \alpha_m) + (\varepsilon_{i,t,h} - \varepsilon_{m,t})$ for $h = 12$				
Degrees of freedom = 263				
	i = 1	i = 2	i = 3	i = 4
	Banks and brokers	Insurance and trading companies	Export industries	Life insurance and import companies
$\alpha_{i,h}$	−0.004	−0.003	0.01	0
t (NW)	−3.52	−2.34	4.549	−0.392
p-value	0	0.019	0	0.695
MH tests $H_0 : \alpha_{i,h} = \alpha_{j,h}$, for all $i, j \neq I$				
χ^2(GMM)	23.402			
p-value	0.001			

banks and brokers at the 1-month horizon and life insurance and import companies at the 6-month horizon.[38] Since Ito's regressions have a similar form (though not a similar economic interpretation) to the tests for univariate unbiasedness in Tables 43.7 and 43.8, it is not surprising that micro-homogeneity tests on the four-equation system of Ito equations produce rejections at a level of virtually zero for all three horizons.

43.6 Conclusions

In this chapter, we undertake a reexamination of the rationality and diversity of JCIF forecasts of the yen-dollar exchange rate. In several ways we update and extend the seminal paper by Ito (1990). In particular, we have attempted to explore the nature of rationality tests on integrated variables. We show that tests based on the "conventional" bivariate regression in change form, while correctly specified in terms of integration accounting, have two major shortcomings. First, following Holden and Peel (1990), they are misspecified as unbiasedness tests, because rejection of the (0, 1) restriction on the slope and intercept is a sufficient, not a necessary, condition for unbiasedness. Only a zero restriction on the intercept in a regression of the forecast error on a constant is both necessary and sufficient for unbiasedness. Second, tests using the bivariate specification suffer from a lack of power. Yet, this is exactly what we would expect in an asset market whose price is a near random walk: the forecasted change is nearly unrelated to (and varies much less than) the actual change.

In contrast, we conduct pretests for rationality based on determining whether the realization and forecast are each integrated and cointegrated. In this case, following Liu and Maddala (1992), a "restricted" cointegration test, which imposes a (0, 1) restriction on the cointegrating vector, is necessary for testing unbiasedness. (We show that the Holden and Peel (1990) critique does not apply if the regressor and regressand are cointegrated.) If a unit root in the restricted residual is rejected, then the univariate test which regresses the forecast error on a constant is equivalent to the restricted cointegration test. Testing this regression for white noise residuals is one type of weak efficiency test. Testing other stationary regressors in the information set for zero coefficients produces additional efficiency tests.

In the univariate specification, we find that, for each group, the ability to produce unbiased forecasts deteriorates with horizon length: no group rejects unbiasedness at the 1-month horizon, but all groups reject at the 6-month horizon, because the forecast errors are nonstationary. Exporters consistently perform worse than the other industry groups, with a tendency toward depreciation bias.

[38]Elliott and Ito (1999), who have access to forecasts for the 42 individual firms in the survey, find that, for virtually the same sample period as ours, the null hypothesis of a zero deviation from the mean forecast is rejected at the 5 % level by 17 firms at the 1-month horizon, 13 firms for the 3-month horizon, and 12 firms for the 6-month horizon. These authors do not report results by industry group.

Using only 2 years of data, Ito (1990) found the same result for exporters, which he described as a type of "wishful thinking."

The unbiasedness results are almost entirely reversed when we test the hypothesis using the conventional bivariate specification. That is, the joint hypothesis of zero intercept and unit slope is rejected for all groups at the 1-month horizon, but only for exporters and the 3- and 6-month horizons. Thus, in stark contrast to the univariate unbiasedness tests, as well as Ito's (1990) bivariate tests, forecast performance does not deteriorate with increases in the horizon.

Also, since Engle and Granger (1987) have showed that cointegrated variables have an error correction representation, we impose joint "unbiasedness" restrictions first used by Hakkio and Rush (1989) on the ECM. However, we show that these restrictions also represent sufficient, not necessary, conditions, so these tests could tend to over-reject. We then develop and test restrictions which are both necessary and sufficient conditions for unbiasedness. The test results confirm that the greater rate of rejections of the joint "unbiasedness" restrictions in the ECM is caused by the failure of the implicit restriction of weak efficiency with respect to the lagged forecast. When we impose the restriction that the coefficient of the forecast equals one, the ECM unbiasedness test results mimic those of the simple univariate unbiasedness tests. For this data set, at least, it does not appear that an ECM provides any value added over the simple unbiasedness test. Furthermore, since the error correction term is not statistically significant in any regressions, it is unclear whether the ECM provides any additional insight into the long-run adjustment mechanism of exchange rate changes.

The failure of more general forms of weak efficiency is borne out by two types of explicit tests for weak efficiency. In the first type, we regress the forecast error on single and cumulative lags of mean forecast error, mean forecasted depreciation, and actual depreciation. We find many rejections of weak efficiency. In the second type, we use the Godfrey (1978) LM test for serial correlation of order h through h + 6 in the residuals of the forecast error regression. Remarkably, all but one LM test at a specified lag length produces a rejection at less than a 10 % level, with most at less than a 5 % level. (As in the case of the univariate unbiasedness test, all weak efficiency tests at the 6-month horizon fail due to the nonstationarity of the forecast error.)

Whereas Ito (1990) and Elliott and Ito (1999) measured diversity as a statistically significant deviation of an individual's forecast from the cross-sectional average forecast, we perform a separate test of micro-homogeneity for each type of rationality test – unbiasedness as well as weak efficiency – that we first conducted at the industry level. In order to conduct the systems estimation and testing required for the micro-homogeneity test, our GMM estimation and inference make use of an innovative variance-covariance matrix that extends the Keane and Runkle (1990) counterpart from a pooled to an SUR-type structure. Our variance-covariance matrix takes into account not only serial correlation and heteroscedasticity at the individual level (via a Newey-West-Bartlett correction) but also forecaster cross-correlation up to h-1 lags. We document the statistical significance of the cross-sectional correlation using Pesaran's (2004) CD test.

In the univariate unbiasedness tests, we find that, irrespective of the ability to produce unbiased forecasts at a given horizon, micro-homogeneity is rejected at virtually a 0 % level for all horizons. We find this result to be somewhat counter-intuitive, in light of our prior belief that micro-homogeneity would be more likely to obtain if there were no rejections of unbiasedness. Evidently, there is sufficient variation in the estimated bias coefficient across groups and/or high precision of these estimates to make the micro-homogeneity test quite sensitive. Micro-homogeneity is also strongly rejected in the weak efficiency tests.

In contrast to the results with the univariate unbiasedness specification, micro-homogeneity is not rejected at any horizon in the bivariate regressions. We conjecture that the imprecise estimation of the slope coefficient makes it difficult to reject joint hypotheses involving this coefficient.

In conclusion, we recommend that all rationality tests be undertaken using simple univariate specifications at the outset (rather than only if the joint bivariate test is rejected, as suggested by Mincer and Zarnowitz (1969) and Holden and Peel (1990) and employed by Gavin (2003)). Before conducting such tests, one should test the restricted cointegrated regression residuals, i.e., the forecast error, for stationarity. Clearly, integration accounting and regression specification matter for rationality testing.

While our rationality tests do not attempt to explain cross-sectional dispersion, the widespread rejection of micro-homogeneity in different specifications of unbiasedness and weak efficiency tests[39] provides more motivation for the classification of forecasters into types (e.g., fundamentalist and chartist/noise traders) than for simply assuming a representative agent (with rational expectations).

There are characteristics of forecasts other than rationality which are of intrinsic interest. Given our various rejections of rational expectations, it is natural to explore what expectational mechanism the forecasters use. Ito (1994) tested the mean JCIF forecasts for extrapolative and regressive expectations, as well as a mixture of the two.[40] Cohen and Bonham (2006) extend this analysis using individual forecast-generating processes and additional learning model specifications. And, much of the literature on survey forecasts has analyzed the accuracy of predictions, typically ranking forecasters by MSE. One relatively unexplored issue is the statistical significance of the ranking, regardless of loss function. However, other loss functions, especially nonsymmetric ones, are also reasonable. For example, Elliott and Ito (1999) have ranked individual JCIF forecasters using a profitability criterion. As mentioned in Sect. 43.2.2, the loss function may incorporate strategic considerations that result in “rational bias.” Such an exploration would require more disaggregated data than the JCIF industry forecasts to which we have access.

[39]We put less weight on the results of the weaker tests for micro-homogeneity in the bivariate regressions.

[40]He also included regressors for adaptive expectations and the forward premium.

Appendix 1: Testing Micro-homogeneity with Survey Forecasts

The null hypothesis of micro-homogeneity is that the slope and intercept coefficients in the equation of interest are equal across individuals. This chapter considers the case of individual unbiasedness regressions such as Eq. 43.2 in the text, repeated here for convenience,

$$s_{t+h} - s_t = \alpha_{i,h} + \beta_{i,h}\left(s^e_{i,t,h} - s_t\right) + \varepsilon_{i,t,h} \tag{43.17}$$

and tests $H_0 : \alpha_1 = \alpha_2 = \ldots = \alpha_N$ and $\beta_1 = \beta_2 = \ldots = \beta_N$.

Stack all N individual regressions into the Seemingly Unrelated Regression system

$$S = \mathbf{F}\theta + \varepsilon \tag{43.18}$$

where S is the $NT \times 1$ stacked vector of realizations, s_{t+h}, and $\mathbf{F}$ is an $NT \times 2N$ block diagonal data matrix:

$$\mathbf{F} = \begin{bmatrix} \mathbf{F}_1 & & \\ & \ddots & \\ & & \mathbf{F}_N \end{bmatrix}. \tag{43.19}$$

Each $\mathbf{F}_i = [\iota\ s^e_{i,t,h}]$ is a $T \times 2$ matrix of ones and individual i's forecasts, $\theta = [\alpha_1\ \beta_1 \ldots \alpha_N\ \beta_N]'$, and ε is an $NT \times 1$ vector of stacked residuals. The vector of restrictions, $R\theta = r$, corresponding to the null hypothesis of micro-homogeneity is normally distributed, with $R\theta - r \sim N[0, R(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'\Omega\mathbf{F}(\mathbf{F}'\mathbf{F})^{-1}R']$, where R is the $2(N-1) \times 2N$ matrix

$$R = \begin{bmatrix} 1 & 0 & -1 & 0 & \ldots & 0 \\ 0 & 1 & 0 & -1 & 0 & \vdots \\ \vdots & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & \ldots & 0 & 1 & 0 & -1 \end{bmatrix}, \tag{43.20}$$

and r is a $2(N-1) \times 1$ vector of zeros. The corresponding Wald test statistic, $\left(R\hat{\theta} - r\right)' \left[R(\mathbf{F}'\mathbf{F})^{-1}\mathbf{F}'\hat{\Omega}\mathbf{F}\left(\mathbf{F}'\mathbf{F}\right)^{-1}R'\right]\left(R\hat{\theta} - r\right)$, is asymptotically distributed as a chi-square random variable with degrees of freedom equal to the number of restrictions, $2(N-1)$.

For most surveys, there are a large number of missing observations. Keane and Runkle (1990), Davies and Lahiri (1995), Bonham and Cohen (1995, 2001), and to the best of our knowledge all other papers which make use of pooled regressions in tests of the REH have dealt with the missing observations using the same approach. The pooled or individual regression is estimated by eliminating the missing data points in both the forecasts and the realization. The regression residuals are then padded with zeros in place of missing observations to allow for the calculation of own and cross-covariances. As a result, many individual variances and cross-covariances are calculated with relatively few pairs of residuals. These individual cross-covariances are then averaged. In Keane and Runkle (1990) and Bonham and Cohen (1995, 2001) the assumption of $2(k + 1)$ second moments, which are common to all forecasters, is made for analytical tractability and for increased reliability. In contrast to the forecasts from the Survey of Professional Forecasters used in Keane and Runkle (1990) and Bonham and Cohen (1995, 2001), the JCIF data set contains virtually no missing observations. As a result, it is possible to estimate each individual's variance-covariance matrix (and cross-covariance matrix) rather than average over all individual variances and cross-covariance pairs as in the aforementioned papers.

We assume that for each forecast group i,

$$\begin{aligned}
&E\left[\varepsilon_{i,t,h}\varepsilon_{i,t,h}\right] = \sigma_{i,0}^2 \text{ for all } i, t,\\
&E\left[\varepsilon_{i,t,h}\varepsilon_{i,t+k}\right] = \sigma_{i,k}^2 \text{ for all } i, t, k \text{ such that } 0 < k \le h, \\
&E\left[\varepsilon_{i,t,h}\varepsilon_{i,t+k}\right] = 0 \text{ for all } i, t, k \text{ such that } k > h,
\end{aligned} \tag{43.21}$$

Similarly, for each pair of forecasters i and j, we assume

$$\begin{aligned}
E\left[\varepsilon_{i,t,h}\varepsilon_{j,t}\right] &= \delta_{i,j}(0) \quad \forall i,j,t,\\
E\left[\varepsilon_{i,t,h}\varepsilon_{j,t+k}\right] &= \delta_{i,j}(k) \ \forall i,j,t,k \text{ such that } k \ne 0, \text{and} -h \le k \le h.\\
E\left[\varepsilon_{i,t,h}\varepsilon_{j,t+k}\right] &= 0 \qquad \forall i,j,t,k \text{ such that } k > |h|.
\end{aligned} \tag{43.22}$$

Thus, each pair of forecasters has a different $T \times T$ cross-covariance matrix:

$$\mathbf{P}_{i,j} = \begin{bmatrix}
\delta_{i,j}(0) & \delta_{i,j}(-1) & \dots & \delta_{i,j}(-h) & 0\\
\delta_{i,j}(1) & \delta_{i,j}(0) & \delta_{i,j}(-1) & \dots & 0\\
\vdots & \ddots & \ddots & \ddots & \vdots\\
 & \dots & \delta_{i,j}(1) & \delta_{i,j}(0) & \delta_{i,j}(-1)\\
0 & \delta_{i,j}(h) & \dots & \delta_{i,j}(1) & \delta_{i,j}(0)
\end{bmatrix}, \tag{43.23}$$

Finally, note that $P_{i,j} \ne P_{j,i}$, rather $P'_{i,j} = P_{j,i}$. The complete variance-covariance matrix, denoted Ω, has dimension $NT \times NT$, with matrices $\mathbf{Q}_i$ on the main diagonal and $\mathbf{P}_{i,j}$ off the diagonal.

The individual Q_i, variance-covariances matrices are calculated using the Newey and West (1987) heteroscedasticity-consistent, MA(j)-corrected form. The $P_{i,j}$ matrices are estimated in an analogous manner.

References

Banerjee, A., Dolado, J. J., Galbraith, J. W., & Hendry, D. F. (1993). *Co-integration, error correction and the econometric analysis of non-stationary data*. Oxford: Oxford University Press.

Batchelor, R., & Dua, P. (1990a). Forecaster ideology, forecasting technique, and the accuracy of economic forecasts. *International Journal of Forecasting, 6*, 3–10.

Batchelor, R., & Dua, P. (1990b). Product differentiation in the economic forecasting industry. *International Journal of Forecasting, 6*, 311–316.
Batchelor, R., & Dua, P. (1992). Conservatism and consensus-seeking among economic forecasters. *Journal of Forecasting, 11*, 169–181.
Batchelor, R., & Peel, D. A. (1998). Rationality testing under asymmetric loss. *Economics Letters, 61*, 49–54.
Bonham, C. S., & Cohen, R. H. (1995). Testing the rationality of price forecasts: Comment. *American Economic Review, 85*(1), 284–289.
Bonham, C. S., & Cohen, R. H. (2001). To aggregate, pool, or neither: Testing the rational-expectations hypothesis using survey data. *Journal of Business & Economic Statistics, 19*, 278–291.
Boothe, P., & Glassman, D. (1987). Comparing exchange rate forecasting models. *International Journal of Forecasting, 3*, 65–79.
Breusch, T. S. (1978). Testing for autocorrelation in dynamic linear models. *Australian Economic Papers, 17*, 334–355.
Breusch, T. S., & Pagan, A. R. (1980). The Lagrange multiplier test and its application to model specifications in econometrics. *Review of Economic Studies, 47*, 239–253.
Bryant, R. C. (1995). *The "exchange risk premium," uncovered interest parity, and the treatment of exchange rates in multicountry macroeconomic models*. Technical report, The Brookings Institution.
Cavaglia, S. M. F. G., Verschoor, W. F. C., & Wolff, C. C. P. (1994). On the biasedness of forward foreign exchange rates: Irrationality or risk premia? *Journal of Business, 67*, 321–343.
Chionis, D., & MacDonald, R. (1997). Some tests of market microstructure hypotheses in the foreign exchange market. *Journal of Multinational Financial Management, 7*, 203–229.
Clements, M. P., & Hendry, D. F. (1998). Forecasting economic time series. Cambridge: Cambridge University Press.
Cohen, R. H., & Bonham, C. S. (2006). *Specifying the forecast generating process for exchange rate survey forecasts*. Working paper, Department of Economics, University of Hawaii Manoa.
Davies, A., & Lahiri, K. (1995). A new framework for testing rationality and measuring aggregate shocks using panel data. *Journal of Econometrics, 68*, 205–227.
Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association, 74*, 427–431.
Ehrbeck, T., & Waldmann, R. (1996). Why are professional forecasters biased? Agency versus behavioral explanations. *The Quarterly Journal of Economics, 111*, 21–40.
Elliott, G., & Ito, T. (1999). Heterogeneous expectations and tests of efficiency in the yen/dollar forward exchange rate market. *Journal of Monetary Economics, 43*, 435–456.
Engle, R., & Granger, C. W. S. (1987). Cointegration and error correction representation, estimation and testing. *Econometrica, 55*, 251–276.
Figlewski, S. (1978). Market "efficiency" in a market with heterogeneous information. *Journal of Political Economy, 86*, 581–597.
Figlewski, S. (1982). Information diversity and market behavior. *Journal of Finance, 37*, 87–102.
Figlewski, S. (1984). Information diversity and market behavior: A reply. *Journal of Finance, 39*, 299–302.
Figlewski, S., & Wachtel, P. (1983). Rational expectations, informational efficiency, and tests using survey data: A reply. *Review of Economics and Statistics, 65*, 529–531.
Frankel, J. A., & Froot, K. A. (1987). Using survey data to test standard propositions regarding exchange rate expectations. *American Economic Review, 77*, 133–153.
Frankel, J. A., & Froot, K. A. (1990). Chartists, fundamentalists, and trading in the foreign exchange market. *American Economic Review, 80*, 181–185.
Froot, K. A., & Frankel, J. A. (1989). Forward discount bias: Is it an exchange risk premium? *Quarterly Journal of Economics, 104*, 139–161.
Gavin, W. T. (2003) FOMC forecasts: Is all the information in the central tendency? *Economic Review, Federal Reserve Bank of St. Louis*, May/June, 27–46.

Godfrey, L. G. (1978). Testing for higher order serial correlation in regression equations when the regressors include lagged dependent variables. *Econometrica, 46*, 1303–1310.

Goldberg, M. D., & Frydman, R. (1996). Imperfect knowledge and behaviour in the foreign exchange market. *Economic Journal, 106*, 869–893.

Granger, C. W. J. (1991). Developments in the study of cointegrated economic variables. In R. F. Engle & C. W. J. Granger (Eds.), *Long run economic relationships: Readings in Cointegration* (pp. 65–80). Oxford: Oxford University Press.

Hakkio, C. S., & Rush, M. (1989). Market efficiency and cointegration: An application to the sterling deutschemark exchange markets. *Journal of International Money and Finance, 8*, 75–88.

Haltiwanger, J. C., & Waldman, M. (1989). Rational expectations in the aggregate. *Economic Inquiry, 27*, 619–636.

Hamilton, J. (1994). *Time series analysis*. Princeton: Princeton University Press.

Hansen, L. P., & Hodrick, R. J. (1980). Forward exchange rates as optimal predictors of future spot rates: An econometric analysis. *Journal of Political Economy, 88*, 829–853.

Holden, K., & Peel, D. A. (1990). On testing for unbiasedness and efficiency of forecasts. *Manchester School, 58*, 120–127.

Ito, T. (1990). Foreign exchange rate expectations: Micro survey data. *American Economic Review, 80*, 434–449.

Ito, T. (1994). Short-run and long-run expectations of the yen/dollar exchange rate. *Journal of the Japanese and International Economies, 8*, 119–143.

Keane, M. P., & Runkle, D. E. (1990). Testing the rationality of price forecasts: New evidence from panel data. *American Economic Review, 80*, 714–735.

Kirman, A. P. (1992). Whom or what does the representative individual represent? *Journal of Economic Perspectives, 6*, 117–136.

Kiviet, J. F. (1987). *Testing linear econometric models*. Ph.D. Thesis, University of Amsterdam.

Lai, K. (1990). An evaluation of survey exchange rate forecasts. *Economics Letters, 32*, 61–65.

Lamont, O. (2002). Macroeconomic forecast and microeconomic forecasters. *Journal of Economic Behavior and Organization, 48*, 265–280.

Laster, D., Bennett, P., & Geoum, I. S. (1999). Rational bias in macroeconomic forecasts. *Quarterly Journal of Economics, 293–318*.

LeBaron, B. (2000) *Technical trading profitability in foreign exchange markets in the 1990's*. Technical report, Brandeis U.

Leitch, G., & Tanner, J. E. (1991). Economic forecast evaluation: Profits versus the conventional error measures. *The American Economic Review, 81*, 580–590.

Liu, P., & Maddala, G. S. (1992). Rationality of survey data and tests for market efficiency in the foreign exchange markets. *Journal of International Money and Finance, 11*, 366–381.

Lyons, R. K. (2002). Foreign exchange: Macro puzzles, micro tools. *Economic Review, Federal Reserve Bank of San Francisco*, 51–69.

MacDonald, R. (1992). Exchange rate survey data: A disaggregated G-7 perspective. *The Manchester School, LX*(Supplement), 47–62.

Maynard, A., & Phillips, P. C. B. (2001). Rethinking and old empirical puzzle: Econometric evidence on the forward discount anomaly. *Journal of Applied Econometrics, 16*, 671–708.

Mincer, J., & Zarnowitz, V. (1969). The evaluation of economic forecasts. In *Economic forecasts and expectations*. New York: National Bureau of Economic Research.

Muth, J. (1961). Rational expectations and the theory of price movements. *Econometrica, 29*, 315–335.

Newey, W., & West, K. (1987). A simple positive semi-definite heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica, 55*, 703–708.

Osterberg, W. P. (2000). New results on the rationality of survey measures of exchange-rate expectations. *Federal Reserve Bank of Cleveland Economic Review, 36*, 14–21.

Pesaran, H. M. (2004). *General diagnostic tests for cross section dependence in panels*. CESifo working paper series 1229, CE-Sifo Group Munich.

Pesaran, H. M., & Weale, M. (2006). Survey expectations. In G. E. C. W. J. Granger & A. Timmermann (Eds.), *Handbook of economic forecasting*. Amsterdam: North-Holland.
Phillips, P., & Hansen, B. E. (1990). Statistical inference in instrumental variables regression with I(1) processes. *Review of Economic Studies, 57*, 99–125.
Phillips, P. C. B., & Maynard, A. (2001). Rethinking an old empirical puzzle: Econometric evidence on the forward discount anomaly. *Journal of Applied Econometrics, 16*, 671–708.
Pilbeam, K. (1995). Exchange rate models and exchange rate expectations: an empirical investigation. *Applied Economics, 27*, 1009–1015.
Saikkonen, P. (1991). Asymptotically efficient estimation of cointegration regressions. *Econometric Theory, 7*, 1–21.
Stockman, A. C. (1987). Economic theory and exchange rate forecasts. *International Journal of Forecasting, 3*, 3–15.
Theil, H. (1966). *Applied economic forecasting*. Amsterdam: North-Holland.
Zacharatos, N., & Sutcliffe, C. (2002). Is the forward rate for the Greek Drachma unbiased? A VECM analysis with both overlapping and non-overlapping data. *Journal of Financial Management and Analysis, 15*, 27–37.
Zarnowitz, V. (1985). Rational expectations and macroeconomic forecasts. *Journal of Business & Economic Statistics, 3*, 293–311.
Zellner, A (1962). An efficient method of estimating seemingly unrelated regressions and tests for aggregation bias. *Journal of the American Statistical Association*, June, 348–368.
Zellner, A. (1986). Biased predictors, rationality and the evaluation of forecasts. *Economics Letters, 21*, 45–48.
Zivot, E. (2000). Cointegration and forward and spot exchange rate regressions. *Journal of International Money and Finance, 19*, 785–812.

Stochastic Volatility Structures and Intraday Asset Price Dynamics

44

Gerard L. Gannon

Contents

Abstract

The behavior of financial asset price data when observed intraday is quite different from these same processes observed from day to day and longer sampling intervals. Volatility estimates obtained from intraday observed data can be badly distorted if anomalies and intraday trading patterns are not accounted for in the estimation process.

In this paper I consider conditional volatility estimators as special cases of a general stochastic volatility structure. The theoretical asymptotic distribution of the measurement error process for these estimators is considered for particular

G.L. Gannon
Deakin University, Burwood, VIC, Australia
e-mail: gerard@deakin.edu.au; gleonon@yahoo.com

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_44,

features observed in intraday financial asset price processes. Specifically, I consider the effects of (i) induced serial correlation in returns processes, (ii) excess kurtosis in the underlying unconditional distribution of returns, (iii) market anomalies such as market opening and closing effects, and (iv) failure to account for intraday trading patterns.

These issues are considered with applications in option pricing/trading strategies and the constant/dynamic hedging frameworks in mind. Empirical examples are provided from transactions data sampled into 5-, 15-, 30-, and 60-min intervals for heavily capitalized stock market, market index, and index futures price processes.

Keywords

ARCH • Asymptotic distribution • Autoregressive parameters • Conditional variance estimates • Constant/dynamic hedging • Excess kurtosis • Index futures • Intraday returns • Market anomalies • Maximum likelihood estimates • Mis-specification • Mis-specified returns • Persistence • Serial correlation • Stochastic volatility • Stock/futures • Unweighted GARCH • Volatility co-persistence

44.1 Introduction

One issue considered in Nelson (1990a) is whether it is possible to formulate an ARCH data generation process that is similar to the true process, in the sense that the distribution of the sample paths generated by the ARCH structure and the underlying diffusion process becomes “close” for increasingly finer discretizations of the observation interval. Maximum likelihood estimates are difficult to obtain from stochastic differential equations of time-varying volatility common in the finance literature. If the results in Nelson hold for “real-time” data when ARCH structures approximate a diffusion process, then these ARCH structures may be usefully employed in option pricing equations. In this paper I consider the ARCH structure as a special case of a general stochastic volatility structure. One advantage of an ARCH structure over a general stochastic volatility structure lies in computational simplicity. In the ARCH structure, it is not necessary for the underlying processes to be stationary or ergodic. The crucial assumption in an option pricing context is that these assumed processes approach a diffusion limit. These assumed diffusion limits have been derived for processes assumed to be observed from day-to-day records. Given that market anomalies such as market opening and market closing effects exist, any volatility structure based on observations sampled on a daily basis will provide different volatility estimates. Evidence of these intraday anomalies and effects on measures of constant volatility is reported in Edwards (1988) and Duffie et al. (1990). Brown (1990) argues that the use of intraday data in estimating volatility within an option pricing framework leads to volatility estimates that are too low. This can be overcome by rescaling, assuming the anomalies are accounted for.

Better estimates of volatility may then be obtained by employing intraday observations and allowance made for anomalies and trading activity within conditional volatility equations. However, mis-specifications in either or both first- and second-moment equations may mitigate against satisfying the conditions for an approximate diffusion limit. Then it is important to investigate cases where the diffusion limit is not attainable and identify the factors which help explain the behavior of the process. If these factors can be accounted for in the estimation process then these formulations can be successfully employed in options pricing and trading strategies.

The specific concern in this paper is the effect on the asymptotic distribution of the measurement error process and on parameter estimates, obtained from the Generalized ARCH (GARCH(1,1)) equations for the conditional variance, as the observation interval approaches transactions records ($d \rightarrow 0$). Three issues are considered for cases where the diffusion limit may not be achieved at these observation intervals. The first issue is the effect of mis-specifying the dynamics of the first-moment-generating equation on resultant GARCH(1,1) parameter estimates. The second issue is the effect on measures of persistence obtained from the GARCH structure when increasing kurtosis is induced in the underlying unconditional distribution as $d \rightarrow 0$. This leads to a third issue which is concerned with evaluating effects of inclusion of weighting (mixing) variables on parameter estimates obtained from these GARCH(1,1) equations. If these mixing variables are important then standard, GARCH equation estimates will be seriously distorted. These mixing variables may proxy the level of activity within particular markets or account for common volatility of assets trading in the same market.

Sampling the process too finely does result in induced positive or negative serial correlation in return processes. The main distortion to the basis change is generated from cash index return equations. However, the dominant factor distorting unweighted GARCH estimates is induced excess kurtosis in unconditional distributions of returns. Many small price changes are dominated by occasional large price changes. This effect leads to large jumps in the underlying distribution causing continuity assumptions for higher derivatives of the conditional variance function to break down.

These observations do not directly address issues related to intraday market trading activity and possible contemporaneous volatility effects transmitted to and from underlying financial asset price processes. This effect was considered within the context of a structural Simultaneous Volatility (SVL) model in Gannon (1994). Further results for the SVL model are documented, along with results of parameter estimates, in Gannon (2010). In this paper the intraday datasets on cash index and futures price processes from Gannon (2010) are again employed to check the effects on parameter estimates obtained from GARCH and weighted GARCH models. A set of intraday sampled stock prices are also employed in this paper. The relative importance of mis-specification of the second-moment equation dynamics over mis-specification of first-moment equation dynamics is the most important issue. If intraday trading effects are important, this has implications for smoothness and continuity assumptions necessary in deriving diffusion limit results for the unweighted GARCH structure.

If these effects are severe then an implied lower sampling boundary needs to be imposed in order to obtain sensible results. This is because the measure of persistence obtained from GARCH structures may approach the Integrated GARCH (IGARCH) boundary and become explosive or conditional heteroskedasticity may disappear. This instability can be observed when important intraday anomalies such as market opening and closing effects are not accounted for within the conditional variance specifications. Distortions to parameter estimates are most obvious when conditional second-moment equations are mis-specified by failure to adequately account for observed intraday trading patterns. These distortions can be observed across a wide class of financial assets and markets.

If these financial asset price processes have active exchange traded options contracts written on these "underlying assets," it is important to study the intraday behavior of these underlying financial asset price processes. If systematic features of the data series can be identified then it is possible to account for these features in the estimation process. Then intraday estimates of volatility obtained from conditional variance equations, which incorporate structural effects in first and/or conditional second-moment equations, can be usefully employed in option pricing equations. Identifying intraday anomalies and linking trading activity to contemporaneous volatility effects mean the improved estimator can be employed within a trading strategy. This can involve analysis of the optimal time to buy options within the day to minimize premium cost. Alternatively, optimal buy or sell straddle strategies based on comparison of estimated volatility estimates relative to market implied volatility can be investigated. In this paper I focus on theoretical results which can explain the empirically observed behavior of these estimators when applied to intraday financial asset price processes. I start by summarizing relevant results which are currently available for conditional variance structures as the observation interval reduces to daily records (h→0). These results are modified and extended in order to accommodate intraday observation intervals. The alternative first-moment-generating equations are described and the basis change defined and discussed within the context of the co-persistence structure. I then focus on the general GARCH structure and state some further results for specific cases of the GARCH and weighted GARCH (GARCH-W) structure.

44.2 Stochastic Volatility and GARCH

Nelson and Foster (1994) derive and discuss properties for the ARCH process as the observation interval reduces to daily records (h→0) when the underlying process is driven by an assumed continuous diffusion process. Nelson and Foster (1991) generalized a Markov process with two state variables, ${}_hX_t$ and ${}_h\sigma_t^2$, only one of which ${}_hX_t$ is ever directly observable. The conditional variance ${}_h\sigma_t^2$ is defined conditional on the increments in ${}_hX_t$ per unit time and conditional on an information set ${}_h\varsigma_t$. Modifying the notation from h to d (to account for intraday discretely observed data), and employing the notation $d \to 0$ to indicate reduction in the observation interval from above, when their assumptions 2, 3, and 1′ hold, when d is small, ${}_dX_t$, $\varphi({}_d\sigma_t^2))$ is referred to as a near diffusion if for any T, $0 \le T \le \infty$, $({}_dX_t,\ \varphi\ ({}_d\sigma_t^2))_{0 \le t \le T} \Rightarrow (X_t,\ \varphi(\sigma_t^2))_{0 \le t \le T}$.

If we assume these data-generating processes are near diffusions, then the general discrete time stochastic volatility structure, defined in Nelson and Foster (1991), may be described using the following modified notation:

$$\begin{bmatrix} {}_dX_{(k+1)d} \\ \varphi\left({}_d\sigma^2_{(k+1)d}\right) \end{bmatrix} = \begin{bmatrix} {}_dX_{kd} \\ \varphi\left({}_d\sigma^2_{kd}\right) \end{bmatrix} + d\ .\begin{bmatrix} \mu({}_dX_{kd}, \sigma_{kd}) \\ \lambda({}_dX_{kd}, {}_d\sigma_{kd}) \end{bmatrix} + d^{1/2}\begin{bmatrix} {}_d\sigma^2_{kd} & \Lambda_{\varphi,x}({}_dX_{kd}, {}_d\sigma_{kd}) \\ \Lambda_{\varphi,x}({}_dX_{kd}, {}_d\sigma_{kd}) & \Lambda^2({}_dX_{kd}, {}_d\sigma_{kd}) \end{bmatrix}^{\frac{1}{2}} \begin{bmatrix} {}_dZ_{1,kd} \\ {}_dZ_{2,kd} \end{bmatrix} \tag{44.1}$$

where $({}_dZ_{1,kd},{}_d\ Z_{2,kd})_{k=0,\infty}$ is i.i.d. with mean zero and identity covariance matrix. In Eq. 44.1 d is the size of the observation interval, X may describe the asset price return, and σ^2 the volatility of the process. It is not necessary to assume the data-generating processes are stationary or ergodic, but the crucial assumption is that the data-generating processes are near diffusions.

In the ARCH specification, ${}_dZ_{2,kd}$ is a function of ${}_dZ_{1,kd}$ so that ${}_d\sigma^2_{kd}({}_dh_{kd})$ can be inferred from past values of the one observable process ${}_dX_{kd}$. This is not true for a general stochastic volatility structure where there are two driving noise terms.

For the first-order Markov ARCH structure, a strictly increasing function of estimates ${}_d\hat{\sigma}_t^2\left({}_d\hat{h}_t\right)$ of the conditional variance process ${}_d\sigma_t^2({}_dh_t)$ is defined as $\phi(\sigma^2)$, and estimates of the conditional mean per unit of time of the increments in X and $\phi(\sigma^2)$ are defined as $\hat{\mu}\left(x,\ \hat{\sigma}\right)$ and $\hat{\kappa}\left(x,\ \hat{\sigma}\right)$. Estimates of ${}_d\sigma^2_{kd}$ are updated by the recursion:

$$\phi\left({}_d\hat{\sigma}^2_{(k+1)d}\right) = \sigma\left({}_d\hat{\sigma}^2_{kd}\right) + d\,\hat{\kappa}\left({}_dX_{kd}, {}_d\hat{\sigma}_{kd}\right) + d^{1/2}a\left({}_dX_{kd}, {}_d\hat{\sigma}_{kd}\right)g\left({}_dZ_{1,kd}, {}_d\hat{X}_{kd}, {}_d\hat{\sigma}^2_{kd}\right) \tag{44.2}$$

where $\hat{\kappa}\,(.), a(.), \hat{\mu}\,(.)$, and g(.) are continuous on bounded $(\varphi(\sigma^2),\ x)$ sets and $g(z_1,\ x.\sigma^2)$ assumed continuous everywhere with the first three derivatives of g with respect to z_1 well defined and bounded. The function $g({}_dZ_{1,kd},.)$ is normalized to have mean zero and unit conditional variance. Nonzero drifts in $\phi\left({}_d\sigma^2_{kd}\right)$ are allowed for in the $\hat{\kappa}\,(.)$ term and non-unit conditional variances accounted for in the a(.) term. The second term on the right measures the change in $\phi\left({}_d\sigma^2_{kd}\right)$ forecast by the ARCH structure while the last term measures the surprise change.

The GARCH(1,1) structure is obtained by setting $\phi(\sigma^2) = \sigma^2$, $\hat{\kappa}(x,\sigma) = (\omega - \theta\sigma^2)$, $g(z_1) = (z_1^2 - 1)/\mathrm{SD}(z_1^2)$, and $a(x,\sigma) = \alpha \,.\, \sigma^2 \,.\, \mathrm{SD}(z_1^2)$. The parameters ω, θ, and α index a family of GARCH(1,1) structures. In a similar manner the EGARCH volatility structure can be defined. That is, by selecting $\phi(.)$, $\hat{\kappa}(.)$, g(.), and a(.), various "filters" can be defined within this framework.

Other conditions in Nelson and Foster (1991) define the rate at which the normalized measurement error process ${}_dQ_t$ mean reverts relative to $({}_dX_t, {}_d\sigma_t^2)$. It becomes white noise as d→0 on a standard time scale but operates on a faster time scale mean-reverting more rapidly. Asymptotically optimal choice of a(.) and ϕ(.) given g(.) can be considered with respect to minimizing the asymptotic variance of the measurement error. This is considered on a faster time scale (T + τ) than T. The asymptotically optimal choice of g(.) depends upon the assumed relationship between Z_1 and Z_2. In the ARCH structure, Z_2 is a function of Z_1 so that the level driving ${}_d\sigma_t^2$ can be recovered from shocks driving ${}_dX_t^2$. Without further structure in the equation specified for σ_t^2, we are unable to recover information about changes in σ_t^2. Their discussion is strictly in terms of constructing a sequence of optimal ARCH filters which minimize the asymptotic variance of the asymptotic distribution of the measurement errors. This approach is not the same as choosing an ARCH structure that minimizes the measurement error variance for each d.

The asymptotic distribution of the measurement error process, for large τ and small d,

$$\left[{}_d\hat{\sigma}^2_{T+\tau d^{1/2}} - {}_d\sigma^2_{T+\tau d^{1/2}} \,\middle|\, \left({}_dQ_T, {}_d\sigma_T^2, {}_dX_T\right)\right] = \left(q, \sigma^2, x\right)$$

with derivatives evaluated as $\phi'\left({}_d\sigma_T^2\right)$, $\varphi'\left({}_d\sigma_T^2\right)$, etc. and the notation simplified as ϕ' and φ' is approximately normal with mean

$$d^{1/2}\frac{(2\sigma^2\varphi')\left[\hat{\kappa}_d/\varphi' - a^2\varphi''/2(\varphi')^3 - \lambda_d/\phi' + \Lambda^2\phi'^3\right] + 2a\sigma.\left[\mu_d - \hat{\mu}_d\right].E[g_z]}{a\,.\,E[Z_1\,.\,g_z]} \tag{44.3}$$

and variance

$$d^{1/2}\frac{(2\sigma^2\phi')\,.\,(a^2/\phi') + \lambda^2/[\varphi']^2 - 2a\Lambda.\mathrm{Cov}(Z_2, g)/[\sigma'][\varphi']}{a.E[Z_1.g_z]}. \tag{44.4}$$

General results in Nelson and Foster (1991, 1994) for the GARCH(1,1) structure are that GARCH(1,1) can be more accurately measured firstly the less variable and the smaller is ${}_d\sigma_t^2$, second the thinner the tails of Z_1, and third the more the true data-generating mechanism resembles an ARCH structure as opposed to a stochastic volatility structure. If the true data-generating process is GARCH(1,1), then Corr $(Z_1^2, Z_2) = 1$.

As d→0, the first result will generally hold, and the second can be checked from the unconditional distribution of the returns process. The latter result is the most difficult to evaluate. Now reconsider some assumptions necessary to obtain these results and reasons these assumptions may not hold when d→0.

(i) Mis-specification of the difference between the estimated and true drift in mean, $\left[{}_d\hat{\mu}_t - {}_d\mu_t\right]$, is assumed fixed as d→0 so that effects of mis-specifying this drift has an effect that vanishes at rate $d^{1/2}$ and is negligible asymptotically. These terms do not appear in the expression for the variance of the asymptotic distribution of the measurement error. As d→0, the effect of bid/ask bounce and order splitting in futures price processes and non-trading-induced effects on market indices becomes more severe. Mis-specification of the drift in the mean is not constant. Whether this effect transfers to estimates of conditional variances is an empirical issue.

(ii) The conditional variance of the increments in ${}_d\sigma_t^2$ involves the fourth moment of ${}_dZ_{1,kd}$ so that the influence of this fourth moment remains as the diffusion limit is approached. Excess kurtosis is a feature of intraday financial price changes.

(iii) Values of $\hat{\kappa}_d$ and λ_d are considered fixed as d→0 so that effects of mis-specifying the drift in σ^2 has an effect that also vanishes at rate $d^{1/2}$. As well, although these drift terms enter the expression for the asymptotic bias of the measurement error, these also do not appear in the expression for the asymptotic variance. The term g_z represents part of the "surprise" change in the recursion defined in Eq. 44.2 and is directly linked to departures from normality observed in point (ii). These departures from normality can be generated by extremes in Z_1 induced by large jumps in the underlying distribution. In this case first and second derivatives of ϕ may be discontinuous throughout the sample space as well. Then the expression for the bias in this asymptotic distribution of the measurement errors may be explosive.

(iv) The ARCH specification of the drift in mean and variance only enters the $0_p(d^{1/2})$ terms of the measurement error. Asymptotically, the differences in the conditional variance specifications are more important, appearing in the $0_p(d^{1/4})$ terms. If the conditional variance specification is not correct then the measurement error variance is affected. This is because matching the ARCH and true variance of the variance cannot proceed.

I will consider these issues further by generalizing a theoretical framework in which to address each in the above order. Firstly I consider the relationship between serial correlation in returns on the market index, the index futures, and basis change as d→0. Second, these effects are considered in the context of the co-persistence structure for the basis change. Finally, I consider effects on conditional variance parameter estimates when mixing and weighting variables are included in the equations for the conditional variance.

44.3 Serial Correlation in Returns

Consider a first-order autoregressive process for the spot asset price return (an AR(D) representation):

$$\mathrm{s_t} = \rho_1 \mathrm{s_{t-1}} + \mathrm{e_t} \tag{44.5}$$

where $\mathrm{s_t}$ is the return on the asset (the difference in the natural logarithm of the spot asset price levels or the difference in the spot asset price levels) between time t and t−1, ρ_1 is the first-order serial correlation coefficient between s at time t and t−1 and $\mathrm{e_t}$ is an assumed homoskedastic disturbance term, σ_e^2 and $-1 < \rho_1 < 1$. This equation provides an approximation to the time-series behavior of the spot asset price return.

An alternative specification to Eq. 44.5 is a simple autoregressive equation for the level of the spot asset price S (or natural logarithm of the levels) at time t and t−1, (an AR(L) representation)

$$S_t = \rho_1 S_{t-1} +_t a_t \tag{44.6}$$

where $\mathrm{a_t}$ is an assumed homoskedastic disturbance term, σ_a^2, and the value of ϕ_1 may be greater or less than one.

The assumed bid/ask bounce in futures price changes is approximated by a MA (1) process in Miller et al. (1994). For the index futures price, this specification is

$$f_t = a_t + \theta_1 a_{t-1}, \tag{44.7}$$

where $\mathrm{f_t}$ is the index futures price change and $\mathrm{a_t}$ is an assumed mean zero, serially uncorrelated shock variable with a homoskedastic variance, σ_a^2, and $-1 < \theta_1 < 0$.

The basis change is defined as

$$b_t = f_t - i_t, \tag{44.8}$$

where *f* and *i* are the index futures return and index portfolio return, respectively.

Whether shocks generated from Eqs. 44.5 or 44.6 generate differences in parameter estimates and measures of persistence obtained from conditional second-moment equations is an issue. Measures of persistence for the basis change may be badly distorted from employing index futures and index portfolio changes. Measures of persistence for index futures and index portfolios may be badly distorted by employing observed price changes.

The simplest GARCH structure derived from Eq. 44.1 for the conditional variance is the GARCH(1,1):

$$h_t = \omega + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 h_{t-1} \tag{44.9}$$

where $\mathrm{h_t}(\sigma_t^2)$ is the conditional variance at time t and ε_{t-1}^2 are squared unconditional shocks generated from any assumed first-moment equation and $0 \leq \alpha_1, \beta_1 \leq 1$ and

$\alpha_1 + \beta_1 \leq 1$. This parameterization is a parsimonious representation of an ARCH (p') process where a geometrically declining weighting pattern on lags of ε^2 is imposed. This is easily seen by successive substitution for $h_{t-j}(j = 1, \ldots, J)$ as $J \rightarrow \infty$,

$$h_t = \omega\left(1 + \beta_1 + \cdots + \beta_1^J\right) + \alpha_1\left(\varepsilon_{t-1}^2 + \beta_1\varepsilon_{t-2}^2 + \cdots + \beta_1^J\varepsilon_{t-J-1}^2\right) + \text{Remainder}. \tag{44.10}$$

Now consider Eq. 44.6, with ϕ_1 fixed at 1, as representing one mis-specified spot asset price process, Eq. 44.5 representing the mis-specified differenced autoregressive process, and Eq. 44.7 representing the mis-specified differenced moving average process. Taking expected values, then the unconditional variance when Eq. 44.6 is the mis-specified representation and ϕ_1 set equal to one is

$$E\left(s_t^2\right) = E\left(a_t^2\right). \tag{44.11}$$

When Eq. 44.7 is the moving average (MA) representation, the unconditional variance relative to shocks generated via Eq. 44.11 is

$$E\left(s_t^2\right)_{MA} = E\left[\left(1 + \theta^2\right)a_t^2\right], \tag{44.12}$$

and when Eq. 44.5 is the autoregressive (AR) representation, the unconditional variance relative to shocks generated via Eq. 44.11 is

$$E\left(s_t^2\right)_{AR} = E\left[\left(\frac{1 - \rho_1}{1 + \rho_1}\right)\left(a_t^2\right)\right]. \tag{44.13}$$

The conditional variance from a GARCH(1,1) structure for Eq. 44.11 can be rewritten as

$$\begin{aligned} h_t(s) = {} & \omega\left(1 + \beta_1 + \cdots + \beta_1^J\right) + \alpha_1\left(a_{t-1}^2 + \beta_1 a_{t-2}^2 + \cdots + \beta_1^J a_{t-J-1}^2\right) \\ & + \text{Remainder}. \end{aligned} \tag{44.14}$$

If Eq. 44.7 is the representation, then, relative to the conditional variance equation from Eq. 44.11,

$$\begin{aligned} h_t(s)_{MA} = {} & \omega\left(1 + \beta_1 + \cdots + \beta_1^J\right) + \alpha_1\left(1 + \theta^2\right)a_{t-1}^2 + \beta\left(1 + \theta^2\right)a_{t-2}^2 \\ & + \cdots\beta_1^J\left(1 + \theta^2\right)a_{t-J-1}^2\Big) + \text{Remainder}, \end{aligned} \tag{44.15}$$

and if Eq. 44.5 is the representation, then, relative to the conditional variance equation from Eq. 44.11,

$$\begin{aligned} \mathrm{h_t(s)_{AR}} &= \omega\left(1+\beta_1+\cdots+\beta_1^J\right) \\ &\quad + \alpha_1\left(\left(\frac{1-\rho_1}{1+\rho_1}\right)a_{t-1}^2 + \beta_1\left(\frac{1-\rho_1}{1+\rho_1}\right)a_{t-2}^2 + \cdots \beta_1^J\left(\frac{1-\rho_1}{1+\rho_1}\right)a_{t-J-1}^2\right) + \text{Remainder}. \end{aligned} \tag{44.16}$$

If ω, α_1, and β_1 were equivalent in Eqs. 44.14, 44.15, and 44.16, when the conditional variance is driven by Eq. 44.7 with θ_1 negative, then $h_t(s)_{MA} > h_t(s)$, and when the conditional variance is driven by Eq. 44.5 with ρ_1 negative, then $h_t(s)_{AR} > h_t(s)$ and with ρ_1 positive then $h_t(s)_{AR} < h_t(s)$. However, given the scaling factor in Eq. 44.15 relative to Eq. 44.16, the potential for distortions to GARCH parameter estimates is greater when the underlying process is driven by Eq. 44.5 relative to Eq. 44.7.

44.4 Persistence, Co-Persistence, and Non-Normality

Now define ε_t as shocks from any of the assumed first-moment equations from Sect. 3 with the following simplified representation of a GARCH(1,1) structure obtained from Eq. 44.1 where $h_t(\sigma_t^2)$ represents the conditional variance and $z_t({}_dZ_{1\ ,kd})$ the stochastic part:

$$\varepsilon_t = \sqrt{h_t}\, z_t \quad z_t \sim \mathrm{NID}(0,1). \tag{44.17}$$

In the univariate GARCH(1,1) structure, h_t converges and is strictly stationary if $E[\ln(\beta_1 + \alpha_1 z_{t-i}^2)] < 0$. Then $\sum_{i=1,kd} \ln(\beta_1 + \alpha_1 z_{t-i}^2)$ is a random walk with negative drift which diverges to $-\infty$ as the observation interval reduces.

Now consider the co-persistence structure in the context of the constant hedging model. Defining the true processes for the differences in the natural logarithm of the spot index price and the natural logarithm of the futures price as

$$\begin{aligned} i_t &= \gamma_1 \xi_t + \eta_{it} \\ f_t &= \gamma_2 \xi_t + \eta_{ft}, \end{aligned} \tag{44.18}$$

the common "news" factor ξ_t is IGARCH, in the co-persistence structure, while the idiosyncratic parts are assumed jointly independent and independent of ξ_t and not IGARCH. The individual processes have infinite unconditional variance. If a linear combination is not IGARCH, then the unconditional variance of the linear combination is finite and a constant hedge ratio (defined below) leads to substantial reduction in portfolio risk.

A time-varying hedge ratio can lead to greater reduction in portfolio risk under conditions discussed in Ghose and Kroner (1994) when the processes are co-persistent in variance. From a practical perspective, account needs to be taken of the rebalancing costs of portfolio adjustment.

A nonoptimal restricted linear combination is the basis change defined as the difference between the change in the log of the index futures price and change in the log of the spot index level. This implied portfolio is short 1 unit of the spot for every unit long in the futures. For the futures and spot price processes reported in McCurdy and Morgan (1987), the basis change is co-persistent in variance.

If there are "news factors" $\xi_{f\,t} \neq \xi_{i\,t}$, then the constant hedge ratio may not exist. Define these processes as

$$\begin{aligned} i_t &= \gamma_1 \xi_{i\,t} + \eta_{i\,t} \\ f_t &= \gamma_2 \xi_{f\,t} + \eta_{f\,t} \end{aligned} \tag{44.19}$$

then the estimated constant hedge ratio which is short g units of the spot for every 1 unit long in the futures is

$$\hat{\gamma} = \left[\frac{\hat{\mathrm{cov}}\ (\mathrm{f}, \mathrm{i})}{\hat{\mathrm{var}}\ (\mathrm{i})}\right]_{|\mathrm{t}} = \left[\hat{\gamma}_2\left[\frac{\mathrm{p}\sqrt{\hat{\mathrm{var}}\ [\xi_\mathrm{f}]\ \hat{\mathrm{var}}\ [\xi_\mathrm{i}]}}{\hat{\mathrm{var}}\ [\xi_\mathrm{i}]}\right] + \mathrm{var}\left(\hat{\eta}_i\right)\right]_{|\mathrm{t}} \tag{44.20}$$

where $\hat{\rho}$ is the correlation between ξ_{ft} and ξ_{it}.

When both ξ_f and ξ_i follow IGARCH processes and no common factor structure exists, then the estimated constant hedge diverges. Ghose and Kroner (1994) investigate this case.

When ξ_{ft} follows an IGARCH process but ξ_{it} is weak GARCH, then the estimated constant hedge ratio cannot be evaluated. There are two problems:

(a) The estimated sample variance of ξ_f in Eq. 44.20 is infinite as $T \to \infty$.

(b) $p = \left[\left[c\ o\hat{v}\left[\xi_f\right]\right] / \sqrt{v}\ \hat{a}\ r\left[\xi_f\right]\ v\hat{a}r\ \left[\xi_i\right]\right]_{|t}$ so that there is no linear combination of ξ_f and ξ_i which can provide a stationary unconditional variance.

This last observation has a direct parallel from the literature for cointegration in the means of two series.

If ξ_f is an approximate I(1) process and ξ_i is I(0), then there is no definable linear combination of ξ_f and ξ_i.

When observing spot index and futures prices over successively finer intervals, the co-persistence structure may not hold for at least two further reasons. This argument relates directly to the horizon t+kd for the hedging strategy. This argument also relates directly to distortions possibly induced onto a dynamic hedging strategy, as d→0.

Perverse behavior can be observed in spot index level changes as oversampling becomes severe. The smoothing effect due to a large proportion of the portfolio entering the non- and thin-trading group generates a smoothly evolving process with short-lived shocks generated by irregular news effects.

General results assume that the z_t's are drawn from a continuous distribution. When sampling futures price data at high frequency, then the discrete nature of the price recording mechanism guarantees that there are discontinuities in return-generating processes. The distribution of the z_t's can become extremely peaked due to multiple small price changes and can have very long thin tails due to abrupt shifts in the distribution. As $d \to 0$, $-\infty < E[1n(\beta_1 + \alpha_1 z_{t-i}^2)] < +\infty$ for a large range of values for $0 \leq \alpha_1$, $\beta_1 \leq 1$ and $\alpha_1 + \beta_1 < 1$, and this depends on the distribution induced by oversampling and resultant reduction in $(\alpha_1 + \beta_1)$. In the limit $E[(z_t)^4]/[E(z_t^2)]^2 \to \infty$. Then even-numbered higher moments of z_t are unbounded as $d \to 0$. This oversampling can lead to two extreme perverse effects generated by bid/ask bounce or zero price changes. The effect depends upon liquidity in the respective markets.

44.4.1 Case 1

$$\alpha_1 \to 1, \alpha_1 + \beta_1 > 1 \text{ and } E\left[\ln\left(\beta_1 + \alpha_1 z_{t-i}^2\right)\right] > 0.$$

The intuitive explanation for this result relies on oversampling (not overdifferencing) in highly liquid markets. The oversampling approaches analysis of transactions. At this level bid/ask bounce and order splitting require an appropriate model. Any arbitrary autoregressive model, for unconditional first moments, generates unconditional shocks relating predominantly to behavior of the most recent shock. These effects carry through to conditional squared innovations.

44.4.2 Case 2

Oversampling can produce many zero price changes in thin markets. In this latter case as $d \to 0$ then $\alpha_1 + \beta_1 \to 0$.

This explanation can apply to relatively illiquid futures (and spot asset) price changes. That is, conditional heteroskedasticity disappears as oversampling becomes severe.

The effect on the basis change when there is a relatively illiquid futures market and oversampling may be badly distorted. As well, failure to account for anomalies in conditional variance equations can severely distort estimates.

44.5 Weighted GARCH

Recall from Eq. 44.1 that $\hat{\kappa}$ is a function defining the estimated drift in $\left\{\phi\left({}_d\sigma_t^2\right)\right\}$ so that λ is a function defining true drift in $\{\varphi({}_d\sigma_t^2)\}$. In the ARCH structure, the drift in $h_t(\sigma_t^2)$ in the diffusion limit is represented by $\hat{\kappa}/\phi' - a^2\phi''/2(\phi')^3$, whereas for the stochastic differential equation defined from assumptions 2, 3, and 1′ in Nelson and

Foster (1991, 1994), this diffusion limit is $\lambda/\varphi' + \Lambda^2\varphi''/2\,(\varphi')^3$. The effect on the expression for the bias in the asymptotic distribution of the measurement error process can be explosive if derivatives in the terms $a^2\phi''/2(\phi')^3 - \Lambda^2\varphi''/2(\varphi')^3$ cannot be evaluated because of discontinuities in the process. This can happen when important intraday effects are neglected in the conditional variance equation specification. As well, the bias can diverge as $d \to 0$ if the ${}_dZ_{1,kd}$ terms are badly distorted.

Occasional large jumps in the underlying distribution contribute large $O_p(1)$ movements while the near diffusion components contribute small $O_p(d^{1/2})$ increments. When sampling intraday financial data, there are often many small price changes which tend to be dominated by occasional large shifts in the underlying distribution.

Failure to account for intraday effects (large shocks to the underlying distribution) can lead to a mixture of $O_p(1)$ and $O_p(d^{1/2})$ effects in the process. One approach is to specify this mixed process as a jump diffusion. An alternative is to account for these effects by incorporating activity measures in the specification of conditional variance equations.

It follows that failure to account for these $O_p(1)$ effects can lead to an explosive measurement error $\left[{}_d\hat{h}_t - {}_dh_t\right]$. However, in empirical applications this measurement error is unobservable since ${}_dh_t$ is unobservable. Failure to account for these jumps in the underlying distribution imply that the unweighted GARCH(1,1) structure cannot satisfy the necessary assumptions required to approximate a diffusion limit.

44.6 Empirical Examples

The first issue is the effect of possible mis-specification of the first-moment equation dynamics and resultant effect on estimates of persistence of individual processes as $d \to 0$. If the effect is not important (mean irrelevance) then the focus of attention is on the estimates from the co-persistence structure and implications for a constant hedge ratio.

The second issue is the effect of inclusion of variables to proxy intraday activity on measures of persistence. However, it is still important to consider possible effects from mis-specifying the first-moment equation on parameter estimates obtained from conditional variance equations as $d \to 0$. If there is strong conditioning from measures of activity onto the market price processes, the conditioning should be independent of specification of the dynamics of the first-moment equation (mean irrelevance).

44.6.1 Index Futures, Market Index, and Stock Price Data

The data has been analyzed for the common trading hours for 1992 from the ASX and SFE, i.e., from 10.00 a.m. to 12.30 p.m. and from 2.00 p.m. to 4.00 p.m.

This dataset was first employed in examples reported in Gannon (1994) and a similar analysis undertaken in Gannon (2010) for the SVL models as is undertaken in this paper for the GARCH and GARCH-W models. Further details of the sampling and institutional rules operating in these markets are briefly reported in the Appendix of this paper. Transactions for the Share Price Index futures (SPI) were sampled from the nearest contract 3 months to expiration. The last traded price on the SPI levels and stock price levels were then recorded for each observation interval. The All Ordinaries Index (AOI) is the recorded level at the end of each observation interval. During these common trading hours, the daily average number of SPI futures contracts traded 3 months to expiration was 816. As well, block trades were extremely rare in this series. Transactions on the heavily capitalized stock prices were extremely dense within the trading day.

These seven stocks were chosen from the four largest market capitalized groupings according to the ASX classification code in November 1991, i.e., general industrial, banking, manufacturing, and mining. The largest capitalized stocks were chosen from the first three categories as well as the four largest capitalized mining stocks. This selection provides for a diversified portfolio of very actively traded stocks which comprised 32.06 % of total company weights from the 300 stocks comprising the AOI.

All datasets were carefully edited in order to exclude periods where the transaction capturing broke down. The incidence of this was rare. As well, lags were generated and therefore the effects of overnight records removed. A natural logarithmic transformation of the SPI and AOI prices is undertaken prior to analysis.

Opening market activity for the SPI is heaviest during the first 40 min of trading. Trade in the SPI commences at 9.50 a.m. but from 10.30 a.m. onwards volume of trade tapers off until the lunchtime close. During the afternoon session, there is a gradual increase in volume of trade towards daily market close at 4.10 p.m. Excluding the market opening provides the familiar U-shaped pattern of intraday trading volume observed on other futures markets. SPI price volatility is highest during the market opening period with two apparent reverse J-shaped patterns for the two daily trading sessions (small J effect in afternoon session). The first and last 10 min of trade in the SPI are excluded from this dataset.

Special features govern the sequence at which stocks open for trade on the ASX. Individual stocks are allocated a random opening time to within plus or minus 30 s of a fixed opening time. Four fixed opening times, separated by 3-min intervals starting at 10.00 a.m., operated throughout 1992. Four alphabetically ordered groups then separately opened within the first 10 min of trading time. The last minute of trading on the ASX is also subject to a random closing time between 3.59 p.m. and 4.00 p.m. The effect of these institutional procedures on observed data series can be potentially severe.

Both of these activity effects in market opening prices of trading on the SFE, and ASX should be accounted for in the estimation process.

Table 44.1 Autoregressive parameter estimates for SPI, AOI, and basis change

Interval	f_t	i_t	b_t
Full day			
30 min	−0.0324	0.0909	−0.1699
	(−1.54)	(4.33)	(−8.19)
15 min	−0.0497	0.0953	−0.1348
	(−3.45)	(6.42)	(−9.15)
05 min	0.0441	0.3317	−0.0137
	(5.14)	(41.2)	(−1.60)
Excluding market open			
30 min	−0.0299	0.1213	−0.1947
	(−1.34)	(5.48)	(−8.89)
15 min	−0.0301	0.2531	−0.2145
	(−1.91)	(16.6)	(−13.9)
05 min	0.0532	0.3284	−0.0968
	(5.84)	(38.2)	(−10.7)

Asymptotic t-statistics in brackets
$f_t = F_t - F_t$ is the difference in the observed log level of the SPI
$i_t = I_t - I$ is the difference in the observed log level of the AOI
Equation 44.5, i.e., an autoregressive specification for the differences, an [AR(D)], is estimated for both data series with one lag only for the autoregressive parameter
Dummy variables are included, for the first data series, in order to account for market opening effects for the SPI, institutional features governing market opening on the ASX and therefore effects transmitted to the basis change. Two separate dummy variables are included for the first pair of 5-min intervals
This form for the basis change is $b_t = f_t - i_t$
In the lower panel, results are reported from synchronized trading from 10.30 a.m. That is, the first 40 min and first 30 min of normal trade in the SPI and AOI, respectively, is excluded

44.6.2 Estimates of the Autoregressive Parameters

In Table 44.1 the first-order autoregressive parameter estimate is reported for the observed differenced series for the SPI (f), AOI (i), and basis change (b). In the top panel, these parameter estimates are from observations for the full (synchronized) trading day. These equations include dummy variables to account for institutional market opening effects. In the following tables of results, SING refers to singularities in the estimation process.

For the SPI futures price process, low first-order negative serial correlation is initially detected in the log of the price change. As the sampling interval is reduced, low first-order positive serial correlation can be detected in the series. This feature of the data accords with order splitting and non-trading-induced effects.

Low positive first-order serial correlation can be detected in differences of the log of the market index and serial correlation increases. Positive serial correlation is high at 15-min intervals for the opening excluded set. When sampling the market index at 5-min intervals, substantial positive first-order serial correlation is detected in the log of the price change process. Miller et al. (1994) demonstrate that thin trading and non-trading in individual stocks induce a positive serial correlation in the observed spot index price change process. However, the smoothing effect from non-trading in individual stocks, averaging bid/ask bounce effects in heavily traded stocks, and under-differencing the aggregate process also contribute.

Of interest is the reduction in serial correlation of the basis spread as d → 0. As the log futures price change moves into the order splitting/non-price change region, positive serial correlation is induced in the log of the futures price change. This, in part, helps offset the increasing positive serial correlation induced in the log of the spot index.

44.6.3 Conditional Variance Estimates

Allowing for and excluding market open/closing effects in first-moment equations make little difference to GARCH parameter estimates. This observation holds for alternative specifications of the dynamics for the first-moment equation at 30- and 15-min intervals. However, mis-specifying the first-moment equation dynamics is important in the conditional second-moment equations for the AOI at 5-min intervals.

The GARCH(1,1) estimates for the SPI, AOI, and basis are generated from Eq. 44.6, i.e., an autoregressive specification of the levels AR(L) and Eq. 44.5 the AR(D), respectively (Table 44.2).

Opening dummy variables are incorporated in both first-moment and conditional second-moment equations for the full-day data series. Two dummy variables, corresponding to the first- two 5-min intervals, are included in both first and conditional second-moment equations for the full-day data series. Results in the lower panel are obtained from excluding all trade in the SPI and AOI prior to 10.30 a.m.

At 5-min sampling intervals, a transient effect is introduced into the unconditional distributions which carries through to the conditional variance estimates. Values for α_1 and β_1 differ both within the same data series for alternative first-moment equations and across data series for the same form of first-moment equation. For one extreme case (with exclusion of market opening), the GARCH parameter estimates for the AOI and therefore the basis change do depend on the mis-specification of the first-moment equation.

From the first set of results (full day), it would appear as if the SPI is persistent in variance, the AOI is not and neither is the basis change persistent in variance.

A second extreme case occurs at 30- and 15-min intervals (with exclusion of market opening). In this case GARCH parameter estimates are almost identical for alternative specifications of the dynamics of the first-moment equation. However, for the AOI the sum of the GARCH parameter estimates is near the IGARCH boundary, and the same feature is observed in the basis change.

It would appear that, given anomalies are adequately accounted for, mis-specification of the (weak) dynamic structure of price changes in these processes is only relevant for estimating these GARCH equations for the AOI at 5-min intervals. The important issue is the correct specification of the market opening effects.

Table 44.2 GARCH estimates for the SPI (F_t), AOI (I_t), and basis change (B_t) for AR(L) and AR(D) specifications

	F_t		I_t		B_t	
Full day						
30 min						
α_1	0.0527	0.0517	0.1019	0.1002	0.0904	0.0741
	(7.84)	(7.58)	(7.23)	(6.91)	(9.04)	(7.70)
β_1	0.9266	0.9284	0.0196	0.0227	0.0204	0.0251
	(115)	(111)	(2.09)	(2.24)	(2.06)	(2.31)
15 min						
α_1	0.0602	0.0592	0.1922	0.1514	0.1050	0.1000
	(14.7)	(14.5)	(13.2)	(11.4)	(13.6)	(13.6)
β_1	0.9151	0.9167	0.0893	0.1100	0.0852	0.1186
	(224)	(225)	(9.89)	(10.9)	(5.45)	(7.17)
05 min						
α_1	0.0358	0.0362	0.3540	0.2338	0.1705	0.1685
	(34.0)	(34.6)	(35.8)	(27.9)	(39.6)	(38.5)
β_1	0.9535	0.9530	0.2074	0.2428	0.4345	0.4469
	(1075)	(1087)	(26.8)	(29.3)	(43.3)	(45.6)
Excluding market open						
30 min						
α_1	0.0448	0.0444	0.0668	0.0693	0.0684	0.0631
	(8.31)	(8.12)	(8.71)	(9.22)	(11.5)	(10.5)
β_1	0.9473	0.9478	0.9074	0.9050	0.9114	0.9172
	(155)	(150)	(92.9)	(93.3)	(114)	(109)
15 min						
α_1	0.0385	0.0394	0.0867	0.0835	0.0310	0.0403
	(14.8)	(14.5)	(12.9)	(13.2)	(12.9)	(18.8)
β_1	0.9551	0.9564	0.8832	0.8872	0.9499	0.9537
	(382)	(369)	(113)	(115)	(232)	(374)
05 min						
α_1	5.8E-5	0.0329	0.3326	0.2229	4.6E-5	0.0236
	(33.7)	(30.3)	(37.5)	(33.6)	(27.4)	(35.9)
β_1	1.000	0.9638	0.4470	0.5795	1.000	0.9742
	SING	(906)	(35.2)	(46.4)	SING	(1353)

Asymptotis t-statistics in brackets

These preliminary results have been obtained from observations sampled for all four futures contracts and a continuous series constructed for 1992. The AOI is not affected by contract expiration. In Table 44.3, GARCH(1,1) estimates for the separate SPI futures contracts, 3 months to expiration, and synchronously sampled observations on the AOI are recorded. The full data series corresponding to synchronized trading on the SFE and ASX is employed.

The same form of dummy variable set was imposed in both first and second-moment equations. As well, a post-lunchtime dummy is included to account for the

Table 44.3 GARCH estimates for the SPI and AOI 3 months to expiration

	MAR		JUN		SEP		DEC	
	F_t	I_t	F_t	I_t	F_t	I_t	F_t	I_t
30 min								
α_1	0.042	0.026	0.145	0.015	0.019	0.089	0.060	0.147
	(2.37)	(0.94)	(3.95)	(0.76)	(2.96)	(3.65)	(3.24)	(3.70)
β_1	0.933	0.054	0.116	0.004	0.977	0.021	0.918	0.037
	(35.6)	(2.05)	(0.56)	(0.21)	(118)	(1.12)	(35.0)	(1.30)
15 min								
α_1	0.028	0.186	SING	0.107	0.037	0.183	0.036	0.195
	(3.11)	(5.85)		(4.75)	(4.96)	(6.84)	(6.08)	(5.80)
β_1	0.958	0.108	0.998	0.035	0.956	0.055	0.958	0.102
	(72.6)	(6.65)	(3235)	(2.11)	(120)	(3.75)	(114)	(3.67)
05 min								
α_1	0.051	0.240	SING	0.307	0.174	0.378	0.039	0.409
	(7.86)	(12.4)		(14.8)	(22.2)	(17.0)	(11.0)	(19.3)
β_1	0.881	0.231	0.999	0.118	0.649	0.161	0.956	0.259
	(52.9)	(12.1)	(21170)	(7.16)	(38.5)	(10.6)	(247)	(16.9)

GARCH(1,1) parameter estimates for the 3 months corresponding to expiration of the March, June, September, and December contracts for 1992. Full-day data series are employed with opening and post-lunchtime dummy variables in both first and conditional second-moment equations. An AR(L) specification of the (log) mean equation is employed for these results

break in daily market trade at the SFE during 1992. The log of levels is specified for the first-moment equations.

There is some instability within this set of parameter estimates. However, a similar pattern emerges within the set of futures and the set of market index estimates as was observed for the full-day series for 1992. The futures conditional variance parameter estimates are close to the IGARCH boundary while the index conditional variance estimates are not. This again implies that these processes cannot be co-persistent in variance for these samples and observation intervals.

In order to obtain further insight into the, seemingly, perverse results for the market index, a similar analysis was undertaken on seven of the largest market capitalized stocks which comprised the AOI during 1992.

Relevant market opening and closing dummy variables were included in both first and conditional second-moment equations accordingly. These effects were not systematic in the first-moment equations and are not reported. The stock price processes have not been transformed to natural logarithmic form for these estimations. This is because the weighted levels of the stock prices are employed in construction of the market index.

As the observation interval is reduced for these stock prices:

(i) The autoregressive parameter estimates for the price levels equation converges to a unit root.
(ii) The first-order serial correlation coefficient for the price difference equation moves progressively into the negative region.

If these were the only stocks comprising the construction of the index, then we might expect to see increasing negative serial correlation in the index. But this observation would only apply if these processes were sampled from a continuous process which was generated by a particular form of ARMA specification. Only in the case of data observed from a continuous process could the results on temporal aggregation of ARMA processes be applied.

These results should not be surprising as the combination of infrequent price change and "price bounce" from bid/ask effects starts to dominate the time series. As well, these bid/ask boundaries can shift up and down. When these processes are aggregated, the effects of price bounce can cancel out. As well, the smoothing effect of thinly and zero traded stocks within the observation intervals dampens and offsets individual negative serial correlation observed in these heavily traded stocks (Table 44.4).

Some of these autoregressive parameter estimates are quite high for the AR(D) specifications. However, it is apparent that there is almost no difference in GARCH(1,1) estimates from either the AR(L) or AR(D) specifications at each observation interval for any stock price. In some instances estimation breaks down at 5 min intervals.

If the conditional variances of these stock price movements contain common news and announcement effects, then it should not be surprising that the weighted aggregated process is not persistent in variance. This can happen when news affects all stocks in the same market. As well, smoothing effects from thin-traded stocks help dampen volatility shocks observed in heavily traded stocks. These news and announcement effects may be irrelevant when observing these same processes at daily market open to open or close to close. These news and announcement effects may be due to private information filtering onto the market prior to and following market open. It is during this period that overnight information effects can be observed in both price volatility and volume. As well, day traders and noise traders are setting positions. However, the ad hoc application of dummy variables is not sufficient to capture the interaction between volatility and volume. In the absence of specific measures of these news "variables," the effects cannot be directly incorporated into a structural model. However, these effects are often captured in the price volatility and reflected in increased trading activity.

44.6.4 Weighted GARCH Estimates

Weighted GARCH estimates for the futures (log) price process with the accumulated number of futures contracts traded within the interval t to t−1 are reported in Table 44.5. This choice ensures that volume measures are recorded within the interval that actual prices define.

The weighting variable employed in the index (log) price process is the squared shock from the futures price mean equation within the interval t to t−1. There is no natural "volume" of trade variable available for the market index. The form of mean

Table 44.4 Unconditional mean and GARCH(1,1) estimates: Australian stock prices

	AR(L) AR(D)		AR(L)		AR(D)	
Interval	p_1	ρ_1	α_1	β_1	α_1	β_1
NAB						
60 min	0.999	−0.160	0.125	0.762	0.122	0.774
			(17.2)	(39.2)	(16.4)	(40.1)
30 min	0.999	−0.135	0.086	0.862	0.081	0.876
			(25.3)	(123)	(24.4)	(135)
15 min	0.999	−0.199	0.226	0.627	0.220	0.629
			(22.5)	(70.0)	(22.2)	(61.1)
05 min	1.00	−0.287	2.4E-5	0.042	2.3E-5	0.049
			(0.94)	(0.28)	(0.53)	(0.24)
BHP						
60 min	1.00	−0.034	0.096	0.843	0.098	0.839
			(9.26)	(45.6)	(9.14)	(44.2)
30 min	1.00	−0.087	0.067	0.898	0.066	0.899
			(14.2)	(129)	(13.7)	(128)
15 min	1.00	−0.104	0.054	0.923	0.054	0.924
			(20.5)	(281)	(20.2)	(279)
05 min	1.00	−0.123	0.132	0.795	0.130	0.799
			(66.0)	(361)	(61.9)	(363)
BTR						
60 min	0.995	−0.035	0.194	SING	0.196	SING
			(9.68)	(SING)	(9.68)	(SING)
30 min	0.998	−0.109	0.232	0.294	0.231	0.297
			(14.3)	(7.36)	(13.9)	(7.55)
15 min	0.999	−0.109	0.136	0.615	0.131	0.626
			(17.6)	(31.5)	(17.5)	(32.8)
05 min	1.00	−0.074	0.076	0.831	0.074	0.835
			(45.9)	(231)	(45.5)	(233)
WMC						
60 min	1.00	0.052	0.015	0.981	0.015	0.981
			(6.58)	(368)	(6.54)	(363)
30 min	1.00	0.016	0.272	0.234	0.265	0.251
			(13.7)	(5.39)	(13.70	(5.85)
15 min	1.00	−0.008	0.185	0.594	0.186	0.597
			(19.6)	(33.7)	(19.3)	(34.0)
05 min	1.00	−0.065	0.013	0.003	SING	SING
CRA						
60 min	0.999	0.031	0.332	0.058	0.333	0.074
			(13.0)	(1.52)	(12.5)	(1.88)
30 min	0.999	0.008	0.231	0.456	0.231	0.458
			(10.2)	(5.89)	(9.98)	(5.69)
15 min	1.00	0.003	0.132	0.687	0.132	0.688
			(24.0)	(58.8)	(23.8)	(58.5)

(*continued*)

Table 44.4 (continued)

	AR(L) AR(D)		AR(L)		AR(D)	
Interval	p_1	ρ_1	α_1	β_1	α_1	β_1
05 min	1.00	0.006	0.087	0.840	0.087	0.841
			(84.1)	(433)	(80.1)	(437)
MIM						
60 min	0.998	−0.015	0.016	0.979	0.037	0.934
			(5.22)	(227)	(6.47)	(94.4)
30 min	0.999	−0.090	0.166	0.279	.160	0.281
			(25.30)	(123)	(24.4)	(135)
15 min	1.00	−0.091	0.151	0.612	0.150	0.614
			(23.9)	(40.9)	(23.5)	(40.9)
05 min	1.00	−0.084	0.001	0.841	0.001	0.814
			(2.66)	(9.86)	(2.35)	(9.46)
CML						
60 min	1.00	−0.033	0.088	0.867	0.085	0.873
			(8.88)	(52.1)	(9.02)	(56.2)
30 Min	1.00	−0.045	0.037	0.953	0.039	0.950
			(14.4)	(52.1)	(14.4)	(269)
15 Min	1.00	−0.056	0.196	0.647	0.189	0.658
			(28.2)	(62.4)	(27.3)	(64.6)
05 min	1.00	−0.078	0.095	0.844	0.110	0.846
			(74.6)	(456)	(75.90	(471)

Asymptotic t-statistics in brackets
Column 1 contains the observation interval
Column 2 contains the autoregressive parameter from an AR(L) specification of the price levels equation
Column 3 contains the first-order autoregressive parameter estimate from an AR(D) specification of the price change
Columns 4 and 5 contain the GARCH(1,1) parameter estimates from an AR(L) specification of the price levels equation
Columns 6 and 7 contain the corresponding estimates from an AR(D) specification of the price changes

equation is the same as the generated corresponding results for Table 44.3, i.e., an AR(L). However, the results are almost identical when alternative forms for the mean equation are employed, i.e., AR(L) and AR(D).

The specifications generating the reported stock price estimates are augmented to include a measure of trade activity within the observation interval. These measures are the accumulated number of stocks traded in each individual stock within the interval t to t−1. The estimates are reported in Table 44.6. The conditional variance parameter estimates are almost identical from the AR(L) and AR(D) specifications of the mean equation. The logarithmic transformation has not been taken for these stock prices.

Direct comparison of these GARCH parameter estimates (α_1 and β_1) with those from the unweighted GARCH(1,1) estimates demonstrates the importance of this measure of activity. The change in GARCH parameter estimates is striking.

Table 44.5 Weighted GARCH estimates for the SPI and AOI 3 months to expiration

	March		June		September		December	
	F_t	I_t	F_t	I_t	F_t	I_t	F_t	I_t
30 min								
α_1	0.093	0.040	0.134	0.027	SING	0.079	0.030	0.007
	(2.07)	(1.51)	(4.58)	(1.10)		(2.73)	(1.01)	(0.40)
β_1	0.014	0.039	SING	0.068	SING	0.033	SING	0.025
	(0.24)	(1.62)		(2.74)		(1.91)		(1.24)
15 min								
α_1	0.052	0.185	0.175	0.022	0.039	0.042	0.076	0.058
	(1.84)	(3.97)	(6.67)	(1.50)	(2.47)	(2.35)	(2.89)	(3.04)
β_1	0.004	0.086	0.000	0.163	0.000	0.106	0.001	0.136
	(0.19)	(3.82)	(0.00)	(7.61)	(0.00)	(4.99)	(0.01)	(5.67)
05 min								
α_1	0.000	0.156	0.095	0.170	0.004	0.162	0.030	0.128
	(4.04)	(8.34)	(7.79)	(9.55)	(32.1)	(9.58)	(8.32)	(7.30)
β_1	0.000	0.238	SING	0.232	0.000	0.270	0.000	0.327
	(5.06)	(11.9)		(11.8)	(0.12)	(16.0)	(7.79)	(17.8)

The measures of persistence from these weighted estimates are never near the IGARCH boundary. These effects are summarized in Table 44.7.

These results are generated from an AR(L) specification of the first-moment equation. By adequately accounting for contemporaneous intraday market activity, the time persistence of volatility shocks becomes less relevant. It follows that deviations of the estimated conditional variance from the true (unobservable) conditional variance are reduced (Table 44.8).

44.6.5 Discussion

In this section some empirical evidence is documented on the behavior of unconditional first and conditional second-moment effects for the market index, futures contracts written on the market index, and for heavily traded stock prices. These results are for Australian financial assets sampled on an intraday basis as the observation interval approaches transactions time $d \to 0$.

The specific empirical findings were:

1. The autoregressive parameter estimate from a difference equation for the log of the index futures is initially negative but moves into the positive region. This can be attributed to bid/ask bounce being dominated by order splitting and non-trading effects.
2. The autoregressive parameter estimate from a difference equation for the log of the market index displays increasing positive serial correlation. This can be attributed to non-trading smoothing effects in low capitalized stocks, averaging of bid/ask bounce effects, and under-differencing the aggregated index.

3. The basis change between the log futures price change and the log index level change displayed two surprising effects:
   (i) The autoregressive parameter for the basis change is initially negative, but the strength of this effect weakens. This can be attributed to the log of the

Table 44.6 Weighted GARCH estimates: Australian stock prices

		AR(L)			AR(D)		
Stock	Interval	α_1	β_1	X	α_1	β_1	X
NAB	60 min	0.147	0.028	3.0E-9	0.144	0.027	3.0E-9
		(11.7)	(2.89)	(36.9)	(11.1)	(2.35)	(36.2)
	30 min	0.168	0.143	2.8E-9	0.171	0.139	2.9E-9
		(11.4)	(12.9)	(47.0)	(11.6)	(11.6)	(46.8)
	15 min	0.140	0.135	3.1E-9	0.131	0.142	3.1E-9
		(15.7)	(17.1)	(114)	(15.6)	(16.80	(112)
BHP	60 min	0.125	0.053	9.4E-9	0.124	0.057	9.5E-9
		(8.97)	(2.94)	(27.7)	(8.49)	(3.22)	(28.2)
	30 min	0.108	0.133	8.7E-9	0.095	0.137	8.9E-9
		(10.8)	(11.2)	(32.8)	(9.95)	(12.8)	(34.1)
	15 min	0.132	0.110	9.9E-9	0.121	0.113	10E-9
		(13.9)	(13.9)	(56.6)	(13.2)	(12.6)	(58/8)
BTR	60 min	0.045	0.024	7.8E-9	0.042	0.011	8.0E-9
		(3.06)	(1.29)	(19.0)	(2.85)	(0.64)	(19.2)
	30 min	0.116	0.051	8.0E-10	0.106	0.056	8E-10
		(9.74)	(3.94)	(24.8)	(9.46)	(4.34)	(25.3)
	15 min	0.070	0.023	1.9E-9	0.058	0.026	1.9E-9
		(14.0)	(5.28)	(38.80	(11.60	(5.45)	(38.9)
WMC	60 min	0.110	0.040	1.9E-9	0.110	0.040	1.9E-9
		(5.68)	(2.18)	(25.3)	(5.67)	(2.22)	(25.5)
	30 min	0.107	0.008	1.9E-9	0.108	0.009	1.9E-9
		(8.35)	(0.65)	(30.8)	(8.39)	(0.76)	(30.7)
	15 min	0.078	0.053	1.8E-9	0.074	059	1.8E-9
		(14.6)	(6.47)	(38.0)	(14.2)	(7.21)	(38.1)
CRA	60 min	0.116	0.020	4.3E-8	0.116	0.015	4.3E-8
		(7.04)	(2.06)	(26.8)	(7.07)	(1.69)	(26.7)
	30 min	0.056	0.023	5.4E-8	0.054	0.024	5.5E-8
		(8.30)	(4.95)	(38.4)	(8.11)	(5.14)	(38.4)
	15 min	0.000	0.000	2.9E-7	0.000	0.000	2.9E-7
		(0.0)	(4.04)	(207)	(1.02)	(6.88)	(208)
MIM	60 min	0.074	SING	6.4E-10	0.077	0.000	6.E-10
		(5.36)		(21.7)	(5.24)	(0.00)	(21.6)
	30 min	0.071	0.074	6.1E-10	0.070	0.087	6.E-10
		(5.84)	(4.13)	(24.6)	(6.28)	(4.90)	(25.0)
	15 min	0.040	0.068	8.7E-10	0.030	0.074	9.E-10
		(6.67)	(9.29)	(40.0)	(4.98)	(9.68)	(40.5)

(*continued*)

Table 44.6 (continued)

Stock	Interval	AR(L) α_1	AR(L) β_1	AR(L) X	AR(D) α_1	AR(D) β_1	AR(D) X
CML	60 min	0.125	SING	1.9E-8	0.122	SING	1.9E-8
		(5.92)		(17.8)	(5.81)		(17.8)
	30 min	0.108	0.029	2.5E-8	0.101	0.034	2.6E-8
		(8.65)	(3.44)	(30.8)	(8.27)	(4.16)	(31.1)
	15 min	0.056	0.000	4.0E-8	0.057	SING	4.4E-8
		(16.3)	(0.00)	(56.2)	(17.4)		(59.2)

Asymptotic t-statistics in brackets
Columns 3 and 4 contain the weighted GARCH parameter estimates from an AR(L) specification of the price levels equation with the volume parameter estimate in column 5
Columns 6–8 contain the corresponding estimates from an AR(D) specification of the price changes
Estimates are quite unstable at 5-min intervals

Table 44.7 Measures of persistence from GARCH(1,1) and weighted GARCH equations for the SPI 3 months to expiration

	G	G-W	G	G-W	G	G-W	G	G-W
Interval	March		June		September		December	
30 min	0.975	0.107	0.261	SING	0.996	SING	0.978	SING
15 min	0.986	0.056	SING	0.175	0.993	0.039	0.994	0.077
05 min	0.932	0.000	SING	SING	0.823	0.004	0.995	0.030

futures price change behavior where the autoregressive parameter moved into the positive region and price change became less frequent.

(ii) The unweighted log of the futures price change was close to the IGARCH boundary. For the full-day data series, the log of the market index change was not close to the IGARCH boundary and neither was the basis change. When market opening trade was excluded, the log of the futures price change, the log of the market index price change, and the basis change were close to the IGARCH boundary although this effect dissipated as $d \rightarrow 0$. However, volume of trade on both the SFE and ASX is heaviest within this excluded interval. It follows that any conclusions concerning the co-persistence structure would be misleading from this intraday data.

4. The autoregressive parameter from a levels equation for the stock prices converges to a unit root. This can be attributed to small and zero price change effects.
5. The autoregressive parameter from a difference equation for the stock prices displays increasing tendency towards and into the negative serial correlation region. This can be attributed to price bounce effects where the boundaries are tight and relatively stable.
6. GARCH(1,1) or weighted GARCH conditional variance parameter estimates do not depend on the specification of the dynamics of first-moment equation for 30- and 15-min intervals of these futures, market index, and stock price processes.

Table 44.8 Measures of persistence from GARCH(1,1) and weighted GARCH equations for Australian stock price processes

	NAB		BHP		BTR		WMC	
60 min	0.887	0.175	0.939	0.178	SING	0.069	0.996	0.150
30 min	0.948	0.311	0.965	0.241	0.526	0.167	0.506	0.115
15 min	0.853	0.275	0.977	0.242	0.751	0.093	0.779	0.131
	CRA		MIM		CML			
60 min	0.390	0.136	0.995	SING	0.995	SING		
30 min	0.687	0.079	0.445	0.145	0.990	0.137		
15 min	0.819	0.000	0.763	0.108	0.843	0.056		

An AR(L) specification has been employed to generate these results
The measure of persistence is the calculated as a+b from the conditional variance equations
G represents persistence obtained from a GARCH(1,1) specification
G-W represents persistence obtained from a weighted GARCH specification
SING indicates that one or more of these GARCH parameters could not be evaluated due to singularities

The GARCH(1,1) parameter estimates for the AOI at 5-min intervals were different when alternative forms of first-moment equations were specified. There was no perceivable difference in the other processes at this sampling frequency. It would then appear that increasing positive serial correlation (smoothing) in the observed returns process has a greater distorting effect on GARCH(1,1) parameter estimates than increasing negative serial correlation (oscillation) in observed returns processes. The most important effect is mis-specification of the conditional variance equations from failure to adequately account for the interaction between market activity and conditional variance (volatility).

Some implications of these results are:

7. When aggregating stock prices, which may be close to the IGARCH boundary, persistence in variance can be low for the market index because (i) there is common persistence present in heavily traded stock prices which when aggregated do not display persistence in variance because heavily traded stock prices react to market specific news instantaneously, (ii) the smoothing effect of less heavily traded stocks dampens the volatility clustering which is often observed in other financial assets such as exchange rates, and (iii) high volatility and volume of trade effects within these markets following opening is better measured by employing relevant measures of activity than an ad hoc approach.
8. There is a strong and quantifiable relationship between activity in these markets and volatility.

44.7 Conclusion

The behavior of financial asset price data observed intraday is quite different from these data observed at longer sampling intervals such as day to day. Market anomalies which distort intraday observed data mean that volatility estimates obtained from data observed from day-to-day trades will provide different volatility estimates. This latter feature then depends upon when the data is sampled within the trading day. If these anomalies and intraday trading patterns are accounted for in the estimation process, then better estimates of volatility are obtainable by employing intraday observed data. However, this is dependent on a sampling interval that is not so fine that these estimators break down.

The specific results indicate that serial correlation in returns processes can distort parameter estimates obtainable from GARCH estimators. However, induced excess kurtosis may be a more important factor in distortions to estimates. The most important factor is mis-specification of the conditional variance (GARCH) equation from omission of relevant variables which explain the anomalies and trading patterns observed in intraday data. Measures of activity do help explain systematic shifts in the underlying returns distribution and in this way help explain "jumps" in the volatility process. This effect can be observed in the likelihood function and in asymptotic standard errors of weighting (mixing) variables. One feature is that the measure of volatility persistence observed in unweighted univariate volatility estimators is reduced substantially with inclusion of weighting variables.

Appendix

At the time the ASX data were collected, the exchange had just previously moved from floor to screen trading with the six main capital city exchanges linked via satellite and trade data streamed to trading houses and brokers instantaneously via a signal G feed. The SFE maintained Pit trading for all futures and options on futures contracts at the time.

Legal restrictions on third party use and development of interfaces meant the ASX had a moratorium on such usage and development. The author was required to obtain special permission from the ASX to capture trade data from a live feed from broking house Burdett, Buckeridge, and Young (BBY). There was a further delay in reporting results of research following the legal agreement obtained from the ASX.

Trade data for stock prices and volume of trade were then sampled into 5-min files and subsequently into longer sampling interval files. The market index was refreshed at 1-min intervals and the above sampling scheme repeated. Futures price trades were supplied in two formats via feed: Pit (voice recorded) data and Chit data. Although the Pit data provides an instantaneous record of trade data during the trading day, some trades are lost during frantic periods of activity. The Chit records are of every trade (price, volume, buyer, seller, time stamped to the nearest second, etc.). The recorded Chits are placed in a wire basket on a carriageway and transferred up the catwalk where recorders on computers enter details via a set of simplified keystrokes. The average delay from trade to recording is around 30 s for the Chit trades. These are then fed online to trading houses and brokers. At the end of the trading day, these recorded trades are supplemented with a smaller set of records that were submitted to the catwalk late, e.g., morning trades that may have gone to lunch in a brokers pocket and submitted during the afternoon session and also some late submitted trades.

We created the intraday sampled files from both the Pit and Chit records. However, we employed the Chit trades for analysis in this paper so as to have the correct volume of trade details for each trading interval. All trades were reallocated using the time stamps to the relevant time of trade, including trades not submitted on time but supplied as an appendix to the trading day data. In this study the average number of late Chits were not a high proportion of daily trades. These futures price records were then sampled into relevant 5-min records and longer sampling frames generated in the same manner as was employed for the stock prices.

For all series the first price and last price closest to the opening and closing nodes for each sampling interval were recorded with volume of trade the accumulated volume of trade within the interval defined by first and last trade.

References

Brown, S. (1990). Estimating volatility. In S. Figlewski et al. (Eds.), *Financial options: From theory to practice* (pp. 516–537). Homewood: Business One Irwin.

Duffee, G., Kupiec, P., & White, A. P. (1990). *A primer on program trading and stock price volatility: A survey of the issues and evidence* (Working Paper No. 109). FEDS Board of Governors of the Federal Reserve System.

Edwards, F. R. (1988). Futures trading and cash market volatility: Stock index and interest rate futures. *Journal of Futures Markets, 8*, 421–439.

Gannon, G. L. (1994). Simultaneous volatility effects in index futures. *Review of Futures Markets, 13*, 1027–1066.

Gannon, G. L. (2010). Simultaneous volatility transmissions and spillovers: Theory and evidence. *Review of Pacific Basin Financial Markets and Policies, 13*, 127–156.

Ghose, D., & Kroner, K. F. (1994). *Common persistence in conditional variances: Implications for optimal hedging*. Paper presented at the 1994 Meeting of the Australasian Econometric Society.

McCurdy, T., & Morgan, I. G. (1987). Tests of the martingale hypothesis for foreign currency futures with time varying volatility. *International Journal of Forecasting, 3*, 131–148.

Miller, M. H., Muthuswamy, J., & Whaley, R. E. (1994). Mean reversion of standard and poor's 500 index basis changes: Arbitrage-induced or statistical illusion? *Journal of Finance, 49*, 479–513.

Nelson, D. B. (1990a). ARCH models as diffusion approximations. *Journal of Econometrics, 45*, 7–38.

Nelson, D. B. (1990b). Stationarity and persistence in the GARCH(1,1) model. *Econometric Reviews, 6*, 318–334.

Nelson, D. B., & Foster, D. P. (1991). *Estimating conditional variances with misspecified ARCH models: Asymptotic theory*. Graduate School of Business, University of Chicago, mimeo.

Nelson, D. B., & Foster, D. P. (1994). Asymptotic filtering theory for univariate ARCH models. *Econometrica, 62*, 1–41.

Optimal Asset Allocation Under VaR Criterion: Taiwan Stock Market

45

Ken Hung and Suresh Srivastava

Contents

Abstract

Value at risk (VaR) measures the worst expected loss over a given time horizon under normal market conditions at a specific level of confidence. These days, VaR is the benchmark for measuring, monitoring, and controlling downside financial risk. VaR is determined by the left tail of the cumulative probability distribution of expected returns. Expected probability distribution can be generated assuming normal distribution, historical simulation, or Monte Carlo simulation. Further, a VaR-efficient frontier is constructed, and an asset allocation model subject to a target VaR constraint is examined.

This paper examines the riskiness of the Taiwan stock market by determining the VaR from the expected return distribution generated by historical simulation. Our result indicates the cumulative probability distribution has a fatter left tail, compared with the left tail of a normal distribution. This implies a riskier market.

K. Hung (✉)
Division of International Banking & Finance Studies, Texas A&M International University, Laredo, TX, USA
e-mail: ken.hung@tamiu.edu

S. Srivastava
University of Alaska Anchorage, Anchorage, AK, USA
e-mail: afscs@uaa.alaska.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_45,

We also examined a two-sector asset allocation model subject to a target VaR constraint. The VaR-efficient frontier of the TAIEX traded stocks recommended mostly a corner portfolio.

Keywords

Value at risk • Asset allocation • Cumulative probability distribution • Normal distribution • VaR-efficient frontier • Historical simulation • Expected return distribution • Two-sector asset allocation model • Delta • Gamma • Corner portfolio • TAIEX

45.1 Introduction

Risk is defined as the standard deviation of unexpected outcomes, also known as volatility. Financial market risks are of four types: interest rate risk, exchange rate risk, equity risk, and commodity risk. For a fixed-income portfolio, the linear exposure to the interest rate movement is measured by duration. Second-order exposure is measured by convexity. In the equity market, linear exposure to market movement is measured by the systematic risk or beta coefficient. In the derivative markets, the first-order sensitivity to the value of underlying asset is measured by *delta*, and second-order exposure is measured by *gamma*. Innovations in the financial markets have introduced complicated portfolio choices. Hence, it is becoming more difficult for managers to get useful and practical tools of market risk measurement. The simple linear considerations such as Basis Point Value, or first- or second-order volatility, are inappropriate. They can't accurately reflect risk at the time of dramatic price fluctuation.

VaR (value at risk) has become a popular benchmark for the downside risk measurement.[1] VaR converts the risks of different financial products into one common standard: potential loss, so it can estimate market risk for various kinds of investment portfolio. VaR is used to estimate the market risk of financial assets. Special concern of the market risk is the downside risk of portfolio values resulting from the fluctuation of interests, exchange rates, stock prices, or commodity prices. VaR is consistent in estimating the financial risk estimation. It indicates risk of dollar loss of portfolio value. Now the risks exposure of different investment portfolios (such as equity and fixed income) or different financial products (such as interest rate swaps and common stock) have a common basis for direct comparison. For decision makers, VaR is not only a statistical summary; it can also be used as a management and risk control tool to decide capital adequacy, asset allocation, synergy-based salary policy, and so on.

[1]Extensive discussion of value at risk can be found in Basak and Shapiro (2001), Beder (1995), Dowd (1998), Fong and Vasicek (1997), Hendricks (1996), Hoppe (1999), Jorion (1997, 1997), Schachter (1998), Smithson and Minton (1996a, b), and Talmor (1996).

VaR is a concept widely accepted by dealers, investors, and legislative authorities. J.P. Morgan has advocated VaR and incorporated it in RiskMetrics (Morgan 1996). RiskMetrics contain most of the data and formulas used to estimate daily VaR, including daily updated fluctuation estimations for hundreds of bonds, securities, currencies, commodities, and financial derivatives. Regulatory authorities and central bankers from various countries at Basel Committee meetings agreed to use VaR as the risk-monitoring tool for the management of capital adequacy. VaR has also been widely accepted and employed by securities corporations, investment banks, commercial banks, retirement funds, and nonfinancial institutions. Risk managers have employed VaR in ex-post evaluation, that is, to estimate and justify the current market risk exposure.[2] Confidence-based risk measure was first proposed by Roy (1952).

The inclusion of VaR into the asset allocation model means the inclusion of downside risk into model constraints. Within the feasible scope of investment portfolio that meets shortfall constraints, the optimal investment portfolio is decided by maximum expected return. The definition of shortfall constraint is that the probability of investment portfolio value dropping to a certain level is set as the specific disaster probability. The asset allocation framework that takes VaR as one of its constraints has increased the importance of VaR and has employed VaR as an ex-ante control tool of market risk.

45.2 Value at Risk

One difficulty in estimating VaR is the choice of various VaR methods and corresponding hypotheses. There are three major methods to estimate VaR: variance-covariance analysis, historical simulation, and Monte Carlo simulation. Variance-covariance analysis assumes that market returns for financial products are normally distributed, and VaR can be determined from market return's variance and covariance. The normal distribution hypothesis of variance-covariance analysis makes it easy to estimate VaR at different reliability and different holding period (see Appendix 1). Its major disadvantage is that the return in the financial market is usually not in normal distribution and has fat tails. This means the probability of extreme loss is more frequent than estimated by variance-covariance analysis.

Historical simulation assumes the future market return of the investment portfolio is identical to the past returns; hence, the attributes of current market can be used to simulate the future market return (Hendricks 1996; Hull and White 1998). Historical simulation approach does not suffer from the tail-bias problem, for it does not assume normal distribution. It relies on actual market return distribution, and the estimation reflects what happened during the past sample period. It has another advantage over variance-covariance analysis: it can be used for nonlinear products, such as commodity derivatives. However, the problem with historical

[2]Institutional use of VaR can be found in Basel (1995, 1998a, b, c, 1999), Danielsson et al. (1998), and Danielsson Hartmann and de Vries (1998).

simulation is its sensitivity to sample data. Many scholars pointed out that if October 1987 is included into the observation period, then it would make great difference to the estimation of VaR. Another problem with historical simulation is that the left tail of actual return distribution is at zero stock prices. In other words, it would not be accurate to assume a zero probability of loss that is greater than the past loss. Lastly, the estimation of historical simulation is more complicated than that of variance-covariance analysis. The VaR needs to be reestimated every time the level of reliability or holding period changes.

Monte Carlo simulation can be used to generate future return distribution for a wide range of financial products. It is done in two steps. First, a stochastic process is specified for each financial variable along with appropriate parameters. Second, simulated prices are determined for each variable, and portfolio loss is calculated. This process is repeated 1,000 times to produce a probability distribution of losses. Monte Carlo simulation is the most powerful tool for generating the entire probability distribution function and can be used to calculate VaR for a wide range of financial products. However, it is time consuming and expensive to implement.

Modern investment portfolio theories try to achieve optimal asset allocation via maximizing the risk premium per unit risk, also known as the Sharpe ratio (Elton and Gruber 1995). Within the framework of mean-variance, market risk is defined as the expected probable variance of investment portfolio. To estimate risk with standard deviation implies investors pay the same attention to the probabilities of negative and positive returns. Yet investors have different aversion to investment's downside risk than to capital appreciation. Some investors may use semi-variance to estimate the downside risk of investment. However, semi-variance has not become popular.

Campbell et al. (2001) have developed an asset allocation model that takes VaR as one of its constraints. This model takes the maximum expected loss preset by risk managers (VaR) as a constraint to maximize expected return. In other words, the optimal investment portfolio deduced from this model meets the constraint of VaR. This model is similar to the mean-variance model that generates the Sharpe index. If the expected return is a normal distribution, then this model is identical with mean-variance model. Details of this model are presented in the Appendix.

Other researchers have examined four models to introduce VaR for ex-ante asset allocation of optimal investment portfolio: mean-variance (MV) model, mini-max (MM) model, scenario-based stochastic programming (SP) model, and a model that combines stochastic programming and aggregation/convergence (SP-A). The investment portfolio constructed using the SP-A model has a higher return in all empirical and simulation tests. Robustness test indicates that VaR strategy results in higher risk tolerance than risk assessment that takes severe loss into consideration. Basak and Shapiro (2001) pointed out that the drawback of risk management lies in its focus on loss probability instead of loss severity. Although loss probability is a constant, when severe loss occurs, it has greater negative consequence than non-VaR risk management.

Lucas and Klaassen (1998) pointed out the importance of correctly assessing the fat-tailed nature of return distribution. If the real return is not in normal distribution, the asset allocation under the hypothesis of normal distribution will result in non-efficiency or non-feasibility. An excellent discussion of VaR and risk measurements is presented by Jorion (2001).

45.3 Empirical Results

In July 1997, financial crisis broke out in Southeast Asian nations and then proliferated to other Asian regions and led to a series of economic problems. Taiwan had also been attacked by financial crisis in late 1998. Domestic financial markets fluctuated. Corporations and individuals greatly suffered. The proliferation of financial crisis within or among countries makes it impossible for corporations and individuals to ignore market risk. Risk measurement is the first thing to do before investment.

The Taiwan market general weighted stock index, individual weighted stock price indexes, and interbank short loan interest rate used in this research paper are obtained from the data base of AREMOS. We divide the historical period into two groups, from 1980 to 1999 and from 1991 to 1999, so as to analyze the impact of violent stock market fluctuation on VaR estimation, such as the New York stock market collapse in October 1987 and Taiwan stock market dramatic uprising from 1988 to 1990. The first period is rather long and can indicate the nature of dramatic stock fluctuation. The second period is rather short and can reflect the change of stock market tendency.

This paper employs historical simulation to reproduce the daily fluctuations of returns for electrical machinery, cement, food, pulp and paper, plastics and petroleum, and textile and fiber stocks trading in the Taiwan stock market during the periods of 1980–1999 and 1991–1999. We estimate their VaRs under reliability levels of 95 %, 97.5 %, and 99 %. The expected return of investment portfolio in 1999 is the sum of annual mean returns of various stocks multiplied with their respective weights. We use this expected return to estimate year 1999 optimal stock holding proportion and analyze the impact of different historical simulation periods on optimal asset allocation.

Table 45.1 presents the summary of TSE general weighted stock index (daily data) and estimated VaR for periods 1980–1999 and 1991–1999. Table 45.2 presents cumulative probability distributions of TSE daily index return for the 1980–1999 period and daily returns under the assumption of normality. It shows that at confidence level lower than 95.8 % (e.g., 90 %), the left-tail probability for

Table 45.1 Summary of Taiwan Stock Exchange (TSE) daily index

Period	Mean return	Standard deviation	Kurtosis	VaR*
1980–1999	0.06 %	1.67 %	2.735	0.0263
1991–1999	0.04 %	1.60 %	2.464	0.0249

Annualized return for the two periods is 23.64 % and 11.87 %, respectively. VaR* is the maximum expected return loss for 1-day holding period at a reliability level of 95 %

Table 45.2 Cumulative probability distribution of TSE daily index

Taiwan index Return (%)	Historical data Cumulative probability	Normal distribution Cumulative probability
−7	0	1.20946E-05
−6	0.004723	0.000144816
−5	0.01102	0.001236891
−4	0.021165	0.007577989
−3	0.038482	0.033571597
−2.9	0.040581	0.0382868
−2.8	0.042855	0.043528507
−2.7	0.045828	0.049334718
−2.5	0.054749	0.062791449
−2.3	0.066993	0.078949736
−2	0.082736	0.108825855
−1	0.185062	0.262753248
0	0.477873	0.485257048
1	0.775407	0.712585369
2	0.905895	0.876745371
3	0.960994	0.960522967
4	0.982158	0.990731272
5	0.992304	0.998424485
6	0.995977	0.999807736
7	1	0.999983254

Period: 1980–1999

historical distribution is higher than the normal return probability (4.2 %). Hence, under the normal distribution assumption, the VaR is overestimated, and this leads to an overcautious investment decision. At confidence level higher than 95.8 % (e.g., 97.5 %), the left-tail probability for historical distribution is lower than the normal return probability (4.2 %). Hence, under the normal distribution assumption, the VaR is underestimated, and this leads to an overactive investment decision.

Figure 45.1 is the graphical presentation of the data in Table 45.2. The solid blue line represents cumulative probability distributions of TSE daily index return, and the dashed red line represents the normal distribution. The bottom panel is the enlarged view of the left tail. This graph also indicates that VaR estimated using extreme values of historical distribution will lead to an overactive investment decision.

Table 45.3 reports annualized returns and standard deviations for TSE daily index and six selected industries: cement, electrical machinery, food, pulp and paper, plastics and petroleum, and textile and fiber. For the 1980–1999 period, the food industry had the greatest risk with a standard deviation of 53.47 % and 15.13 % annual return, whereas the overall market had a standard deviation of 50.80 % with 23.22 % annual return. Textile and fiber was the least risky industry with a standard deviation of 41.21 % and 12.78 % annual return. For the 1991–1999 period, the electrical machinery industry had the greatest risk with a standard

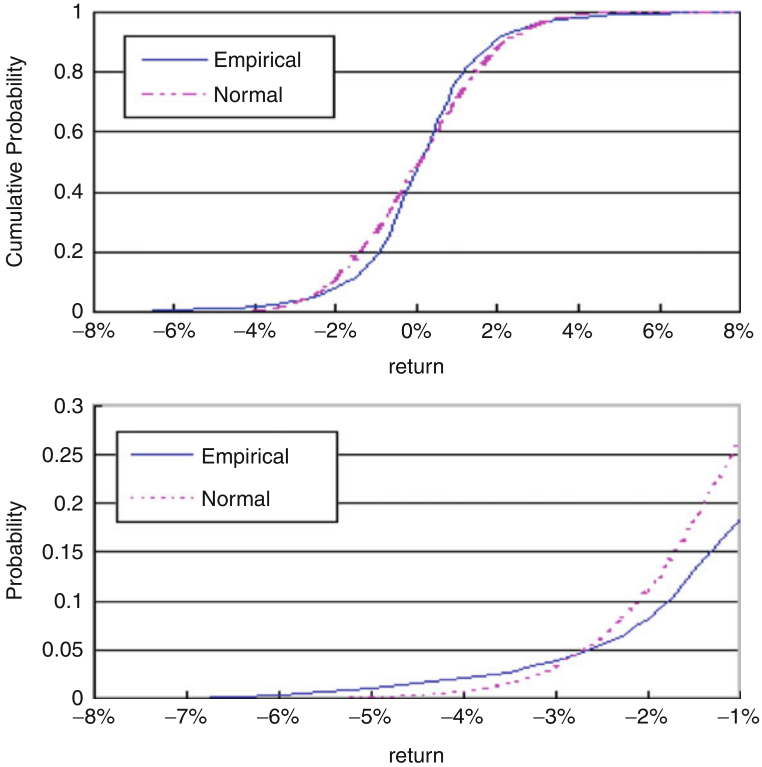


Fig. 45.1 Cumulative probability distribution of Taiwan daily stock index. Period: 1980–1999. *Left* tail of the cumulative probability of Taiwan weighted daily stock index. Period 1980–1999. *Lower* panel shows a fat *left* tail

Table 45.3 TSE index and selected industry's returns and standard deviations

Period	1980–1998		1991–1998	
Industry	Annual return (%)	Standard deviation (%)	Annual return (%)	Standard deviation (%)
TSE index	23.22	50.80	9.40	36.58
Cement	12.48	48.37	0.45	24.92
Electrical machinery	19.49	46.76	26.12	42.26
Food	15.13	53.47	9.55	32.29
Pulp and paper	8.50	45.78	4.64	39.41
Plastics and petroleum	13.32	43.99	11.39	36.26
Textile and fiber	12.78	41.21	7.88	36.62

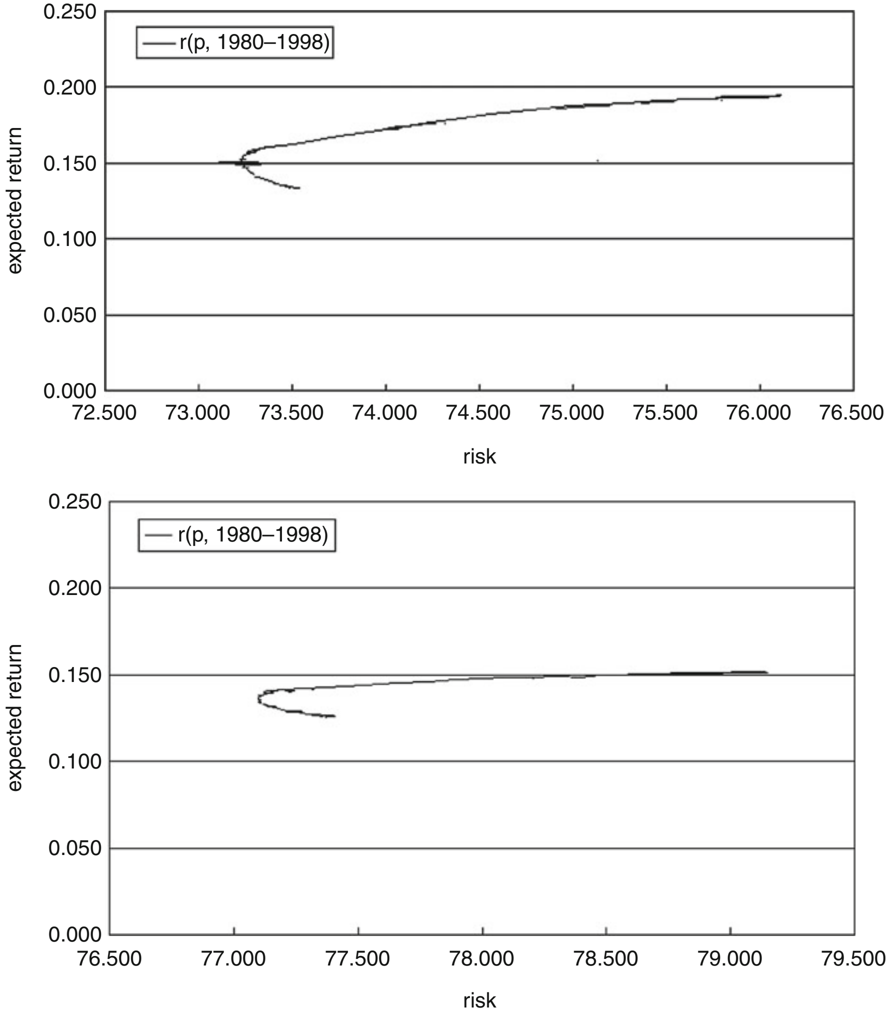


Fig. 45.2 VaR-efficient frontier. The *upper* figure refers to investment portfolio of electrical machinery and plastics and petroleum stocks, and the *lower* figure refers to investment portfolio of cement and food stocks. VaR is set at reliability level of 95 %. Expected return and VaR are estimated from TSE industry indexes from 1980 to 1998

deviation of 42.267 % and 26.12 % annual return, whereas the overall market had a standard deviation of 36.58 % with 9.40 % annual return. The cement industry was the least risky industry with a standard deviation of 24.92 % and 0.45 % annual return. Next we constructed a two-industry optimal portfolio subject to VaR constraint. The optimal asset allocation for the two-industry portfolio is obtained by maximizing S(p) (derivation discussed in Appendix). The resulting VaR-efficient frontiers are plotted in Fig. 45.2. The upper panel in Fig. 45.2 refers

Table 45.4 Optimal asset allocation for the two-industry portfolio obtained by maximizing *S*(*p*) at different level of confidence

	Confidence level		
Portfolio choices	95 %	97.5 %	99 %
Electrical machinery cement	{1,0}; {1,0}[a]	{1,0}; {1,0}	{1,0}; {1,0}
Electrical machinery food	{1,0}; {1,0}	{1,0}; {1,0}	{1,0}; {1,0}
Electrical machinery pulp and paper	{1,0}; {1,0}	{1,0}; {1,0}	{1,0}; {1,0}
Electrical machinery plastics and petroleum	{1,0}; {1,0}	{1,0}; {1,0}	{1,0}; {1,0}
Electrical machinery textile and fiber	{1,0}; {1,0}	{1,0}; {1,0}	{1,0}; {1,0}
Cement food	{0,1}; {0,1}	{0,1}; {0,1}	{0,1}; {0,1}
Cement pulp and paper	{0,1}; {0,1}	{0,1}; {0,1}	{0,1}; {0,1}
Cement plastics and petroleum	{0.12, 0.88}; {0,1}	{0.04, 0.96}; {0,1}	{0.17, 0.83}; {0,1}
Cement textile and fiber	{0.67, 0.33}; {0,1}	{0.6, 0.4}; {0,1}	{0.15, 0.85}; {0,1}
Food pulp and paper	{1,0}; {1,0}	{1,0}; {1,0}	{0.98, 0.02}; {0,1}
Food plastics and petroleum	{1,0}; {1,0}	{1,0}; {1,0}	{1,0}; {1,0}
Food textile and fiber	{1,0}; {1,0}	{1,0}; {1,0}	{1,0}; {1,0}
Pulp and paper plastics and petroleum	{0,1}; {0,1}	{0,1}; {0,1}	{0,1}; {0,1}
Pulp and paper textile and fiber	{0,1}; {0,1}	{0,1}; {0,1}	{0,1}; {0,1}
Plastics and petroleum textile and fiber	{0.99, 0.01}; {1,0}	{0.89, 0.11}; {1,0}	{1,0}; {1,0}

First set {x, y} refers to the historical simulation for period 1980–1998, and second set {x, y} refers to the historical simulation for period 1991–1998

[a]{1, 0} represents 100 % investment in electrical machinery industry and 0 % investment in cement industry

to VaR-efficient portfolios of electrical machinery and plastics and petroleum stocks, and the lower panel in Fig. 45.2 refers to VaR-efficient portfolios of cement and food stocks. VaR is set at the 95 % confidence level. Table 45.4 presents portfolio weights for different combinations of industry stocks at different levels of confidence. Asset allocations for most of the industry combinations represent corner solutions, i.e., 100 % investment in one industry. For example, when electrical machinery stocks are combined with stocks from any other industry, the optimal portfolio is 100 % investment in electrical machinery stocks, for both the time periods and the three confidence levels. Asset allocation for cement stocks is dominated by the other five industry stocks. Allocations of food stocks dominate all other stock weights except electrical machinery. The general nature of asset allocation is the same for both time periods and the confidence levels.

Suppose an investor selects the VaR constraint for maximizing S(p) at a specified level of confidence (say 95 %) and the actual VaR(c, p*) is at a higher level (97.5 %); then VaR(portfolio) will be greater that target VaR*. In this case investors will have to invest a portion of the fund in T-bills ($B > 0$, defined in Appendix). This will make investment VaR(c, p*) equal to the VaR* in the preset

Table 45.5 Optimal allocation for two-industry portfolio (historical simulation, period 1980–1998)

Confidence level (%)	Cement (%)	Food (%)	Portfolio VaR (c, p*)	Lending, B yuan	VaR*	Cement (%)	Food (%)	Cash (%)
95	100	0	31.12	0	31.12	100	0	0
97.5	100	0	42.57	121	31.12	87.9	0	12.1
99	100	0	53.92	216	31.12	78.4	0	21.6
95	0	100	27.49	0	27.49	0	100	0
97.5	0	100	40.22	139	27.49	0	86.1	13.9
99	0	100	56.27	267	27.49	0	73.3	26.7
95	12	88	27.52	0	27.52	12	88	0
97.5	4	96	41.99	154	27.52	3.4	81.2	15.4
99	17	83	53.42	246	27.52	12.8	62.6	24.6
Confidence level (%)	Pulp and paper (%)	Textile and fiber (%)	Portfolio VaR(c, p*)	Lending, B yuan	VaR*	Pulp and paper (%)	Textile and fiber (%)	Cash (%)
95	0	100	29.58	0	29.58	0	100	0
97.5	0	100	41.97	132	29.58	0	86.8	13.2
99	0	100	53.85	230	29.58	0	77	23
Confidence level (%)	Cement (%)	Textile and fiber (%)	Portfolio VaR(c, p*)	Lending, B yuan	VaR*	Cement (%)	Textile and fiber (%)	Cash (%)
95	67	33	25.86	0	25.86	12	88	0
97.5	60	40	41.97	172	25.86	49.7	33.1	17.2
99	15	85	52.51	256	25.86	11.2	63.2	25.6
Confidence level (%)	Plastics and petroleum (%)	Textile and fiber (%)	Portfolio VaR (c, p*)	Lending, B yuan	VaR*	Plastics and petroleum (%)	Textile and fiber (%)	Cash (%)
95	99	1	28.43	0	28.43	99	1	0
97.5	89	11	41.35	139	28.43	76.6	9.5	13.9
99	100	0	55.50	253	28.43	74.7	0	25.3
Confidence level (%)	Food (%)	Textile and fiber (%)	Portfolio VaR(c, p*)	Lending, B yuan	VaR*	Food (%)	Textile and fiber (%)	Cash (%)
95	100	0	27.49	0	27.49	100	0	0
97.5	100	0	40.22	121	27.49	87.9	0	12.1
99	100	0	56.27	267	27.49	73.3	0	26.7

Two-industry portfolio with initial investment of 1,000 yuan
Allocations for other industry combinations are available to interested readers

constraint. In opposite case, the VaR* constraint is specified at a higher level than the portfolio VaR(c, p*); then investors will borrow money to invest in risky assets ($B < 0$). Tables 45.5 and 45.6 list examples of investment in two-industry stocks and T-bills for periods 1980–1999 and 1991–1999, respectively. In each case, the target VaR* in the preset constraint is at a 95 % level of confidence, and 1,000 yuan is invested in the portfolio. In the first panel of Table 45.5, 1,000 yuan is invested in

Table 45.6 Optimal allocation for two-industry portfolio (historical simulation, period 1991–1998)

Confidence level (%)	Cement (%)	Food (%)	Portfolio VaR (c, p*)	Lending, B yuan	VaR*	Cement (%)	Food (%)	Cash (%)
95	100	0	30.55	0	30.55	100	0	0
97.5	100	0	42.57	128	30.55	87.2	0	12.8
99	100	0	53.92	221	30.55	77.9	0	22.1
95	0	100	24.57	0	24.57	0	100	0
97.5	0	100	34.68	117	24.57	0	88.3	11.7
99	0	100	48.41	238	24.57	0	76.2	23.8
95	0	100	27.19	0	27.19	0	100	0
97.5	0	100	35.95	100	27.19	0	90	10
99	0	100	48.59	213	27.19	0	78.7	21.3

Confidence level (%)	Pulp and paper (%)	Textile and fiber (%)	Portfolio VaR(c, p*)	Lending, B yuan	VaR*	Pulp and paper (%)	Textile and fiber (%)	Cash (%)
95	0	100	27.75	0	27.75	0	100	0 %
97.5	0	100	38.98	124	27.75	0	87.6	12.4
99	0	100	49.06	212	27.75	0	78.8	21.2

Confidence level (%)	Cement (%)	Textile and fiber (%)	Portfolio VaR(c, p*)	Lending, B yuan	VaR*	Cement (%)	Textile and fiber (%)	Cash (%)
95	0	100	27.75	0	27.75	0	100	0
97.5	0	100	38.98	124	27.75	0	87.6	12.4
99	0	100	49.06	212	27.75	0	78.8	21.2

Confidence level (%)	Plastics and petroleum (%)	Textile and fiber (%)	Portfolio VaR(c, p*)	Lending, B yuan	VaR*	Plastics and petroleum (%)	Textile and fiber (%)	Cash (%)
95	100	0	27.19	0	27.19	0	100	0
97.5	100	0	35.95	100	27.19	0	90	10
99	100	0	48.59	213	27.19	0	78.7	21.3

Confidence level (%)	Food (%)	Textile and fiber (%)	Portfolio VaR(c, p*)	Lending, B yuan	VaR*	Food (%)	Textile and fiber (%)	Cash (%)
95	100	0	24.57	0	24.57	100	0	0
97.5	100	0	34.68	117	24.57	88.3	0	11.7
99	100	0	48.41	238	24.57	76.2	0	23.8

Two-industry portfolio with initial investment of 1,000 yuan
Allocations for other industry combinations are available to interested readers

electrical machinery stocks and 0 in cement stocks. These allocations are from Table 45.4. Portfolio VaR(c, p*) is 31.12, 42.57, and 53.92 yuan at 95 %, 97.5 %, and 99 % respectively. This leads to the lending of 0, 121, and 216 yuan to meet the target VaR of 31.12. Thus, a 97.5 % portfolio consists of 87.9 % in electrical machinery stocks, 0 in cement stocks, and 21.1 % in T-bills.

45.4 Conclusion

TSE daily index was found to be riskier than the market risk under the assumption of normal distribution for market returns. This resulted in the left tail of cumulative return distribution being fatter and a higher value at risk, indicating an overactive investment activity. For asset allocation model under a constrained VaR framework, most of the optimal portfolios have predominant investment in stocks from one industry. Hence, it will be inappropriate to comment on the optimal allocation of future investment portfolio based on the past stock performance of this unique period studied.

Appendix 1: Value at Risk

Let W_0 be the initial investment and R be the rate of return of a portfolio. The value of the portfolio at the end of the target horizon will be $W = W_0(1 + R)$. Let μ and σ be the expected return and standard deviation of R. The lowest portfolio value at the confidence level c is defined as $W^* = W_0\ (1 + R^*)$. The *relative VaR* is the dollar loss relative to the mean:

$$VaR(\text{mean}) = E(W) - W^* = -W_0(R^* - \mu) \tag{45.1}$$

The *absolute VaR* is the dollar loss relative to zero:

$$VaR(\text{zero}) = W_0 - W^* = -W_0R^* \tag{45.2}$$

W^* and R^* are minimum value and cutoff return, respectively. In this paper we are discussing *absolute VaR*. The general form of *VaR* can be derived from the probability distribution of the future portfolio value f(w). For a given confidence level c, the worst possible portfolio value W^* is such that probability of exceeding W^* is c:

$$c = \int_{\mathrm{W}*}^{\infty} f(w)dw \tag{45.3}$$

The probability of a value lower than W^*, $\mathrm{p} = \mathrm{P}(\mathrm{w} \leq \mathrm{W}^*)$ is 1–c:

$$\mathrm{p} = \mathrm{P}(\mathrm{w} \leq \mathrm{W}^*) = 1 - \mathrm{c} = \int_{-\infty}^{\mathrm{W}*} f(w)dw \tag{45.4}$$

Typical confidence level c is 95 %. This computation of VaR does not require estimation of variance-covariance matrix.

When portfolio returns are normally distributed, then distribution f(w) can be translated into a standard normal distribution Φ(ε), where ε has mean zero and standard deviation of one. VaR can be determined from the tables of the cumulative standard normal distribution:

$$\mathrm{VaR} = \mathrm{N}(1-\mathrm{c}) = \int_{-\infty}^{1-c} \varphi(\varepsilon)d\varepsilon \tag{45.5}$$

and the cutoff return $R^* = -z\sigma + \mu$. This Appendix is based on Jorion (2001). Details of normal distribution can be found in Johnson and Wichern (2007).

Appendix 2: Optimal Portfolio Under a VaR Constraint

We present an asset allocation model under a value-at-risk constraint. This model sets maximum expected loss not to exceed the VaR for a selected investment horizon, T at a given confidence level. Then asset proportions are allocated across the portfolio such that the wealth at the end of investment horizon is maximized. Suppose W_0 is the investor's initial wealth and B is the amount that the investor can borrow ($B > 0$) or lend ($B < 0$) at the risk-free interest rate r_f. Let n be the number of risky assets, γ_i be the fraction invested in risky asset i, and P(i, t) be the price of asset I at time t. Then the initial value of the portfolio

$$W_0 + B = \sum_{i=1}^{n} \gamma_i P(i, 0) \tag{45.6}$$

represents the budget constraint.

Let VaR* be the target VaR consistent with investor's risk aversion and W_T be the wealth at the end of the holding period, T. The downside risk constraint can be written as

$$\Pr\{(W_0 - W_T) \geq \mathrm{VaR}^*\} \leq (1-\mathrm{c}) \tag{45.7}$$

where Pr {.} denotes the expected probability conditioned on information available at time, $t = 0$, and c is the confidence level. Equation 45.2 can be written as

$$\Pr\{W_T \geq (W_0 - \mathrm{VaR}^*)\} \leq (1-\mathrm{c}) \tag{45.8}$$

Let r_p be the total portfolio return at the end of the holding period and T then the expected wealth at the end of holding period; T can be written as

$$\mathrm{E}(W_T) = (W_0 + B)\left(1 + r_p\right) - B(1 + r_f) \tag{45.9}$$

Investor's constrained wealth maximizing objective can be written as

$$\begin{aligned} &\text{Max.}\ \ E(W_T) \\ &\text{s.t.}\quad \Pr\{W_T \leq (W_0 - VaR^*)\} \leq (1-c) \end{aligned} \tag{45.10}$$

Performance measure $S(p)$ and borrowed amount can be deduced from Eq. 45.5. Let p^* be the maximizing portfolio and $q(\mathrm{c}, p)$ defines the quantile that corresponds to probability (1 − c) which can be obtained from portfolio return's cumulative density function. Maximizing portfolio p^* is defined as

$$p^* : \max_p \quad S(p) = \frac{r_p - r_f}{W_0 r_f - W_0 q(c, p)} \tag{45.11}$$

Initial wealth in the denominator of Eq. 45.6 is a scale constant and does not affect the asset allocation. Let VaR(c, p) denote portfolio p's VaR, and then the denominator of Eq. 45.6 can be written as

$$\Phi(c, p) = \mathrm{W_0 r_f} - \mathrm{VaR}(c, p) \tag{45.12}$$

If we consider $\mathrm{r_f}$ as the benchmark return, then $\Phi(c, p)$ represents potential for portfolio losses at the confidence level c. Performance measure S(p) represents Sharpe-like reward-risk ratio, and optimization problem becomes

Optimal portfolio:

$$p^* : \max_p \quad S(p) = \frac{r_p - r_f}{\Phi(c, p)}$$

Optimal portfolio allocation is independent of the initial wealth. It is also independent of the target VaR*. Risk measure Φ(c, p*) depends on VaR(c, p*) and not on VaR*. Investors first allocates the wealth among risky assets and then decides borrowing or lending depending on the value of {VaR* − VaR(c, p*)}. Borrowed amount B can be written as

$$B = \frac{W_0(VaR^* - VaR(c, p^*))}{\Phi'(c, p')}$$

If {VaR* − VaR(c, p*)} is positive, then there is an opportunity to increase the portfolio return by borrowing at the risk-free rate and invest it in risky asset. If {VaR* − VaR(c, p*)} is negative, then the portfolio risk needs to be reduced by investing a portion of the initial wealth in the risk-free asset. In either case the relative proration of funds invested in risky assets remains the same. Since VaR(c, p*) depends on the choice of holding period, confidence level, VaR estimation technique, and the assumption regarding the expected return distribution, the borrowing ($B > 0$) or lending ($B < 0$) will also change. This Appendix is based on Campbell et al. (2001).

References

Basak, S., & Shapiro, A. (2001). Value-at-risk based risk management: Optimal policies and asset prices. *The Review of Financial Studies, 14*(2), 371–405.

Beder, T. S. (1995). VAR: Seductive but dangerous. *Financial Analysts Journal, 51*(5), 12–24.

Campbell, R., Huisman, R., & Koedijk, K. (2001). Optimal portfolio selection in a value-at-risk framework. *Journal of Banking and Finance, 25*, 1789–1804.

Dowd, K. (1998). *Beyond value at risk*: *The new science of risk management*. New York: Wiley.

Elton, E. J., & Gruber, M. J. (1995). *Modern portfolio theory and investment analysis* (5th ed.). New York: Wiley.
Fong, G., & Vasicek, O. A. (1997). A multidimensional framework for risk analysis. *Financial Analysts Journal, 53*, 51–57.
Hendricks, D. (1996). Evaluation of value-at-risk models using historical data. *FRBNY Economic Policy Review, 2*, 39–69.
Hoppe, R. (1999). It's time we buried value-at-risk. *Risk Professional*, 1. October page 18. London.
Hull, J., & White, A. (1998). Incorporating volatility updating into the historical simulation method for value-at-risk. *Journal of Risk, 1*, 5–19.
Johnson, R., & Wichern, D. (2007). *Applied multivariate statistical analysis* (6th ed.). New York: Pearson Education.
Jorion, P. (1997). In defense of VAR. *Derivatives Strategy, 2*, 20–23.
Jorion, P. (2001). *Value at risk* (2nd ed.). New York: McGraw Hill.
Lucas, A., & Klaassen, P. (1998). Extreme returns, downside risk, and optimal asset allocation. *Journal of Portfolio Management, 25*(1), 71–79.
Morgan, J. P. (1996). *RiskMetrics^{TM} – technical document*. New York: Morgan Guaranty Trust Company.
Roy, A. D. (1952). Safety first and the holding of assets. *Econometrica, 20*, 431–449.
Schachter, B. (1998). An irreverent guide to value at risk. *Risks and Rewards*, (1), 17–18.
Smithson, C., & Minton, L. (1996a). Value-at-risk (1): The debate on the use of VaR. *Risk, 9*, 25–27.
Smithson, C., & Minton, L. (1996b). Value-at-risk (2): The debate on the use of VaR. *Risk, 9*, 38–39.
Talmor, S. (1996). History repeat itself. *The Banker*, (1), 75–76.

Alternative Methods for Estimating Firm's Growth Rate

46

Ivan E. Brick, Hong-Yi Chen, and Cheng-Few Lee

Contents

Abstract

The most common valuation model is the dividend growth model. The growth rate is found by taking the product of the retention rate and the return on equity. What is less well understood are the basic assumptions of this model. In this paper, we demonstrate that the model makes strong assumptions regarding the financing mix of the firm. In addition, we discuss several methods suggested in the literature on estimating growth rates and analyze whether these approaches are consistent with the use of using a constant discount rate to evaluate the firm's assets and equity.

This chapter is a slightly revised version of Chapter 64 of *Encyclopedia of Finance*, 2nd Edition and Brick et al. (2014).

I.E. Brick (✉)
Department of Finance and Economics, Rutgers, The State University of New Jersey, Newark/New Brunswick, NJ, USA
e-mail: ibrick@andromeda.rutgers.edu

H.-Y. Chen
Department of Finance, National Central University, Taoyuan, Taiwan
e-mail: fnhchen@ncu.edu.tw

C.-F. Lee
Department of Finance and Economics, Rutgers Business School, Rutgers, The State University of New Jersey, Piscataway, NJ, USA

Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
e-mail: lee@business.rutgers.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_46,

The literature has also suggested estimating growth rate by using the average percentage change method, compound-sum method, and/or regression methods. We demonstrate that the average percentage change is very sensitive to extreme observations. Moreover, on average, the regression method yields similar but somewhat smaller estimates of the growth rate compared to the compound-sum method. We also discussed the inferred method suggested by Gordon and Gordon (1997) to estimate the growth rate. Advantages, disadvantages, and the interrelationship among these estimation methods are also discussed in detail.

Keywords
Growth rate • Discount cash flow model • Internal growth rate • Sustainable growth rate • Compound sum method

46.1 Introduction

One of the more highly used valuation models is that developed by Gordon and Shapiro (1956) and Gordon (1962) known as the dividend growth model. In security analysis and portfolio management, growth rate estimates of earnings, dividends, and price per share are important factors in determining the value of an investment or a firm. These publications demonstrate that the growth rate is found by taking the product of the retention rate and the return on equity. What is less well understood are the basic assumptions of this model. In this paper, we demonstrate that the model makes strong assumptions regarding the financing mix of the firm.

In addition, we will also discuss several methods suggested in the literature on estimating growth rates. We will analyze whether these approaches are consistent with the use of using a constant discount rate to evaluate the firm's assets and equity. In particular, we will demonstrate that the underlying assumptions of the internal growth rate model (whereby no external funds are used to finance growth) is incompatible with the constant discount rate model of valuation.

The literature has also suggested estimating growth rate by taking the average of percentage change of dividends over a sample period, taking the geometric average of the change in dividends or using regression analysis to estimate the growth rate (e.g., Lee et al. 2009; Lee et al. 2012; Lee et al. 2000; and Ross et al. 2010). Gordon and Gordon (1997) suggest first using the Capital Asset Pricing Model (*CAPM*) to determine the cost of equity of the firm and then using the dividend growth model to infer the growth rate. Advantages, disadvantages, and the interrelationship among these estimation methods are also discussed in detail.

This paper is organized as follows. In Sect. 46.2 we present the Gordon and Shapiro model (1956). We discuss the inherent assumptions of the model and its implied method to estimate the growth rate. Section 46.3 analyzes the internal growth rate and sustainable growth rate models. Section 46.4 describes leading statistical methods for estimating firm's growth rates. We will also present the inferred method suggested by Gordon and Gordon (1997) to estimate the growth rate. Concluding remarks appear in Sect. 46.5.

46.2 The Discounted Cash Flow Model and the Gordon Growth Model

The traditional academic approach to evaluate a firm's equity is based upon the constant discount rate method. One approach uses the after-tax weighted average cost of capital as a discount rate. This model is expressed as:

$$Value\ of\ Equity = \sum_{t=1}^{\infty} \frac{CF_{ut}}{(1 + ATWACOC)^t} - Debt_t, \tag{46.1}$$

where CF_{ut} is the expected unlevered cash flow of the firm at time t and $Debt_t$ is the market value of debt outstanding. $ATWACOC$ equals $L(1 - \tau)R_d + (1 - L)r$ where L is the market value proportion of debt, τ is the corporate tax rate, R_d is the cost of debt and r is the cost of equity. The first term on the right hand side of Eq. 46.1 is the value of the assets. Subtracting out the value of debt yields the value of equity. The price per share is therefore the value of equity divided by the number of shares outstanding. Alternatively, the value of equity can be directly found by discounting the dividends per share by the cost of equity, or more formally:

$$Value\ of\ Common\ Stock\ (P_0) = \sum_{t=1}^{\infty} \frac{d_t}{(1 + r)^t}, \tag{46.2}$$

where d_t is the dividend per share at time t. Boudreaux and Long (1979), and Chambers et al. (1982) demonstrate the equivalence of these two approaches assuming that the level of that the level of debt is a constant percentage of the value of the firm.[1] Accordingly:

$$\frac{\sum_{t=1}^{\infty} \frac{X_t}{(1 + ATWACOC)^t} - Debt_t}{\text{\#of Shares Outstaning}} = \sum_{t=1}^{\infty} \frac{d_t}{(1 + r)^t} \tag{46.3}$$

If we assume that dividends per share grow at a constant rate g, then Eq. 46.2 is reduced to the basic dividend growth model[2]:

$$P_0 = \frac{d_1}{(r - g)}. \tag{46.4}$$

[1]See Brick and Weaver (1984, 1997) concerning the magnitude of error in the valuation using a constant discount rate when the firm does not maintain a constant market based leverage ratio.

[2]Gordon and Shapiro's (1956) model assume that dividends were paid continuously and hence $P_0 = d_1/(r - g)$.

Gordon and Shapiro (1956) demonstrates that if b is the fraction of earnings retained within the firm, and r is the rate of return the firm will earn on all new investments, then $g = br$. Let I_t denote the level of new investment at time t. Because growth in earnings arises from the return on new investments, earnings can be written as:

$$E_t = E_{t-1} + rI_{t-1}, \tag{46.5}$$

where E_t is the earnings in period t.[3] If the firm's retention rate is constant and used in new investment, then the earnings at time t is

$$E_t = E_{t-1} + rbE_{t-1} = E_{t-1}(1 + rb). \tag{46.6}$$

Growth rate in earnings is the percentage change in earnings and can be expressed as

$$g_E = \frac{E_t - E_{t-1}}{E_{t-1}} = \frac{E_{t-1}(1 + rb) - E_{t-1}}{E_{t-1}} = rb. \tag{46.7}$$

If a constant proportion of earnings is assumed to be paid out each year, the growth in earnings equals the growth in dividends, implying $g = br$. It is worthwhile to examine the implication of this model for the growth in stock prices over time. The growth in stock price is

$$g_P = \frac{P_{t+1} - P_t}{P_t}. \tag{46.8}$$

Recognizing that P_t and P_{t+1} can be defined by Eq. 46.4, while noting that d_{t+2} is equal to $d_{t+1}(1 + br)$ then:

$$g_P = \frac{\dfrac{d_{t+2}}{k - rb} - \dfrac{d_{t+1}}{k - rb}}{\dfrac{d_{t+1}}{k - rb}} = \frac{d_{t+2} - d_{t+1}}{d_{t+1}} = \frac{d_{t+1}(1 + br) - d_{t+1}}{d_{t+1}} = br. \tag{46.9}$$

Thus, under the assumption of a constant retention rate, for a one-period model, dividends, earnings, and prices are all expected to grow at the same rate.

The relationship between the growth rate, g, the retention rate, b, and the return on equity, r, can be expanded to a multi-period setting as the following numerical example illustrates. In this example, we assume that the book value of the firm's assets equal the market value of the firm. We will assume that the growth rate of the firm sales and assets is 4 % and the tax rate is equal to 40 %. The book value of the

[3]Earnings in this model are defined using the cash-basis of accounting and not on an accrual basis.

assets at time 0 is \$50 and we assume a depreciation rate of 10 % per annum. The amount of debt outstanding is \$12.50 and amount of equity outstanding is \$37.50. We assume that the cost of debt, R_d, is 12 % and the cost of equity, r, is 25 %, implying an *ATWACOC* of 20.55 %. The expected dividend at $t = 1$, d_1, must satisfy Eq. 46.4. That is, $37.50 = d_1/(0.25 - 0.04)$.

The unlevered cash flow is defined as Sales less Costs (excluding the depreciation expense) less Investment less the tax paid. Tax paid is defined as the tax rate (which we assume to be 40 %) times Sales minus Costs minus the Depreciation Expense. Recognizing that the value of the firm is given by $CF_{u1}/(ATWACOC - g)$, if firm value is \$50, $g = 4$ % and *ATWACOC* is 20.55 %, then the expected unlevered cash flow is at time 1 is \$8.28. We assume that the asset turnover ratio is 1.7. Hence, if assets at time 0 is \$50, the expected sales at time 1 is \$85. To obtain the level of investment, note that the depreciation expense at time 1 is \$5. If the book value of assets equals \$52, then the firm must invest \$7. To obtain an expected unlevered cash flow at $t = 1$ of \$8.28, the Gross Profit Margin is assumed to be approximately 26.03 %, resulting in expected costs at time 1 of \$62.88. The interest expense at time 1, is the cost of debt times the amount of debt outstanding at time zero, or \$1.50. The Earnings Before Taxes (*EBT*) is defined as Sales – Costs – Interest Expense – Depreciation Expense, which equals \$15.63 at time 1. 40 % of *EBT* is the taxes paid or \$6.25 resulting in a net income (*NI*) of \$9.38. *ROE*, which equals Net Income/Book Value of Equity at the beginning of the period is 25 %. Since the aggregate level of dividends at time 1 is \$7.88, then the dividend payout ratio $(1 - b)$ is 84 %. Note that b is therefore equal to 16 % and $b \times ROE = 4\%$.[4]

Further note that the firm will increase its book value of equity via retention of *NI* by \$1.50 (*RE* in the table). In order to maintain a leverage ratio of 25 %, the firm must increase the level of debt from time 0 to time 1 by \$0.50. The entries for time periods 2–5 follow the logical extension of the above discussion, and as shown in the table, the retention rate b is 16 % and $ROE = 25$ % for each period. Again the product of b and *ROE* results in the expected growth rate of 4 %. Further note, that $g = 4$ % imply that sales, costs, book value of asset, depreciation, unlevered cash flow, cash flow to stockholders, value of debt and value of equity to increase by 4 % per annum.

Investors may use a one-period model in selecting stocks, but future profitability of investment opportunities plays an important role in determining the value of the firm and its *EPS* and dividend per share. The rate of return on new investments can be expressed as a fraction, c (perhaps larger than 1), of the rate of return security holders require (r):

$$k = cr. \tag{46.10}$$

[4]Generally, practioners define ROE as the ratio of the Net Income to the end of year Stockholders Equity. Here we are defining ROE as the ratio of the Net Income to the beginning of the year Stockholders Equity. Brick et al. (2012) demonstrate that the practitioner's definition is one of the sources for the Bowman Paradox reported in the Organization Management literature.

Table 46.1 The book value of the firm's assets equal the market value of the firm (growth rate is 4 %)

	0	1	2	3	4	5
Assets	\$50.00	\$52.00	\$54.08	\$56.24	\$58.49	\$60.83
Debt	\$12.50	\$13.00	\$13.52	\$14.06	\$14.62	\$15.21
Equity	\$37.50	\$39.00	\$40.56	\$42.18	\$43.87	\$45.62
R_d	0.12	0.12	0.12	0.12	0.12	0.12
r	0.25	0.25	0.25	0.25	0.25	0.25
ATWACOC	0.2055	0.2055	0.2055	0.2055	0.2055	0.2055
Asset turnover		1.7	1.7	1.7	1.7	1.7
GPM		0.26029	0.26029	0.26029	0.26029	0.26029
Sales		\$85.00	\$88.40	\$91.94	\$95.61	\$99.44
Cost		\$62.88	\$65.39	\$68.01	\$70.73	\$73.55
Depreciation		\$5.00	\$5.20	\$5.41	\$5.62	\$5.85
Interest exp.		\$1.50	\$1.56	\$1.62	\$1.69	\$1.75
EBT		\$15.63	\$16.25	\$16.90	\$17.58	\$18.28
Tax		\$6.25	\$6.50	\$6.76	\$7.03	\$7.31
NI		\$9.38	\$9.75	\$10.14	\$10.55	\$10.97
DIV		\$7.88	\$8.19	\$8.52	\$8.86	\$9.21
New debt		\$.50	\$0.52	\$0.54	\$0.56	\$0.59
CF_u		\$8.28	\$8.61	\$8.95	\$9.31	\$9.68
Firm value	\$50.00	\$52.00	\$54.08	\$56.24	\$58.49	\$60.83
Investment		\$7.00	\$7.28	\$7.57	\$7.87	\$8.19
Vequity	\$37.50	\$39.00	\$40.56	\$42.18	\$43.87	\$45.62
RE		\$1.50	\$1.56	\$1.62	\$1.69	\$1.75
ROE		0.25	0.25	0.25	0.25	0.25
1-b		0.84	0.84	0.84	0.84	0.84
g		0.04	0.04	0.04	0.04	0.04

Substituting this into the well-known relationship that $r = \frac{d_1}{P_0} + g$ and rearranging, we have

$$k = \frac{(1-b)E_1}{(1-cb)P_0}. \tag{46.11}$$

If a firm has no extraordinary investment opportunities ($r = k$), then $c = 1$ and the rate of return that security holders require is simply the inverse of the stock's price to earnings ratio. In our example of Table 46.1, *NI* at time 1 is \$9.38 and the value of equity at time 0 is \$37.50. The ratio of these two numbers (which is equivalent to *EPS*/*P*) is *ROE* or 25 %.

On the other hand, if the firm has investment opportunities that are expected to offer a return above that required by the firm's stockholders ($c > 1$), the earnings to price ratio at which the firm sells will be below the rate of return required by investors. To illustrate consider the following example whereby market value of the

Table 46.2 The market value of the firm and equity is greater than its book value

	0	1	2	3	4	5
Assets	$50.00	$52.00	$54.08	$56.24	$58.49	$60.83
Firm value	$60.00	$62.40	$64.90	$67.49	$70.19	$73.00
Debt	$12.50	$13.00	$13.52	$14.06	$14.62	$15.21
Equity	$47.50	$49.40	$51.38	$53.43	$55.57	$57.79
R_d	0.12	0.12	0.12	0.12	0.12	0.12
r	0.25	0.25	0.25	0.25	0.25	0.25
ATWACOC	0.2129	0.2129	0.2129	0.2129	0.2129	0.2129
Asset turnover		1.7	1.7	1.7	1.7	1.7
GPM		0.3093	0.3093	0.3093	0.3093	0.3093
Sales		$85.00	$88.40	$91.94	$95.61	$99.44
Cost		$58.71	$61.06	$63.50	$66.04	$68.68
Depreciation		$5.00	$5.20	$5.41	$5.62	$5.85
Interest exp.		$1.50	$1.56	$1.62	$1.69	$1.75
EBT		$19.79	$20.58	$21.41	$22.26	$23.15
Tax		$7.92	$8.23	$8.56	$8.91	$9.26
NI		$11.88	$12.35	$12.84	$13.36	$13.89
DIV		$9.98	$10.37	$10.79	$11.22	$11.67
New debt		$.50	$0.52	$0.54	$0.56	$0.59
CF_u		$10.38	$10.79	$11.22	$11.67	$12.14
Firm value	$60.00	$62.40	$64.90	$67.49	$70.19	$73.00
Investment		$7.40	$7.70	$8.00	$8.32	$8.66
Vequity	$47.50	$49.40	$51.38	$53.43	$55.57	$57.79
RE		$1.90	$1.98	$2.06	$2.14	$2.22
Market based ROE		0.25	0.25	0.25	0.25	0.25
1-b		0.84	0.84	0.84	0.84	0.84
g		0.04	0.04	0.04	0.04	0.04

firm and equity is greater than its book value. This example is depicted in Table 46.2. The basic assumptions of the model is as follows: We will assume that the growth rate of the firm sales and book value of the assets is 4 %. The book value of the assets at time 0 is again $50 and we assume a depreciation rate of 10 % per annum. However, note that the market value of the firm is $60. The entries for Debt and Equity represent market values. The amount of debt outstanding is $12.50 and amount of equity outstanding is now $47.50. We assume that the cost of debt, R_d, is 12 % and the cost of equity, r, is 25 %, implying an *ATWACOC* of 21.29 %. For the valuation of the firm to be internally consistent, the unlevered cash flow at time 1 is $10.38. Similarly, the value of equity to be internally consistent, the expected dividends at $t = 1$ is $9.98. Note that net income is $11.88 implying a dividend payout ratio of 84 % and a retention rate of 16 %. The *book value based ROE*, k, is found by taking the net income divided by the book value of equity. In our example, implied book value of equity is $37.50. Hence, $k = 31.68$ %, implying that the book value ROE is greater than the cost of equity which is the required rate of return. But g is given by the *market value based ROE* which is defined as Net Income over market value of equity. That is $r = 25$ %. Note again, br is 4 %.

An investor could predict next year's dividends, the firm's long-term growth rate, and the rate of return stockholders require (perhaps using the *CAPM* to estimate r) for holding the stock. Equation 46.4 could then be solved for the theoretical price of the stock that could be compared with its present price. Stocks that have theoretical prices above actual price are candidates for purchase; those with theoretical prices below their actual price are candidates for sale or for short sale.

46.3 Internal Growth Rate and Sustainable Growth Rate Models

The internal growth rate model assumes that the firm can only finance its growth by its internal funds. Consequently, the cash to finance growth must come from only retained earnings. Therefore, retained earnings can be expressed as

$$\begin{aligned}\text{Retained Earnings} &= \text{Earnings} - \text{Dividends}\\ &= \text{Profit Margin} \times \text{Total Sales-Dividends}\\ &= p(S+\Delta S) - p(S+\Delta S)(1-b)\\ &= pb(S+\Delta S),\end{aligned} \tag{46.12}$$

where

$p =$ the profit margin on all sales;

$S =$ annual sales; and

$\Delta S =$ the increase in sales during the year.

Because retained earnings is the only source of new funds, the use of cash represented by the increase in assets must equal the retained earnings:

$$\text{Uses of Cash} = \text{Sources of Cash}$$

$$\text{Increases in Assets} = \text{Retained earnings}$$

$$\begin{aligned}\Delta ST &= p(S+\Delta S)b\\ &= pbS + pb\Delta S,\end{aligned}$$

$$\Delta S[T - pb] = pSb,$$

$$\frac{\Delta S}{S} = \frac{pb}{T - pb}, \tag{46.13}$$

where $T =$ the ratio of total assets to sales. If we divide both numerator and denominator of Eq. 46.13 by T and make rearrange the terms, then we can show that the internal growth rate is:

$$g = \frac{\Delta S}{S} = \frac{pb/T}{1 - pb/T} = \frac{b \times ROA}{1 - b \times ROA}, \tag{46.14}$$

where *ROA* is the return on assets. The internal growth rate is the maximum growth rate that can be achieved without debt or equity kind of external financing. But note this assumption of not issuing new debt or common stock to finance growth is inconsistent with the basic assumption of the constant discount rate models that the firm maintains a constant market based leverage ratio. Hence, this model cannot be used to estimate the growth rate and be employed by the Gordon Growth Model.

Higgins (1977, 1981, 2008) has developed a sustainable growth rate under assumption that firms can generate new funds by using retained earnings or issuing debt, but not issuing new shares of common stock. Growth and its management present special problems in financial planning. From a financial perspective, growth is not always a blessing. Rapid growth can put considerable strain on a company's resources, and unless management is aware of this effect and takes active steps to control it, rapid growth can lead to bankruptcy. Assuming a company is not raising new equity, the cash to finance growth must come from retained earnings and new borrowings. Further, because the company wants to maintain a target debt-to-equity ratio equal to *L*, each dollar added to the owners' equity enables it to increase its indebtedness by $*L*. Since the owners' equity will rise by an amount equal to retained earnings, the new borrowing can be written as:

$$\begin{aligned}\text{New Borrowings} &= \text{Retained Earnings} \times \text{Target Debt-to-Equity Ratio} \\ &= pb(S + \Delta S)L.\end{aligned}$$

The use of cash represented by the increase in assets must equal the two sources of cash (retained earnings and new borrowings)[5]:

$$\text{Uses of Cash} = \text{Sources of Cash}$$

$$\text{Increases in Assets} = \text{Retained Earnings} + \text{New Borrowing}$$

$$\begin{aligned}\Delta ST &= pb(S + \Delta S) + pb(S + \Delta S)L \\ &= pb(1 + L)S + pb(1 + L)\Delta S\end{aligned}$$

$$\Delta S[T - pb(1 + L)] = pb(1 + L)S$$

$$g = \frac{\Delta S}{S} = \frac{pb(1 + L)}{T - pb(1 + L)}. \tag{46.15}$$

[5]Increased in Assets is the net increase in assets. The total investment should also include the depreciation expense as can be seen in our examples delineated in Tables 46.1 and 46.2. But depreciation expense is also a source of funding. Hence, it is netted out in the relationship between increases in assets and retained earnings and new borrowings.

In Eq. 46.15 the $\Delta S/S$ or g is the firm's sustainable growth rate assuming no infusion of new equity. Therefore, a company's growth rate in sales must equal the indicated combination of four ratios, p, b, L, and T. In addition, if the company's growth rate differs from g, one or more of the ratios must change. For example, suppose a company grows at a rate in excess of g, then it must either use its assets more efficiently, or it must alter its financial policies. Efficiency is represented by the profit margin and asset-to-sales ratio. It therefore would need to increase its profit margin (p) or decrease its asset-to-sales ratio (T) in order to increase efficiency. Financial policies are represented by payout or leverage ratios. In this case, a decrease in its payout ratio (1-b) or an increase in its leverage (L) would be necessary to alter its financial policies to accommodate a different growth rate. It should be noted that increasing efficiency is not always possible and altering financial policies are not always wise.

If we divide both numerator and denominator of Eq. 46.15 by T and rearrange the terms, then we can show that the sustainable growth rate can be shown as

$$g = \frac{\Delta S}{S} = \frac{pb(1+L)/T}{1 - pb(1+L)/T} = \frac{b \times ROE}{1 - b \times ROE}. \tag{46.16}$$

Please note that, in the framework of internal growth rate and sustainable growth rate presented above, the source of cash are taken from the end of period values of assets and assumed that the required financing occurs at the end of the period. However, Ross et al. (2010) show that if the source of cash is from the beginning of the period, the relationship between the use and the source of cash can be expressed for the internal growth rate model as $\Delta ST = pSb$ and for the sustainable growth rate model, $\Delta ST = pbS + pbSL$. Such relationship will result an internal growth rate of $b \times ROA$ and a sustainable growth rate of $b \times ROE$. For example, Table 46.3 assumes identical assumptions to that of Table 46.1, but now we will assume a growth rate of 4.1667 % and use total asset, total equity, and total debt from the beginning of the period balance sheet to calculate the net income. Recall that ROE is the net income divided by stockholders' equity at the beginning of the period. Note that the product of ROE and b will yield 4.1667 %.

Note that the intent of the Higgins' sustainable growth rate allows only internal source and external debt financing. Chen et al. (2013) incorporate Higgins (1977) and Lee et al. (2011) frameworks, allowing company use both external debt and equity, and derive a generalized sustainable growth rate as

$$g(t) = \frac{b \times ROE}{1 - b \times ROE} + \frac{\lambda \cdot \Delta n \cdot P/E}{1 - b \times ROE}, \tag{46.17}$$

where

$\lambda =$ degree of market imperfection;
$\Delta n =$ number of shares of new equity issued;
$P =$ price per share of new equity; and
$E =$ total equity.

Table 46.3 The book value of the firm's assets equal the market value of the firm (sustainable growth rate is 4.1667 %)

	0	1	2	3	4	5
Assets	$50.00	$52.08	$54.25	$56.51	$58.87	$61.32
Value	$50.00	$52.08	$54.25	$56.51	$58.87	$61.32
Debt	$12.50	$13.02	$13.56	$14.13	$14.72	$15.33
Equity	$37.50	$39.06	$40.69	$42.39	$44.15	$45.99
R	0.12	0.12	0.12	0.12	0.12	0.12
Re	0.25	0.25	0.25	0.25	0.25	0.25
ATWACOC	0.2055	0.2055	0.2055	0.2055	0.2055	0.2055
Asset turnover		1.7	1.7	1.7	1.7	1.7
GPM		0.26029	0.26029	0.26029	0.26029	0.26029
Sales		$85.00	$88.54	$92.23	$96.07	$100.08
Cost		$62.88	$65.49	$68.22	$71.07	$74.03
Depreciation		$5.00	$5.21	$5.43	$5.65	$5.89
Interest exp.		$1.50	$1.56	$1.63	$1.70	$1.77
EBT		$15.63	$16.28	$16.95	$17.66	$18.40
Tax		$6.25	$6.51	$6.78	$7.06	$7.36
NI		$9.38	$9.77	$10.17	$10.60	$11.04
DIV		$7.81	$8.14	$8.48	$8.83	$9.20
New debt		$7.60	$7.92	$8.25	$8.59	$8.95
CFu		$8.19	$8.53	$8.89	$9.26	$9.64
Value	$50.00	$52.08	$54.25	$56.51	$58.87	$61.32
Investment		$7.08	$7.38	$7.69	$8.01	$8.34
RE		$1.56	$1.63	$1.70	$1.77	$1.84
ROE		0.25	0.25	0.25	0.25	0.25
(1-b)		0.833333	0.833333	0.833333	0.833333	0.833333
g		0.041667	0.041667	0.041667	0.041667	0.041667

Comparing Eq. 46.17, the generalized sustainable growth rate has an additional positive term, $\frac{\lambda \cdot \Delta n \cdot p/E}{1-(1-D)ROE}$, when the new equity issue is taken into account. Therefore, Chen et al. (2013) show that Higgins' (1977) sustainable growth rate is underestimated because of the omission of the source of the growth related to new equity issue.

46.4 Statistical Methods

Instead of relying on financial ratios to estimate firm's growth rates, one may use statistical methods to determine firm's growth rates. A simple growth rate can be estimated by calculating the percentage change in earnings over a time period, and taking the arithmetic average. For instance, the growth rate in earnings over one period can be expressed as:

$$g_t = \frac{E_t - E_{t-1}}{E_{t-1}}. \tag{46.18}$$

The arithmetic average is given by

$$\overline{g} = \frac{1}{n}\sum_{t=1}^{n} g_t. \tag{46.19}$$

A more accurate estimate can be obtained by solving for the compounded growth rate:

$$X_t = X_0(1+g)^t, \tag{46.20}$$

or

$$g = \left(\frac{X_t}{X_0}\right)^{1/t} - 1, \tag{46.21}$$

where

X_0 = measure in the current period (measure can be sales, earnings, or dividends); and

X_t = measure in period t.

This method is called the *discrete compound sum* method of growth-rate estimation. For this approach to be consistent with the dividend growth model, the duration of each period (e.g., quarterly or yearly) must be consistent with the compounding period used in the dividend growth model.

Another method of estimating the growth rate uses the continuous compounding process. The concept of continuous compounding process can be expressed mathematically as

$$X_t = X_0 e^{gt}. \tag{46.22}$$

Equation 46.21 describes a discrete compounding process and Eq. 46.22 describes a continuous compounding process. The relationship between Eqs. 46.21 and 46.22 can be illustrated by using an intermediate expression such as:

$$X_t = X_0\left(1 + \frac{g}{m}\right)^{mt}, \tag{46.23}$$

where m is the frequency of compounding in each year. If $m = 4$, Eq. 46.23 implies a quarterly compounding process; if $m = 365$, it describes a daily process; and if m approaches infinity, it describes a continuous compounding process. Thus Eq. 46.22 can be derived from Eq. 46.23 based upon the definition

$$\lim_{m \to \infty} \left(1 + \frac{1}{m}\right)^m = e. \tag{46.24}$$

Then the continuous analog for Eq. 46.20 can be rewritten as

$$\lim_{m \to \infty} X_t = \lim_{m \to \infty} X_0 \left(1 + \frac{g}{m}\right)^{mt} = X_0 \lim_{m \to \infty} \left(1 + \frac{1}{m/g}\right)^{\left(\frac{m}{g}\right)gt} = X_0 e^{gt}. \tag{46.25}$$

Therefore, the growth rate estimated by continuous compound-sum method can be expressed by

$$g = \frac{1}{t} \ln \frac{X_t}{X_0}. \tag{46.26}$$

If you estimate the growth rate via Eq. 46.26, you are implicitly assuming the dividends are growing continuously and therefore the dividend growth model. In this case, according to Gordon and Shapiro's (1956) model, $P_0 = d_0/(r - g)$.

To use all the information available to the security analysts, two regression equations can be employed. These equations can be derived from Eqs. 46.20 and 46.22 by taking the logarithm (*ln*) on both sides of equation:

$$\ln X_t = \ln X_0 + t \ln(1 + g). \tag{46.27}$$

If Eq. 46.27 can be used to estimate the growth rate, then the antilog of the regression slope estimate would equal the growth rate. For the continuous compounding process,

$$\ln X_t = \ln X_0 + gt. \tag{46.28}$$

Both Eqs. 46.27 and 46.28 indicate that X_n is linearly related to t; and the growth rate can be estimated by the ordinary least square (OLS) regression. For example, growth rates for *EPS* and *DPS* can be obtained from an OLS regression by using

$$\ln\left(\frac{\text{EPS}_t}{\text{EPS}_0}\right) = a_0 + a_1 T + \varepsilon_{1t}, \tag{46.29}$$

and

$$\ln\left(\frac{\text{DPS}_t}{\text{DPS}_0}\right) = b_0 + b_1 T + \varepsilon_{2t}, \tag{46.30}$$

where EPS_t and DPS_t are earnings per share and dividends per share, respectively, in period t, and T is the time indicators (i.e., $T = 1, 2, \ldots, \text{n}$). We denote $\hat{a}_1$ and $\hat{b}_1$ as

Table 46.4 Dividend behavior of firms Pepsico and Wal-Mart in dividends per share (*DPS*)

Year	T	PEP	WMT	Year	T	PEP	WMT
1981	1	3.61	1.73	1996	16	0.72	1.33
1982	2	2.4	2.5	1997	17	0.98	1.56
1983	3	3.01	1.82	1998	18	1.35	1.98
1984	4	2.19	1.4	1999	19	1.4	1.25
1985	5	4.51	1.91	2000	20	1.51	1.41
1986	6	1.75	1.16	2001	21	1.51	1.49
1987	7	2.30	1.59	2002	22	1.89	1.81
1988	8	2.90	1.11	2003	23	2.07	2.03
1989	9	3.40	1.48	2004	24	2.45	2.41
1990	10	1.37	1.9	2005	25	2.43	2.68
1991	11	1.35	1.14	2006	26	3.42	2.92
1992	12	1.61	1.4	2007	27	3.48	3.17
1993	13	1.96	1.74	2008	28	3.26	3.36
1994	14	2.22	1.02	2009	29	3.81	3.73
1995	15	2.00	1.17	2010	30	3.97	4.2

the estimated coefficients for Eqs. 46.29 and 46.30. The estimated growth rates for *EPS* and *DPS*, therefore, are $\exp(\hat{a}_1) - 1$ and $\exp(\hat{b}_1) - 1$ in terms of discrete compounding process and $\hat{a}_1$ and $\hat{b}_1$ in terms of continuous compounding process.[6]

Table 46.4 provides dividends per share of Pepsico and Wal-Mart during the period from 1981 to 2010. Using the data in Table 46.4 for companies Pepsico and Wal-Mart, we can estimate the growth rates for their respective dividend streams. Table 46.5 presents the estimated the growth rates for Pepsico and Wal-Mart by arithmetic average method, geometric average method, compound-sum method, and the regression method in terms of discrete and continuous compounding processes. Graphs of the regression equations for Pepsico and Wal-Mart are shown in Fig. 46.1.

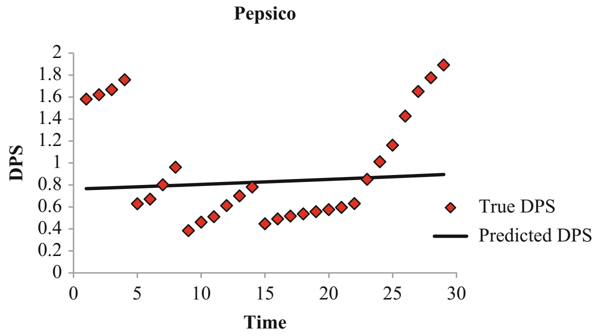
The slope of the regression for Pepsico shows an estimated coefficient for the intercept is 0.56. The estimated intercept for Wal-Mart is 7.04. The estimated growth rates for Pepsico and Wal-Mart, therefore, are 0.56 % and 7.29 % in terms of discrete compounding process. Figure 46.1 also shows the true DPS and predicted DPS for Pepsico and Wal-Mart. We find that the regression method, to some extent, can estimate the growth rate for Wal-Mart more precisely than for Pepsico. Comparing to the geometric average method, the regression method yields a similar value of the estimated growth rate for Wal-Mart, while not for Pepsico.

There are some complications to be aware of when employing the arithmetic average, the geometric average, and regression model in estimating the growth rate. The arithmetic average is quite sensitive to extreme values. The arithmetic average, therefore, has an upward bias that increases directly with the variability of the data.

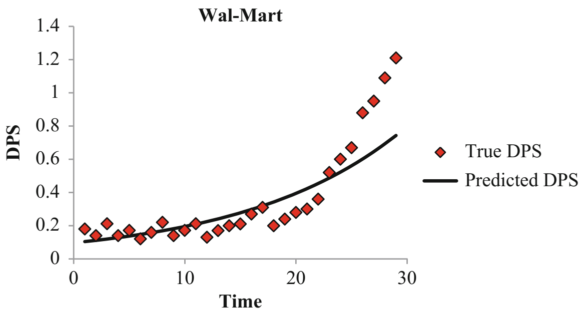
[6]If the earnings (or dividend) process follows Eq. 46.27, we can get same results from the non-restricted model as Eqs. 46.29 and 46.30.

Table 46.5 Estimated dividend growth rates for Pepsico and Wal-Mart

	Pepsico (%)	Wal-Mart (%)
Arithmetic average	4.64	8.99
Geometric average	0.99	5.45
Compound-sum method	0.99	5.30
Regression method (continuous)	0.56	7.04
Regression method (discrete)	0.56	7.29



$$\ln\left(\frac{DPS_t}{DPS_0}\right) = \underset{(0.0056)}{-0.6236} + \underset{(0.0113)}{0.1947}\, T + \varepsilon_t$$



$$\ln\left(\frac{DPS_t}{DPS_0}\right) = \underset{(0.1286)}{-0.9900} + \underset{(0.0075)}{0.0704}\, T + \varepsilon_t$$

Fig. 46.1 Regression models for Pepsico and Wal-Mart

Consider the following situation. Dividends in years 1, 2 and 3 are \$2, \$4 and \$2. The arithmetic average of growth rate is 25 % but the true growth rate is 0 %. The difference in the two average techniques will be greater when the variability of the data is larger. Therefore, it is not surprising that we find differences in the estimated growth rates using arithmetic average and geometric average methods for Pepsico and Wal-Mart in Table 46.5.

Table 46.6 Estimated dividend growth rates for 50 randomly selected companies

	50 Firms (%)	Firms with positive growth (35 firms) (%)	Firms with negative growth (15 firms) (%)
Arithmetic average	4.95	7.27	−0.47
Geometric average	0.93	3.00	−3.88
Compound-sum method	0.83	2.91	−4.02
Regression method (continuous)	0.66	2.32	−3.22
Regression method (discrete)	0.71	2.37	−3.15

The regression method uses more available information than the geometric average, discrete compounding and continuous compounding methods in that it takes into account the observed growth rates between the first and last period of the sample. A null hypothesis test can be used to determine whether the growth rate obtained from the regression method is statistically significantly different from zero or not. However, logarithms cannot be taken with zero or negative numbers. Under this circumstance the arithmetic average will be a better alternative.

We further randomly select 50 companies from S&P 500 index firms, which paid dividends during 1981–2010, to estimate their dividend growth rates by arithmetic average method, geometric average method, compound-sum method, and the regression method in terms of discrete and continuous compounding processes. Table 46.6 shows averages of estimated dividend growth rates for 50 random companies by different methods. As we discussed before, the arithmetic average is sensitive to extreme values and has an upward bias. We, therefore, find a larger average of the estimated dividend growth rate using the arithmetic average method. We also find that on average, the geometric, and compound sum methods yield relatively smaller growth rate estimates as compared to the estimates obtained using the regression methods to estimate growth rate. However, it appears that estimates obtained using the geometric, compound sum and regression methods are very similar.

Finally, Gordon and Gordon (1997) suggest that one can infer the growth rate using the dividend growth model. In particular, the practitioner can use regression analysis to calculate the beta of the stock and use the *CAPM* to estimate the cost of equity. Since

$$P_0 = \frac{d_0(1+g)}{(r-g)} \tag{46.31}$$

and the price of the stock is given by the market, the cost of equity is obtained using the *CAPM*, and d_0 and the current dividend is known, one can infer the growth rate using Eq. 46.31. If the inferred growth rate is less than the practitioner's estimate, then the recommendation will be to buy the stock. On the other hand, if the inferred growth is greater than the practitioner's estimate, the recommendation will be to sell the stock. However, it should be noted that the explanatory power of the *CAPM* to explain the relationship between stock returns and risk has been extensively questioned in the literature. See for example, Fama and French (1992).

46.5 Conclusion

The most common valuation model is the dividend growth model. The growth rate is found by taking the product of the retention rate and the return on equity. What is less well understood are the basic assumptions of this model. In this paper, we demonstrate that the model makes strong assumptions regarding the financing mix of the firm. In addition, we discuss several methods suggested in the literature on estimating growth rates and analyze whether these approaches are consistent with the use of using a constant discount rate to evaluate the firm's assets and equity. In particular, we demonstrate that the underlying assumptions of the internal growth rate model (whereby no external funds are used to finance growth) are incompatible with the constant discount rate model of valuation. The literature has also suggested estimating growth rate by using the average percentage change method, compound-sum method, and/or regression methods. We demonstrate that the average percentage change is very sensitive to extreme observations. Moreover, on average, the regression method yields similar but somewhat smaller estimates of the growth rate compared to the compound-sum method. We also discussed the inferred method suggested by Gordon and Gordon (1997) to estimate the growth rate. Advantages, disadvantages, and the interrelationship among these estimation methods are also discussed in detail. Choosing an appropriate method to estimate firm's growth rate can yield a more precise estimation and be helpful for the security analysis and valuation. However, all of these methods use historical information to obtain growth estimates. To the extent that the future may differ from the past, will ultimately determine the efficacy of any of these methods.

References

Boudreaux, K. J., & Long, R. W. (1979). The weighted average cost of capital as a cutoff rate: A further analysis. *Financial Management, 8*, 7–14.

Brick, I. E., & Weaver, D. G. (1984). A comparison of capital budgeting techniques in identifying profitable investments. *Financial Management, 13*, 29–39.

Brick, I. E., & Weaver, D. G. (1997). Calculating the cost of capital of an unlevered firm for use in project evaluation. *Review of Quantitative Finance and Accounting, 9*, 111–129.

Brick, I. E., Palmon, O., & Venezia, I. (2012). The risk-return (Bowman) paradox and accounting measurements. In I. Venezia & Z. Wiener (Eds.), *Bridging the GAAP: Recent advances in finance and accounting*. London: World Scientific Publishing Company.

Brick, I. E., Chen H. Y., Hsieh C. H., & Lee C. F. (2014, forthcoming). A comparison of alternative models for estimating firm's growth rate. *Review of Quantitative Finance and Accounting*.

Chambers, D. R., Harris, R. S., & Pringle, J. J. (1982). Treatment of financing mix in analyzing investment opportunities. *Financial Management, 11*, 24–41.

Chen, H. Y., Gupta, M. C., Lee, A. C., & Lee, C. F. (2013). Sustainable growth rate, optimal growth rate, and optimal payout ratio: A joint optimization approach. *Journal of Banking and Finance, 37*, 1205–1222.

Fama, E. F., & French, K. R. (1997). The cross-section of expected stock returns. *The Journal of Finance, 47*, 427–465.

Gordon, J., & Gordon, M. (1997). The finite horizon expected return model. *The Financial Analysts Journal, 53*, 52–61.

Gordon, M., & Shaprio, E. (1956). Capital equipment analysis: The required rate of profit. *Management Science, 3*, 102–110.

Higgins, R. C. (1977). How much growth can a firm afford? *Financial Management, 6*, 7–16.

Higgins, R. C. (1981). Sustainable growth under inflation. *Financial Management, 10*, 36–40.

Higgins, R. C. (2008). *Analysis for financial management* (9th ed.). New York: McGraw-Hill.

Lee, C. F., & Lee, A. C. (2013). *Encyclopedia of finance* (2nd ed.). New York: Springer.

Lee, C. F., Lee, J. C., & Lee, A. C. (2000). *Statistics for business and financial economics* (2nd ed.). Singapore: World Scientific Publishing Company.

Lee, A. C., Lee, J. C., & Lee, C. F. (2009). *Financial analysis, planning and forecasting: Theory and application* (2nd ed.). Singapore: World Scientific Publishing Company.

Lee, C. F., Finnerty, J. E., Lee, A. C., Lee, J., & Wort, D. H. (2012a). *Security analysis, portfolio management, and financial derivatives* (2nd ed.). New York: Springer.

Lee, C. F., Gupta, M. C., Chen, H. Y., & Lee, A. C. (2012b). Optimal payout ratio under uncertainty and the flexibility hypothesis: Theory and empirical evidence. *Journal of Corporate Finance, 17*, 483–501.

Ross, S. A., Westerfield, R. W., & Jaffe, J. (2010). *Corporate finance* (9th ed.). New York: McGraw-Hill/Irwin.

Econometric Measures of Liquidity

47

Jieun Lee

Contents

Abstract

A security is *liquid* to the extent that an investor can trade significant quantities of the security quickly, at or near the current market price, and bearing low transaction costs. As such, liquidity is a multidimensional concept. In this chapter, I review several widely used econometrics or statistics-based measures that researchers have developed to capture one or more dimensions of a security's liquidity (i.e., limited dependent variable model (Lesmond, D. A. et al. *Review of Financial Studies, 12*(5), 1113–1141, 1999) and autocovariance of price changes (Roll, R., *Journal of Finance, 39*, 1127–1139, 1984). These alternative proxies have been designed to be estimated using either low-frequency or high-frequency

J. Lee
Economic Research Institute, Bank of Korea, Seoul, South Korea
e-mail: jelee@bok.or.kr

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_99,

data, so I discuss four liquidity proxies that are estimated using low-frequency data and two proxies that require high-frequency data. Low-frequency measures permit the study of liquidity over relatively long time horizons; however, they do not reflect actual trading processes. To overcome this limitation, high-frequency liquidity proxies are often used as benchmarks to determine the best low-frequency proxy. In this chapter, I find that estimates from the effective tick measure perform best among the four low-frequency measures tested.

Keywords

Liquidity • Transaction costs • Bid-ask spread • Price impact • Percent effective spread • Market model • Limited dependent variable model • Tobin's model • Log likelihood function • Autocovariance • Correlation analysis

47.1 Introduction

A security is *liquid* to the extent that an investor can trade significant quantities of the security quickly, at or near the current market price, and bearing low transaction costs. A security's liquidity is an important characteristic variable, relevant in asset pricing studies, studies of market efficiency, and even corporate finance. In the asset pricing literature, researchers have considered whether liquidity is a priced risk factor (e.g., Amihud and Mendelson 1986; Brennan and Subrahmanyam 1996; Amihud 2002; Pastor and Stambaugh 2003). In corporate finance, researchers have found that liquidity is related to capital structure, mergers and acquisitions, and corporate governance (e.g., Lipson 2003; Lipson and Mortal 2007, 2009; Bharath 2009; Chung et al. 2010).

In these and many other studies, researchers have chosen from a variety of liquidity measures that have been developed. In turn, the variety of available liquidity measures reflects the multidimensional aspect of liquidity. Note that the definition of liquidity given above features four dimensions of liquidity: trading quantity, trading speed, price impact, and trading cost. Some extant measures focus on a single dimension of liquidity, while others encompass several dimensions. For instance, the bid-ask spread measure in Amihud and Mendelson (1986), the estimator of the effective spread in Roll (1984), and the effective tick estimator in Goyenko et al. (2009) relate to the trading cost dimension. The turnover measure of Datar et al. (1998) captures the trading quantity dimension. The measures in Amihud (2002) and Pastor and Stambaugh (2003) are relevant to price impact. The number of zero trading volume days in Liu (2006) emphasizes trading speed. Finally, and different from the others, the measure in Lesmond et al. (1999) encompasses several dimensions of liquidity.

Among the available measures, this chapter focuses on six liquidity proxies, including four that are commonly estimated using low-frequency data (i.e., daily closing prices) and two that are commonly estimated using high-frequency data (i.e., intraday trades and quotes). The low-frequency measures are in Roll (1984), Goyenko et al. (2009), Lesmond et al. (1999), and Amihud (2002).

The high-frequency measures are the percent quoted spread and the percent effective spread. The low-frequency proxies are advantageous because they are more amenable to the study of liquidity over relatively long time horizons and across countries. However, they are limited because they do not directly reflect actual trading processes, while the high-frequency measures do. Thus, high-frequency liquidity proxies are often used as benchmarks to determine the best low-frequency proxy. This is not a universal criterion, however, because each measure captures a different dimension of liquidity and may lead to different results in specific cross-sectional or time-series applications.

The remainder of this chapter is organized as follows. In Sects. 47.2 and 47.3, I introduce and briefly discuss each of the low-frequency and high-frequency liquidity measures, respectively. Section 47.4 provides an empirical analysis of these liquidity measures, including the aforementioned test of the best low-frequency measure. Section 47.5 concludes.

47.2 Low-Frequency Liquidity Proxies

Below I describe four widely used measures of liquidity: the Roll (1984) measure; effective tick; the Amihud (2002) measure; and the Lesmond et al. (1999) measure.

47.2.1 The Roll Measure

Roll (1984) develops a measure of the *effective* bid-ask spread. He assumes that the true value of a stock follows a random walk and that P_t, the observed closing price on day t, is equal to the stock's true value plus or minus half of the effective spread. He also assumes that a security trades at either the bid price or the ask price, with equal frequency. This relationship can be expressed as follows:

$$P_t = P_t^* + Q_t \frac{S}{2}$$

$$Q_t \sim \mathrm{IID}\begin{bmatrix} +1 \text{ with probability } 1/2 \ \text{(buyer initiated)} \\ -1 \text{ with probability } 1/2 \ \text{(seller initiated)} \end{bmatrix}$$

where Q_t is an order-type indicator variable, indicating whether the transaction at time t is at the ask (buyer-initiated) or at the bid (seller-initiated) price. His assumption that P_t^* is the fundamental value of the security implies that $E[Q_t] = 0$; hence, $\Pr(Q_t = 1) = \Pr(Q_t = -1) = 1/2$. Also, there are no changes in the fundamental value of the security (i.e., over a short horizon).

It follows that the process for price changes ΔP_t is

$$\Delta P_t = \Delta P_t^* + \left(Q_t - Q_{(t-1)}\right)\frac{S}{2} = \left(Q_t - Q_{(t-1)}\right)\frac{S}{2}.$$

Under the assumption that Q_t is IID, the variance and covariance of ΔP_t can be easily calculated:

$$\mathrm{var}[\Delta P_t] = \frac{s^2}{2}.$$

$$\mathrm{cov}[\Delta P_t, \Delta P_{t-1}] = -\frac{s^2}{4}.$$

$$\begin{aligned}
\mathrm{cov}[\Delta P_t, \Delta P_{t-1}] &= \mathrm{cov}\left[\frac{s}{2}(Q_t - Q_{t-1}), \frac{s}{2}(Q_{t-1} - Q_{t-2})\right] \\
&= \frac{s^2}{4}[\mathrm{cov}(Q_t - Q_{t-1}), \mathrm{cov}(Q_{t-1} - Q_{t-2})] \\
&= \frac{s^2}{4}[\mathrm{cov}(Q_t, Q_{t-1}) - \mathrm{cov}(Q_{t-1}, Q_{t-1}) \\
&\quad + \mathrm{cov}(Q_{t-1}, Q_{t-2}) - \mathrm{cov}(Q_t, Q_{t-2})] \\
&= \frac{s^2}{4}[-\mathrm{var}(Q_{t-1})] = -\frac{s^2}{4}\left[\frac{1}{2}(1-0)^2 + \frac{1}{2}(-1-0)^2\right] = -\frac{s^2}{4}
\end{aligned}$$

$$\mathrm{cov}[\Delta P_t, \Delta P_{k-1}] = 0, k > 1.$$

Solving for S yields Roll's effective spread estimator:

$$S = 2\sqrt{-\mathrm{Cov}\left(\Delta P_{t,}, \Delta P_{t-1}\right)}.$$

Roll's measure is simple and intuitive: If P* is fixed so that prices take only two values, bid or ask, and if the current price is the bid, then the change between current price and previous price must be either 0 or –s and the price change between current price and next price must be either 0 or s. Analogous possible price changes apply when the current price is the ask.

The Roll measure S is generally estimated using daily data on price changes. Roll (1984) and others have found that for some individual stocks, the autocovariance that defines S is positive, rather than negative, so that S is undefined. In this case, researchers generally choose one of three solutions: (1) treat the observation as missing, (2) set the Roll spread estimate to zero, or (3) multiply the covariance by negative one, calculate S, and multiply this estimate by negative one to produce a negative spread estimate. In my empirical analysis to follow, I find that results are insensitive to the alternative solutions, so I only report results of setting S to zero when the observed autocovariance is positive.

47.2.2 Effective Tick

Goyenko et al. (2009) and Holden (2009) develop an *effective tick* measure that is based on price clustering and changes in tick size. Below I describe the effective tick measure in Goyenko et al. (2009), which is elegant in its simplicity.

Consider four possible bid-ask spreads for a stock: \$1/8, \$1/4, \$1/2, and \$1. If the spread is \$1/4, the authors assume that bid and ask prices are associated with only even quarters. Thus, if an odd-eighth transaction price shows up, it is instead inferred that the spread is \$1/8. The number of quotes that occur at \$1/8 spread is given by N_1. The number of quotes at odd-quarter fractions (\$1/8 and \$1/4) is N_2. The number of quotes at odd-half (\$1/8, \$1/4, and \$1/2) is N_3. Finally, the number of whole dollar quotes is given by N_4. The following is the proportion of each price fraction observed during the day:

$$F_i = \frac{Ni}{\sum_{i=1}^{I} Ni} \quad \text{for } i = 1, \ldots, I.$$

Next, suppose that the unconstrained probability of the effective *i*th estimated spread is

$$\begin{array}{c} 2F_i \;\; i = 1 \\ U = 2F_i - F_{i-1} \;\; i = 2, \ldots, i \\ F_i - F_{i-1} \;\; i = I. \end{array}$$

The effective tick is a simply probability-weighted average of effective spread size divided by average price in a given time interval:

$$\text{Effective Tick}_{it} = \frac{\sum_{i=1}^{I} a * S_i}{P},$$

where the probability a is constrained to be nonnegative and to be no more than 1 minus the probability of a finer spread, S is the spread, and P is the average price in the time interval.

To obtain estimates of effective tick in this chapter, I must deal with changes in minimum tick size that were instituted over time in the US equity markets. For NYSE, AMEX, and NASDAQ stocks from 1/93 to 5/97, I used a fractional grid accounting for price increments as small as \$1/8. For NYSE and AMEX (NASDAQ) stocks from 6/97 to 1/01 (6/97 to 3/01), I used a minimum tick size increment of \$1/16. Thereafter, I used a decimal grid for all stocks.

47.2.3 Amihud (2002)

The measures in Amihud (2002) and Pastor and Stambaugh (2003) both purport to capture the price impact dimension of liquidity. Goyenko et al. (2009) show that the Amihud (2002) measure performs well in measuring price impact while the Pastor and Stambaugh (2003) measure is dominated by other measures. Pastor and Stambaugh (2003, p. 679) also caution against their measure as a liquidity measure for individual stocks, reporting large sampling errors in individual estimates.

Referring also to Hasbrouck (2009), I do not discuss the Pastor and Stambaugh (2003) measure. The Amihud (2002) measure is a representative proxy for price impact, i.e., the daily price response associated with one dollar of trading volume:

$$\mathrm{Ami}_{\mathrm{it}} = \frac{|\mathrm{Ret}_{\mathrm{it}}|}{\mathrm{Vol}_{\mathrm{it}}},$$

where $\mathrm{Ret}_{\mathrm{it}}$ is the stock i's return on day t and $\mathrm{Vol}_{\mathrm{it}}$ is the stock i's dollar volume on day t. The average is calculated over all positive-volume days, since the ratio is undefined for zero volume days.

47.2.4 Lesmond, Ogden, and Trzcinka (LOT 1999)

The Lesmond et al. (LOT 1999) liquidity measure is based on the idea that an informed trader observing a mispriced stock will execute a trade only if the difference between the current market price and the true price exceeds the trader's transaction costs; otherwise, no trade will occur. Therefore, they argue that a stock with high transaction costs will have less frequent price movements and more zero returns than a stock with low transaction costs. Based on this relationship, they develop a measure of the marginal trader's effective transaction costs for an individual stock. Their measure utilizes the limited dependent variable regression model of Tobin (1958) and Rosett (1959) applied to the "market model."

47.2.4.1 Market Model

The basic market model is a regression of the return, R^*_{it} on security i and period t, on the contemporaneous market return, R_{mt}:

$$R^*_{\mathrm{it}} = a_i + \beta_i R_{\mathrm{mt}} + \varepsilon_{\mathrm{it}} \tag{47.1}$$

The market model implies that a security's return reflects the effect of new information on the value of the stock, which can be divided into two components: contemporaneous market-wide information ($\beta_i R_{\mathrm{mt}}$) and firm-specific information $\varepsilon_{\mathrm{it}}$. In an ideal market without frictions such as transaction costs, new information will be immediately reflected into the security's price, so R^*_{it} is the true return on security i.

47.2.4.2 Relationship Between Observed and True Returns

In the presence of transaction costs, investors will trade only when the marginal profits exceed the marginal transaction costs. In this context, transaction costs would include various dimensions such as bid-ask spread, commissions, and price impact, as well as taxes or short-selling costs, because investors make trading decisions after considering overall transaction costs. Transaction costs inhibit informative trades and therefore drive a wedge between observed returns and true

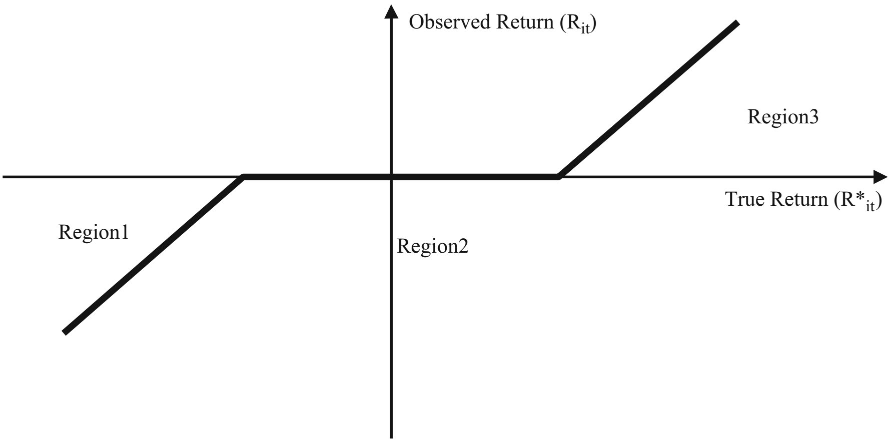


Fig. 47.1 This figure illustrates the relationship between the observed return on a stock in the presence of transaction costs that inhibit trading, R_{it}, and its true return in the absence of transaction costs, R^*_{it}, where the latter reflects the true effects of new market-wide or firm-specific information. The relationship can be divided into three regions: (1) Region 1, where the value of new information is negative and exceeds transaction costs; (2) Region 2, where the transaction costs exceed the value of new information regardless of the direction of the value of information; and (3) Region 3, where the value of new information is positive and exceeds transaction costs

returns. In the presence of transaction costs, Lesmond et al. (1999) propose the following relationship between observed and true returns:

$$R_{it} = R^*_{it} - \mu S_i, \tag{47.2}$$

where μS_i is the spread adjustment for security i, R_{it} is the observed return, and R^*_{it} is true return.

Specifically, the relationship is as follows:

$$\begin{array}{lll} R_{it} = R^*_{it} - \alpha_{1i} & \text{if} & R^*_{it} < \alpha_{1i} \\ R_{it} = 0 & \text{if} & \alpha_{1i} < R^*_{it} < \alpha_{2i} \\ R_{it} = R^*_{it} - \alpha_{2i} & \text{if} & R^*_{it} > \alpha_{2i} \end{array} \tag{47.3}$$

where $\alpha_{1i} < 0$ and $\alpha_{2i} > 0$. α_{1i} is the transaction costs for the marginal investor when information has a negative shock (selling), and α_{2i} is the transaction costs for the marginal investor when information has a positive shock (buying). Consequently, the difference between α_{1i} and α_{2i} is a measure of round-trip transaction costs.

If the true return exceeds transaction costs, the marginal investor will continuously trade, and the market price will respond until, for the next trade, marginal profit is equal to marginal transaction costs. If the transaction costs are greater than true returns, then the marginal investors will not trade, price will not move, and consequently the zero returns will occur. Therefore, in this model the frequency of zero returns is a simple alternative measure of transaction costs. The relationship between observed returns and true returns is illustrated in Fig. 47.1.

47.3 High-Frequency Liquidity Proxies

Next I describe two well-known spread proxies that can be estimated using high-frequency data. These are the percent quoted spread and percent effective spread.

47.3.1 Percent Quoted Spread

The ask (bid) quotation is the price at which shares can be purchased (sold) with immediacy. The difference, known as the percent quoted spread, is the cost of a round-trip transaction and is generally expressed as a proportion of the average of the bid and ask prices:

$$\text{Percent quoted spread}_{\text{it}} = \frac{\text{Ask}_{\text{it}} - \text{Bid}_{\text{it}}}{\text{M}_{\text{it}}}.$$

In the empirical analysis in the next section, I estimate percent quoted spreads using high-frequency data. Following convention, for each stock and trading day, I find the highest bid and lowest ask prices over all venues at every point during the day, denoting these "inside" ask and bid prices as Ask_{it} and Bid_{it}, respectively. M_{it} is then the average of, or midpoint between, Ask_{it} and Bid_{it}. I then calculate the average percent quoted spread for a stock and day as the time-weighted average of all spreads observed for that stock during the day. Finally, *percent quoted spread* for each stock is calculated by averaging the daily estimates across all trading days within a given month.

47.3.2 Percent Effective Spread

Some trades occur *within* the range of inside bid and ask quotes, as when simultaneous buy and sell market orders are simply crossed. Thus, the inside bid-ask spread may overestimate the realized amount of this component of transaction costs. Hasbrouck's (2009) measure of *percent effective spread* attempts to adjust for this bias. For a given stock, percent effective spread is computed for all trades relative to the prevailing quote midpoint:

$$\text{Percent effective spread}_{\text{it}} = 2\text{D}_{\text{it}}\left(\frac{\text{P}_{\text{it}} - \text{M}_{\text{it}}}{\text{M}_{\text{it}}}\right),$$

where, for stock i, D_{it} is the buy-sell indicator variable which takes a value of 1 (–1) for buyer-initiated (seller-initiated) trades, P_{it} is the transaction price, and M_{it} is the midpoint of the most recently posted bid and ask quotes. The average percent effective spread for each day is a trade-weighted average across all trades during the day. The monthly percent effective spread for each security is calculated by averaging across all trading days within a given month.

47.4 Empirical Analysis

47.4.1 Data

I estimate liquidity measures for NYSE, AMEX, and NASDAQ common stocks over the years 1993–2008. I estimate the low-frequency measures using daily data from the Center for Research in Security Prices (CRSP) database. I estimate the high-frequency measures using the New York Stock Exchange Trades and Automated Quotes (TAQ) database. TAQ data is available only since 1993, which is therefore the binding constraint in terms of staring year. In order to be included in the sample, a stock must have at least 60 days of past return data. I discard certificates, American Depositary Receipts (ADRs), shares of beneficial interest, units, companies incorporated outside United States, American Trust components closed-end funds, preferred stocks, and Real Estate Investment Trusts (REITs).

Regarding estimating the high-frequency measures, I determine the highest bid and lowest ask across all quoting venues at every point during the day (NBBO quotes) and then follow filters referring to Huang and Stoll (1997) and Brownless and Gallo (2006). To reduce errors and outliers, I remove (1) quotes if either the bid or ask price is negative; (2) quotes if either the bid or ask size is negative; (3) quotes if bid-ask spread is greater than \$5 or negative; (4) the quotes if transaction price is negative; (5) quotes before-the-open and after-the-close trades and quotes; (6) quotes if the bid, ask, or trade price differ by more than 20 % from the previous quote or trade price; (7) quotes originating in market other than the primary exchange because regional quotes tend to closely follow the quotes posted by the primary exchange; and (8) %effective spread/%quoted spread>4.0.

47.4.2 Empirical Results

Table 47.1 reports correlations among the various liquidity estimates. In this table, observations are pooled across all stocks and all months. All correlations are reliably positive and substantial in magnitude, ranging from 0.382 to 0.971. The two high-frequency measures, percent effective spread and percent quoted spread, are very highly correlated (0.971). Using percent effective spread as our high-frequency "benchmark," its correlations with the low-frequency measures are 0.742 (Roll), 0.757 (effective tick), 0.621 (Amihud), and 0.586 (LOT). Based on the aforementioned criterion, these results indicate that the effective tick and Roll measures are the "best" low-frequency measures, as they have the highest correlation with percent effective spread.

Table 47.2 presents the time-series means of monthly correlations of percent effective spread with each of the low-frequency measures for the full sample period as well as subperiods 1993–2000 (pre-decimalization) and 2001–2008 (post-decimalization). For three of the four low frequencies measured, the correlation with percent effective spread is higher in the first subperiod than the second subperiod, which may reflect differential effects of decimalization on the various

Table 47.1 Pooled correlations

	ES	QS	Roll	Eff. tick	Ami	LOT
ES	1					
QS	0.971	1				
Roll	0.742	0.744	1			
Eff. tick	0.757	0.760	0.638	1		
Ami	0.621	0.631	0.512	0.472	1	
LOT	0.586	0.562	0.710	0.541	0.382	1

This table presents correlations among the liquidity estimates based on the pooled sample of monthly time-series and cross-sectional observations. Observation can be dropped if there are fewer than 60 days observations for the firm or if liquidity estimates are missing

Table 47.2 Average cross-sectional correlations with percent effective spread, monthly estimates

	Roll	Eff. tick	Ami	LOT
1993–2008	0.662	0.685	0.684	0.561
1993–2000	0.748	0.754	0.662	0.670
2001–2008	0.576	0.615	0.706	0.452

For each month, I estimate the cross-sectional correlation between the liquidity proxies from the low-frequency data and percent effective spread from TAQ. This table presents the average cross-sectional correlations across all months. A stock is excluded only if it trades for less than 60 days prior to an observation or if liquidity estimates are missing

Table 47.3 Summary statistics for stock-by-stock time-series correlations

	N	Roll	Eff. tick	Ami	LOT
Full	13,374	0.319	0.578	0.603	0.308
NYSE	3,137	0.180	0.694	0.690	0.201
AMEX	1,597	0.231	0.517	0.643	0.310
NASDAQ	9,870	0.353	0.534	0.567	0.335

For each stock, I estimate the time-series correlation between the estimated liquidity measure and percent effective spread from TAQ. The table presents the average time-series correlation across all stocks. Observations are dropped if there are fewer than 60 days observations for the firm or if a spread estimate is missing

dimensions of liquidity. For the full period as well as the first subperiod, effective tick has the correlation with percent effective spread, while for the second period the Amihud measure has the highest correlation with percent effective spread.

Table 47.3 shows stock-by-stock time-series correlations between the high-frequency measure percent effective spread and each of the low-frequency measures, using the full-period data but also breaking the observations down by the exchange on which a stock trades. For the full sample as well as every exchange, the Amihud and effective tick estimates have relatively high correlations with percent effective spread, while the correlations are relatively low for the Roll and LOT measures.

Overall, based on the suggested criteria of correlation with a high-frequency measure, the effective tick measure is best among the low-frequency measures tested, as it exhibits higher correlations with percent effective spread based on time-series, cross-sectional, and pooled tests. Again, though, I caution that, since each liquidity measure captures only part of the multidimensional nature of liquidity, it is difficult to judge which measure is best.

47.5 Conclusions

This chapter discusses several popular measures of liquidity that are based on econometric approaches and compares them via correlation analysis. Among the four low-frequency liquidity proxies, I find that the effective tick measure is generally more highly correlated with the high-frequency measure (i.e., percent effective spread). Thus, by this criterion the effective tick measure is the "best" low-frequency measure of liquidity. However, since each liquidity measure captures only part of the multidimensional nature of liquidity, it is difficult to judge which measure is best. Consequently, from among the available measures of liquidity, a researcher should choose the measure that is consistent with their research purpose or perhaps consider employing several of them.

Appendix 1: Solution to LOT (1990) Model

To estimate transaction costs based on their model in Eq. 47.3, Lesmond et al. (1999) introduce the limited dependent variable regression model of Tobin (1958) and Rosett (1959). Tobin's model specifies that data are available for the explanatory variable, x, for all the observation while data are only partly observable for the dependent variable, y, and for the other unobservable region, the information is given whether or not data are above a certain *threshold*.

Considering this aspect of Tobin's model, the limited dependent variable model is an appropriate econometric method for the LOT model because a nonzero observed return occurs only when marginal profit exceeds marginal transaction costs.

Assuming that market model is correct in the presence of transaction costs, Lesmond et al. (1999) estimate transaction costs on the basis of Eqs. 47.1 and 47.3. The equation system is

$$R_{it}^* = \beta_{it} R_{mt} + \varepsilon_{it},$$

where

$$\begin{aligned} R_{it} &= R_{it}^* - \alpha_{1i} && \text{if} \quad R_{it}^* < \alpha_{1i} \\ R_{it} &= 0 && \text{if} \quad \alpha_{1i} < R_{it}^* < \alpha_{2i} \\ R_{it} &= R_{it}^* - \alpha_{2i} && \text{if} \quad R_{it}^* > \alpha_{2i} \end{aligned} \tag{47.4}$$

The solution to this limited dependent regression variable model requires a likelihood function to be maximized with respect to α_{1i}, α_{2i}, β_i, and σ_i.

$$\begin{aligned} L(\alpha_{1i}, \alpha_{2i}, \beta_i, \sigma_i / R_{it}, R_{mt}) = & \prod_1 \text{Ø}\left(\frac{R_{it} + \alpha_{1i} - \beta_i R_{mt}}{\sigma_i}\right) \\ & \prod_2 \left[\Phi\left(\frac{R_{it} + \alpha_{2i} - \beta_i R_{mt}}{\sigma_i}\right) - \Phi\left(\frac{R_{it} + \alpha_{1i} - \beta_i R_{mt}}{\sigma_i}\right)\right] \\ & \prod_3 \text{Ø}\left(\frac{R_{it} + \alpha_{2i} - \beta_i R_{mt}}{\sigma_i}\right), \end{aligned} \tag{47.5}$$

where Ø refers to the standard normal density function and Φ refers to the cumulative normal distribution. The product is over the Region 1, 2, and 3 of observations for which $R^*_{it} < \alpha_{1i}$, $\alpha_{1i} < R^*_{it} < \alpha_{2i}$, and $R^*_{it} > \alpha_{2i}$, respectively. The log likelihood function is

$$\begin{aligned} \log\ L = & \Sigma_1 \log\left[\frac{1}{(2\pi\sigma_i{}^2)\frac{1}{2}}\right] - \frac{1}{2\sigma_i{}^2}\Sigma_1(R_{it} + \alpha_{1i} - \beta_i R_{mt})^2 + \Sigma_2 \log[\Phi_2 - \Phi_1] \\ & + \Sigma_3 \log\left[\frac{1}{(2\pi\sigma_i{}^2)^{\frac{1}{2}}}\right] - \frac{1}{2\sigma_i{}^2}\Sigma_3(R_{it} + \alpha_{2i} - \beta_i R_{mt})^2. \end{aligned} \tag{47.6}$$

Given Eq. 47.6, α_{1i}, α_{2i}, β_i, and σ_i can be estimated. The difference between α_{2i} and α_{1i} is the proxy of a round-trip transaction cost in the LOT model.

References

Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets, 5*, 31–56.

Amihud, Y., & Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics, 17*(2), 223–249.

Bharath, S. T., Pasquariello, P., et al. (2009). Does asymmetric information drive capital structure decisions? *Review of Financial Studies, 22*(8), 3211–3243.

Brennan, M. J., & Subrahmanyam, A. (1996). Market microstructure and asset pricing: on the compensation for illiquidity in stock returns. *Journal of Financial Economics, 41*, 441–464.

Brennan, M., Chordia, T., et al. (1998). Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. *Journal of Financial Economics, 49*(3), 345–373.

Brownlees, C. T., & Gallo, G. M. (2006). Financial econometric analysis at ultra-high frequency: Data handling concerns. *Computational Statistics & Data Analysis, 51*, 2232–2245.

Chung, K. H., Elder, J., & Kim, J. (2010). Corporate governance and liquidity. *Journal of Financial and Quantitative Analysis, 45*(02), 265–291.

Datar, V. T., Naik, N. Y., & Radcliffe, R. (1998). Liquidity and stock returns: An alternative test. *Journal of Financial Markets, 1*, 203–219.

Goyenko, R. Y., Holden, C. W., et al. (2009). Do liquidity measures measure liquidity? *Journal of Financial Economics, 92*(2), 153–181.

Hasbrouck, J. (2009). Trading costs and returns for US equities: Estimating effective costs from daily data. *Journal of Finance, 64*(3), 1445–1477.

Holden, C. W. (2009). New low-frequency spread measures. *Journal of Financial Markets, 12*(4), 778–813.

Huang, R. D., & Stoll, H. R. (1997). The components of the bid-ask spread: A general approach. *Review of Financial Studies, 10*(4), 995–1034.

Lesmond, D. A., Ogden, J. P., et al. (1999). A new estimate of transaction costs. *Review of Financial Studies, 12*(5), 1113–1141.

Lipson, M. L. (2003). Market microstructure and corporate finance. *Journal of Corporate Finance, 9*(4), 377–384.

Lipson, M. L., & Mortal, S. (2007). Liquidity and firm characteristics: Evidence from mergers and acquisitions. *Journal of Financial Markets, 10*(4), 342–361.

Lipson, M. L., & Mortal, S. (2009). Liquidity and capital structure. *Journal of Financial Markets, 12*(4), 611–644.

Liu, W. (2006). A liquidity-augmented capital asset pricing model. *Journal of Financial Economics, 82*, 631–671.

Maddala, G. S. (1986). *Limited-dependent and qualitative variables in econometrics*. Cambridge: Cambridge University Press.

Pástor, L., & Stambaugh, R. (2003). Liquidity risk and expected stock returns. *Journal of Political Economy, 111*, 642–685.

Roll, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *Journal of Finance, 39*, 1127–1139.

Rosett, R. N. (1959). A statistical model of friction in economics. *Econometrica: Journal of the Econometric Society, 27*, 263–267.

Tobin, J. (1958). Estimation of relationships for limited dependent variables. *Econometrica: Journal of the Econometric Society, 26*, 24–36.

A Quasi-Maximum Likelihood Estimation Strategy for Value-at-Risk Forecasting: Application to Equity Index Futures Markets

48

Oscar Carchano, Young Shin (Aaron) Kim, Edward W. Sun, Svetlozar T. Rachev, and Frank J. Fabozzi

Contents

O. Carchano
Department of Financial Economics, University of Valencia, Valencia, Spain
e-mail: oscar.carchano@uv.es

Y.S.A. Kim
College of Business, Stony Brook University, Stony Brook, NY, USA
e-mail: aaron.kim@stonybrook.edu

E.W. Sun
KEDGE Business School and BEM Management School, Bordeaux, France
e-mail: edward.sun@kedgebs.com

S.T. Rachev
Department of Applied Mathematics and Statistics, College of Business, Stony Brook University, SUNY, Stony Brook, NY, USA

FinAnalytica, Inc, New York, NY, USA
e-mail: svetlozar.rachev@stonybrook.edu

F.J. Fabozzi (✉)
EDHEC Business School, EDHEC Risk Institute, Nice, France
e-mail: frank.fabozzi@edhec.edu; fabozzi321@aol.com

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_48,

Abstract

We present the first empirical evidence for the validity of the ARMA-GARCH model with tempered stable innovations to estimate 1-day-ahead value at risk in futures markets for the S&P 500, DAX, and Nikkei. We also provide empirical support that GARCH models based on normal innovations appear not to be as well suited as infinitely divisible models for predicting financial crashes. The results are compared with the predictions based on data in the cash market. We also provide the first empirical evidence on how adding trading volume to the GARCH model improves its forecasting ability.

In our empirical analysis, we forecast 1 % value at risk in both spot and futures markets using normal and tempered stable GARCH models following a quasi-maximum likelihood estimation strategy. In order to determine the accuracy of forecasting for each specific model, backtesting using Kupiec's proportion of failures test is applied. For each market, the model with a lower number of violations is preferred. Our empirical result indicates the usefulness of classical tempered stable distributions for market risk management and asset pricing.

Keywords

Infinitely divisible models • Tempered stable distribution • GARCH models • Value at risk • Kupiec's proportion of failures test • Quasi-maximum likelihood estimation strategy

48.1 Introduction

Predicting future financial market volatility is crucial for risk management of financial institutions. The empirical evidence suggests that a suitable market risk model must be capable of handling the idiosyncratic features of volatility, that is, daily returns time variant amplitude and volatility clustering. There is a well-developed literature in financial econometrics that demonstrates how autoregressive conditional heteroskedastic (ARCH) and generalized ARCH (GARCH) models – developed by Engle (1982) and Bollerslev (1986), respectively – can be employed to explain the clustering effect of volatility[1]. Moreover, the selected model should consider the stylized fact that asset return distributions are not normally distributed, but instead have been shown to exhibit patterns of leptokurtosis and skewness.

[1]For a description of ARCH and GARCH modeling, see Chap. 8 in Rachev et al. (2007). The chapter of the same reference describes ARCH and GARCH modeling with infinite variance innovations. Engle et al. (2008) provide the basics of ARCH and GARCH modeling with applications to finance.

Taking a different tact than the ARCH/GARCH with normal innovations approach for dealing with the idiosyncratic features of volatility, Kim et al. (2010) formulate an alternative model based on subclasses of the infinitely divisible (ID) distributions. More specifically, for the S&P 500 return, they empirically investigate five subclasses of the ID distribution, comparing their results to that obtained using GARCH models based on innovations that are assumed to follow a normal distribution. They conclude that, due to their failure to focus on the distribution in the tails, GARCH models based on the normal innovations may not be as well suited as ID models for predicting financial crashes.

Because of its popularity, most empirical studies have examined value at risk (VaR) as a risk measure. These studies have focused on stock indices. For example, Kim et al. (2011), and Asai and McAleer (2009) examine the S&P 500, DAX 30, and Nikkei 225 stock indices, respectively. A few researchers have studied this risk measure for stock index futures contracts: Huang and Lin (2004) (Taiwan stock index futures) and Tang and Shieh (2006) (S&P 500, Nasdaq 100, and Dow Jones stock index futures). As far as we know, there are no empirical studies comparing VaR spot and futures indices. For this reason, we compare the predictive performance of 1-day-ahead VaR forecasts in these two markets.

We then introduce trading volume into the model, particularly, within the GARCH framework. There are several studies that relate trading volume and market volatility for equities and equity futures markets. Studies by Epps and Epps (1976), Smirlock and Starks (1985), and Schwert (1989) document a positive relation between volume and market volatility. Evidence that supports the same relation for futures is provided by Clark (1973), Tauchen and Pitts (1983), Garcia et al. (1986), Ragunathan and Peker (1997), and Gwilym et al. (1999). Collectively, these studies clearly support the theoretical prediction of a positive and contemporaneous relationship between trading volume and volatility. This result is a common empirical finding for most financial assets, as Karpoff (1987) showed when he summarized the results of several studies on the positive relation between price changes and trading volume for commodity futures, currency futures, common stocks, and stock indices.

Foster (1995) concluded that not only is trading volume important in determining the rate of information (i.e., any news that affects the market), but also lagged volume plays a role. Although contemporary trading volume is positively related to volatility, lagged trading volume presents a negative relationship. Empirically, investigating daily data for several indices such as the S&P 500 futures contract, Wang and Yau (2000) observe that there is indeed a negative link between lagged trading volume and intraday price volatility. This means that an increase in trading volume today (as a measure of liquidity) will imply a reduction in price volatility tomorrow. In their study of five currency futures contracts, Fung and Patterson (1999) do in fact find a negative relationship between return volatility and past trading volume. In their view, the reversal behavior of volatility with trading volume is generally consistent with the overreaction hypothesis (see Conrad et al. 1994) and supports the sequential information hypothesis (see Copeland 1976), which explains the relationship between return volatility and trading volume.

Despite the considerable amount of research in this area, there are no studies that use trading volume in an effort to improve the capability of models to forecast 1-day-ahead VaR. Typically, in a VaR context, trading volume is only employed as a proxy for "liquidity risk" – the risk associated with trying to close out a position. In this paper, in contrast to prior studies, we analyze the impact of introducing trading volume on the ability to enhance performance in forecasting VaR 1 day ahead. We empirically test whether the introduction of trading volume will reduce the number of violations (i.e., the number of times when the observed loss exceeds the estimated one) in the spot and futures equity markets in the USA, Germany, and Japan.

The remainder of this paper is organized as follows. ARMA-GARCH models with normal and tempered stable innovations are reviewed in Sect. 48.2. In Sect. 48.3, we discuss parameter estimation of the ARMA-GARCH models and forecasting daily return distributions. VaR values and backtesting of the ARMA-GARCH models are also reported in Sect. 48.2, along with a comparison of the results for (1) the spot and futures markets and (2) the normal and tempered stable innovations. Trading volume is introduced into the ARMA-GARCH model with tempered stable innovations in Sect. 48.4. VaR and backtesting of the ARMA-GARCH with different variants of trading volume are presented and compared to the results for models with and without trading volume. We summarize our principal findings in Sect. 48.5.

48.2 ARMA-GARCH Model with Normal and Tempered Stable Innovations

In this section, we provide a review of the ARMA-GARCH models with normal and tempered stable innovations. For a more detailed discussion, see Kim et al. (2011).

Let $(S_t)_{t\geq 0}$ be the asset price process and $(y_t)_{t\geq 0}$ be the return process of $(S_t)_{t\geq 0}$ defined by $y_t = \log \frac{S_t}{S_t - 1}$. The ARMA(1,1)-GARCH(1,1) model is

$$\begin{cases} y_t = ay_{t-1} + b\sigma_{t-1}\varepsilon_{t-1} + \sigma_t\varepsilon_t + c_t \\ \sigma_t^2 = \alpha_0 + \alpha_1\sigma^2{}_{t-1}\varepsilon^2{}_{t-1} + \beta_1\sigma^2{}_{t-1} \end{cases}. \tag{48.1}$$

where $\varepsilon_0 = 0$ and a sequence $(\varepsilon_t)_{tcN} = 0$ of independent and identically distributed (iid) real random variables. The innovation ε_t is assumed to follow the standard normal distribution. This ARMA(1,1)-GARCH(1,1) model is referred to as the "normal-ARMA-GARCH model."

If the ε_ts are assumed to be tempered stable innovations, then we obtain a new ARMA(1,1)-GARCH(1,1) model. In this paper, we will consider the standard classical tempered stable (denoted by stdCTS) distributions. This ARMA(1,1)-GARCH(1,1) model is defined as follows: CTS-ARMA-GARCH model, $\varepsilon_t \sim stdCTS(\alpha, \lambda_+, \lambda_-)$. This distribution does not have a closed-form solution for its probability density function. Instead, it is defined by its characteristic function as follows: Let $\alpha \in (0,2)\setminus\{1\}, C, \lambda_+, \lambda_- > 0$, and $m \in \mathbb{R}$. Then a random variable X is said to follow the classical tempered stable (CTS) distribution if the characteristic function of X is given by

$$\begin{aligned}\phi_x(u) &= \phi_{CTS}\left(u : \alpha,\ C, \lambda_+, \lambda_-, m\right)\\ &= \exp\left(ium - iuCT(1-\alpha)\left(\lambda_+^{\alpha-1} - \lambda_-^{\alpha-1}\right)\right.\\ &\quad \left. + C\Gamma(-\alpha)\left((\lambda_+ - iu)^\alpha - \lambda_+^\alpha + (\lambda_- - iu)^\alpha - \lambda_-^\alpha\right)\right),\end{aligned} \tag{48.2}$$

and we denote $X \sim CTS(\alpha, C, \lambda_+, \lambda_-, m)$.

The cumulants of X are defined by

$$C_n(X) = \frac{1}{i^n}\frac{\partial^n}{\partial u^n}\ \log E\left[e^{iuX}\right]|u = 0,\ \ n = 1, 2, 3, \ldots.$$

For the tempered stable distribution, we have $E[X] = c_1(X) = m$. The cumulants of the tempered stable distribution for $n = 2, 3, \ldots$ are

$$c_n(X) = C\Gamma(n-\alpha)\left(\lambda_+^{\alpha-n} + (-1)^n \lambda_-^{a-n}\right).$$

By substituting the appropriate value for the two parameters m and C into the three tempered stable distributions, we can obtain tempered stable distributions with zero mean and unit variance. That is, $X \sim CTS(\alpha, C, \lambda_+, \lambda_-, 0)$ has zero mean and unit variance by substituting

$$C = \left(\Gamma(2-\alpha)\left(\lambda_+^{\alpha-2} + \lambda_-^{\alpha-2}\right)\right)^{-1}. \tag{48.3}$$

The random variable X is referred to as the standard CTS distribution with parameters $(\alpha, \lambda_+, \lambda_-)$ and denoted by $X \sim stdCTS(\alpha, \lambda_+, \lambda_-)$.

48.3 VaR for the ARMA-GARCH Model

In this section, we discuss VaR for the ARMA-GARCH model with normal and tempered stable innovations.

48.3.1 VaR and Backtesting

The definition of VaR for a significance level η is

$$\text{VaR}_\eta(X) = -\inf\{x \in \mathbb{R} | P(X \le x) > \eta\}.$$

If we take the ARMA-GARCH model described in Sect. 48.2, we can define VaR for the information until time t with significance level η as[2]

$$\mathrm{VaR}_{t,\eta}(y_{t+1}) = -\inf\{x \in \mathbb{R} | P_t(y_{t+1} \le x) > \eta\},$$

where $P_t(A)$ is the conditional probability of a given event A for the information until time t.

Two models are considered: normal-ARMA(1,1)-GARCH(1,1) and stdCTS-ARMA(1,1)-GARCH(1,1). For both models, the parameters have been estimated for the time series between December 14, 2004 and December 31, 2008. For each daily estimation, we worked with 10 years of historical daily performance for the S&P 500, DAX 30, and Nikkei 225 spot and futures indices. More specifically, we used daily returns calculated based on the closing price of those indices. In the case of futures indices, we constructed a unique continuous time series using the different maturities of each futures index following the methodology proposed by Carchano and Pardo (2009).[3] Then, we computed VaRs for both models.

The maximum likelihood estimation method (MLE) is employed to estimate parameters of the normal-ARMA(1,1)-GARCH(1,1) model. For the CTS distribution, the parameters are estimated as follows[4]:

1. Estimate parameters α_0, α_1, β_1, α, b, c with normal innovations by the MLE. Volatility clustering is captured by the GARCH model.
2. Extract residuals using those parameters. The residual distribution still presents fat tail and skewness.
3. Fit the parameters of the innovation distribution (CTS) to the extracted residuals using MLE. The fat tailed and skewed features of the residual distribution are captured.

In order to determine the accuracy of VaR for the two models, backtesting using *Kupiec's proportion of failures test* (Kupiec 1995) is applied. We first calculate the number of violations. Then, we compare the number of violations with the conventional number of exceedances at a given significance level. In Table 48.1 the number of violations and p-values for Kupiec's backtest for the three stock indices over the 41-year periods are reported. Finally, we sum up the number of violations and their related p-values for 1 % VaRs for the normal and CTS-ARMA-GARCH models.

[2]VaR on the CTS distribution is described in the Appendix.

[3]Thus, the last trading day of the front contract is chosen as the rollover date. Then, the return of the day after the rollover date is calculated as the quotient between the closing price of the following contract and the previous closing price of such contract. By doing so, all the returns are taken from the same maturity.

[4]A quasi-MLE strategy is followed because the ARMA-GARCH CTS model has too many parameters. If all the parameters are estimated at once, then the GARCH parameters go to zero. This strategy is also followed in Kim et al. (2009, 2010, 2011). For a discussion of the quasi- MLE methodology, see Rachev et al. (2007, pp. 292–293) or Verbeek (2004, pp. 182–184).

Table 48.1 Normal-ARMA-GARCH versus CTS-ARMA-GARCH

	1 year (255 days)			
	Dec. 14, 2004	Dec. 16, 2005	Dec. 21, 2006	Dec. 28, 2007
	~ Dec. 15, 2005	~ Dec. 20, 2006	~ Dec. 27, 2007	~ Dec. 31, 2008
	N(*p*-value)	*N*(*p*-value)	*N*(*p*-value)	*N*(*p*-value)
Model	S&P 500 spot			
Normal-ARMA-GARCH	1(0.2660)	3(0.7829)	8(0.0061)	10(0.0004)
CTS-ARMA-GARCH	0	2(0.7190)	6(0.0646)	4(0.3995)
	S&P 500 futures			
Normal-ARMA-GARCH	3(0.7829)	3(0.7829)	7(0.0211)	9(0.0016)
CTS-ARMA-GARCH	1(0.2660)	3(0.7829)	4(0.3995)	5(0.1729)
	DAX 30 spot			
Normal-ARMA-GARCH	4(0.3995)	4(0.3995)	3(0.7829)	6(0.0646)
CTS-ARMA-GARCH	4(0.3995)	4(0.3995)	3(0.7829)	4(0.3995)
	DAX 30 futures			
Normal-ARMA-GARCH	3(0.7829)	5(0.1729)	6(0.0646)	6(0.0646)
CTS-ARMA-GARCH	3(0.7829)	4(0.3995)	6(0.0646)	3(0.7829)
	Nikkei 225 spot			
Normal-ARMA-GARCH	2(0.7190)	4(0.3995)	5(0.1729)	5(0.1729)
CTS-ARMA-GARCH	1(0.2660)	3(0.7829)	4(0.3995)	5(0.1729)
	Nikkei 225 futures			
Normal-ARMA-GARCH	2(0.7190)	2(0.7190)	7(0.0211)	5(0.1729)
CTS-ARMA-GARCH	5(0.1729)	5(0.1729)	6(0.0646)	6(0.0646)

The number of violations (*N*) and *p*-values of Kupiec's proportion of failures test for the S&P 500, DAX 30, and Nikkei 225 spot and futures indices data has been shown. Normal-ARMA-GARCH and CTS-ARMA-GARCH compared

Based on Table 48.1, we conclude the following for the three stock indices. First, a comparison of the normal and tempered stable models indicates that there are no cases using the tempered stable model at the 5 % significance level, whereas the normal model is rejected five times. This evidence is consistent with the findings of Kim et al. (2011). Second, a comparison of the spot and futures indices indicates that spot data provide less than or the same number of violations than futures data. One potential explanation is that futures markets are more volatile, particularly, when the market falls.[5] This overreaction to bad news could cause the larger number of violations.

[5]We compared the spot and futures series when the markets discount bad news (negative returns). We find that for the three stock indices, futures volatility is significantly greater than spot volatility at a 5 % significance level. Moreover, for all three stock indices, the minimum return and the 1 % percentile return are also lower for futures data than spot data.

48.4 Introduction of Trading Volume

In the previous section, we showed the usefulness of the tempered stable model for stock index futures. Motivated by the vast literature linking trading volume and volatility, for the first time we investigate whether the introduction of trading volume in the CTS model could improve its ability to forecast 1-day-ahead VaR.

Let $(St)_{t\geq 0}$ be the asset price process and $(y_t)_{t\geq 0}$ be the return process of $(S_t)_{t\geq 0}$ defined by $y_t = \log \frac{S_t}{S_t - 1}$. We propose the following ARMA(1,1)-GARCH(1,1) with trading volume model:

$$\begin{cases} y_t = ay_{t-1} + b\sigma_{t-1}\varepsilon_{t-1} + \sigma_t\varepsilon_t + c \\ \sigma_t^2 = \alpha_0 + \alpha_1{\sigma^2}_{t-1}{\varepsilon^2}_{t-1} + \beta_1{\sigma^2}_{t-1} + \gamma_1 Vol_{t-1}, \end{cases} \tag{48.4}$$

where $\varepsilon_0 = 0$ and a sequence $(\varepsilon_t)_{tcN} = 0$ of *iid* real random variables. The innovation ε_t is assumed to be the tempered stable innovation. We will consider the standard classical tempered stable distributions. This new ARMA(1,1)-GARCH(1,1)-V model is defined as follows:

$$\text{CTS-ARMA-GARCH-V model} : \varepsilon_t \sim stdCTS(\alpha, \lambda_+, \lambda_-).$$

The inclusion of lagged volume as an independent variable along with lagged volatility into the model may cause a problem of multicollinearity. In order to determine the seriousness of the problem, we calculated the model without volume, extracted the GARCH series, and determined the degree of collinearity between both variables. The most recommended measure in the literature is to calculate the condition index following Belsley et al. (1980) and observe if the index exceeds 20, in which case collinearity is considered to be grave. In our case, the calculated value was 4.9268, 3.2589, and 4.5569 for the S&P, DAX, and Nikkei, respectively. Therefore, we concluded that collinearity is a minor problem.

Moreover, the ARMA-GARCH model is only affected in the GARCH framework, particularly the equation coefficients (i.e., the volume variable can appear insignificant when it is indeed significant), but not the numerical estimation of the variance; neither is the forecast power of the global model. As our objective is to forecast the VaR, we believe that the multicollinearity problem can be ignored because the results will not be affected.

48.4.1 Different Variants of Trading Volume

For the S&P 500 cash and futures markets, we test the following versions of trading volume in order to determine which one would be the most appropriate:

- Lagged trading volume in levels: V(t − 1)
- Logarithm of lagged trading volume: log [V(t − 1)]
- Relative change of lagged trading volume: log [V(t − 1)/V(t − 2)]

Table 48.2 CTS-ARMA-GARCH with lagged volume

	1 year (255 days)			
	Dec. 14, 2004	Dec. 16, 2005	Dec. 21, 2006	Dec. 28, 2007
	∼ Dec. 15, 2005	∼ Dec. 20, 2006	∼ Dec. 27, 2007	∼ Dec. 31, 2008
	N(p-value)	N(p-value)	N(p-value)	N(p-value)
Model	S&P 500 spot			
CTS-ARMA-GARCH-V(t − 1)	16(0.0000)	10(0.0004)	23(0.0000)	26(0.0000)
CTS-ARMA-GARCH-log [V(t − 1)]	1(0.2660)	4(0.3995)	10(0.0004)	6(0.0646)
CTS-ARMA-GARCH-ln [V(t − 1)/V(t − 2)]	0	2(0.7190)	6(0.0646)	5(0.1729)
	S&P 500 futures			
CTS-ARMA-GARCH-V(t − 1)	1(0.2660)	11(0.0001)	4(0.3995)	6(0.0646)
CTS-ARMA-GARCH-log [V(t − 1)]	1(0.2660)	3(0.7829)	4(0.3995)	5(0.1729)
CTS-ARMA-GARCH-ln [V(t − 1)/V(t − 2)]	1(0.2660)	3(0.7829)	3(0.7829)	5(0.1729)
CTS-ARMA-GARCH-V$($t$ − 1)	0	1(0.2660)	15(0.0000)	3(0.7829)
CTS-ARMA-GARCH-log [V$($t$ − 1)]	3(0.7829)	3(0.7829)	4(0.3995)	5(0.1729)
CTS-ARMA-GARCH-ln [V$($t$ − 1)/V$($t$ − 2)]	2(0.7190)	0	8(0.0061)	8(0.0061)

The number of violations (N) and p-values of Kupiec's proportion of failures test for the S&P 500 spot and futures indices with the different variants of volume using the stdCTS-ARMA(1,1)-GARCH(1,1) model has been shown. V(t − 1), log [V(t − 1)], and ln [V(t − 1)/V(t − 2)] stand for levels, logarithm, and relative change of the lagged trading volume, respectively. V$($t$ − 1), log [V$($t$ − 1)], and ln [V$($t$ − 1)/V$($t$ − 2)] stand for levels, logarithm, and relative change of the lagged trading volume in dollars, respectively

The spot series trading volume is in dollars; for the futures series, the trading value is in number of contracts. We can calculate the volume of the futures market in dollars too. The tick value of the S&P 500 futures contract is 0.1 index points or $25. Multiplying the number of contracts by the price and finally by $250 (the contract's multiple), we obtain the trading volume series for the futures contract in dollars. Thus, for the futures contract we get three new versions of trading volume to test:

- Lagged trading volume in dollars: V$($t$ − 1)
- Logarithm of trading volume in dollars: log [V$($t$ − 1)]
- Relative change of lagged trading volume in dollars: log [V$($t$ − 1)/V$($t$ − 2)]

By doing that, we can determine which series (in dollars or in contracts) seems to be more useful for the futures index.

In Table 48.2 we report the number of violations and p-values of Kupiec's backtest for the different versions of the CTS-ARMA-GARCH-V model for the S&P 500 spot and futures indices. We count the number of violations and the corresponding *p*-values for 1 %-VaRs of both markets. From Table 48.2, we conclude the following:

- The model with the lagged trading volume in level is rejected at the 1 % significance level in all 4 years for the S&P 500 spot, and for the second period (2005–2006) for the S&P 500 futures.
- The logarithm of trading volume in the model is rejected at the 5 % significance level for the spot market for the third period (2006–2007), but it is not rejected in any period for the futures market.
- The relative change of the lagged volume is not rejected at the 5 % significance level in any period in either market. Of the three versions of trading volume tests, this version seems to be the most useful for both spot and futures markets.
- The results for trading volume in contracts and the trading volume in dollars in the futures market indicate that the former is rejected at the 1 % significance level only for the lagged trading volume in level in the second period (2005–2006). Trading volume in dollars is rejected three times, for the lagged trading volume in levels for the third period (2006–2007) and for the lagged relative trading volume change in the last two periods (2006–2007 and 2007–2008). These findings suggest that the trading volume in contracts is the preferred measure.

48.4.2 Lagged Relative Change of Trading Volume

As we have just seen, the variant of trading volume that seems more useful for forecasting 1-day-ahead VaR using CTS-ARMA-GARCH is the relative change of trading volume. Next, we compare the original CTS-ARMA-GARCH model with the new CTS-ARMA-GARCH-V model where V is the lagged relative change of trading volume. Table 48.3 shows the number of violations and *p*-values of Kupiec's backtest for the two models for the three stock indices and both markets. We sum up the number of violations and the corresponding *p*-values for 1 % VaRs for each case.

Our conclusions from Table 48.3 are as follows. For the spot markets, the introduction of trading volume does not mean a reduction in the number of violations in any period for any index. However, for the futures markets, the numbers of violations are the same or lower for the model with trading volume than with the original model. Thus, by introducing trading volume, we get a slightly more conservative model, increasing the VaR forecasted for futures equity markets.

48.4.3 Lagged Trading Volume or Forecasting Contemporaneous Trading Volume

Although there is some evidence which supports the relationship between lagged trading volume and volatility, the literature is not as extensive as the studies that establish a strong link between volatility and contemporaneous trading

Table 48.3 CTS-ARMA-GARCH versus CTS-ARMA-GARCH-V

	1 year (255 days)			
	Dec. 14, 2004	Dec. 16, 2005	Dec. 21, 2006	Dec. 28, 2007
	~ Dec. 15, 2005	~ Dec. 20, 2006	~ Dec. 27, 2007	~ Dec. 31, 2008
	N(p-value)	N(p-value)	N(p-value)	N(p-value)
Model	S&P 500 spot			
CTS-ARMA-GARCH	0	2(0.7190)	6(0.0646)	4(0.3995)
CTS-ARMA-GARCH-ln [V(t − 1)/V(t − 2)]	0	2(0.7190)	6(0.0646)	5(0.1729)
	S&P 500 futures			
CTS-ARMA-GARCH	1(0.2660)	3(0.7829)	4(0.3995)	5(0.1729)
CTS-ARMA-GARCH-ln [V(t − 1)/V(t − 2)]	1(0.2660)	3(0.7829)	3(0.7829)	5(0.1729)
	DAX 30 spot			
CTS-ARMA-GARCH	4(0.3995)	4(0.3995)	3(0.7829)	4(0.3995)
CTS-ARMA-GARCH-ln [V(t − 1)/V(t − 2)]	5(0.1729)	4(0.3995)	3(0.7829)	5(0.1729)
	DAX 30 futures			
CTS-ARMA-GARCH	3(0.7829)	4(0.3995)	6(0.0646)	3(0.7829)
CTS-ARMA-GARCH-ln [V(t − 1)/V(t − 2)]	3(0.7829)	4(0.3995)	5(0.1729)	3(0.7829)
	Nikkei 225 spot			
CTS-ARMA-GARCH	1(0.2660)	3(0.7829)	4(0.3995)	5(0.1729)
CTS-ARMA-GARCH-ln [V(t − 1)/V(t − 2)]	3(0.7829)	3(0.7829)	4(0.3995)	5(0.1729)
	Nikkei 225 futures			
CTS-ARMA-GARCH	5(0.1729)	5(0.1729)	6(0.0646)	6(0.0646)
CTS-ARMA-GARCH-ln [V(t − 1)/V(t − 2)]	4(0.3995)	4(0.3995)	6(0.0646)	6(0.0646)

The number of violations (N) and p-values of Kupiec's proportion of failures test for the S&P 500, DAX 30, and Nikkei 225 spot and futures indices has been reported. CTS-ARMA-GARCH and CTS-ARMA-GARCH with lagged relative change of trading volume compared

volume. As there are countless ways to try to forecast trading volume, we begin by introducing contemporaneous trading volume relative change in the model as a benchmark to assess whether it is worthwhile to forecast trading volume.

In Table 48.4 we show the number of violations and p-values of Kupiec's backtest for the CTS-ARMA-GARCH with contemporaneous and lagged relative change of trading volume for the three stock indices for both markets. We count the number of violations and the corresponding p-values for 1 % VaRs for the six indices.

Our conclusions based on the results reported in Table 48.4 are as follows. First, with the exception of the S&P 500 futures, the introduction of the contemporaneous relative change of trading volume in the model is rejected at the 1 % significance level for the last period analyzed (2007–2008). In the case of the S&P

Table 48.4 CTS-ARMA-GARCH with contemporaneous volume versus CTS-ARMA-GARCH with lagged volume

	1 year (255 days)			
	Dec. 14, 2004	Dec. 16, 2005	Dec. 21, 2006	Dec. 28, 2007
	~ Dec. 15, 2005	~ Dec. 20, 2006	~ Dec. 27, 2007	~ Dec. 31, 2008
	N(p-value)	N(p-value)	N(p-value)	N(p-value)
Model	S&P 500 spot			
CTS-ARMA-GARCH-ln [V(t)/V($t-1$)]	0	3(0.7829)	3(0.7829)	8(0.0061)
CTS-ARMA-GARCH-ln [V($t-1$)/V($t-2$)]	0	5(0.1729)	6(0.0646)	5(0.1729)
	S&P 500 futures			
CTS-ARMA-GARCH-ln [V(t)/V($t-1$)]	0	1(0.2660)	7(0.0211)	6(0.0646)
CTS-ARMA-GARCH-ln [V($t-1$)/V($t-2$)]	1(0.2660)	3(0.7829)	3(0.7829)	5(0.1729)
	DAX 30 spot			
CTS-ARMA-GARCH-ln [V(t)/V($t-1$)]	0	1(0.2660)	3(0.7829)	11(0.0001)
CTS-ARMA-GARCH-ln [V($t-1$)/V($t-2$)]	5(0.1729)	4(0.3995)	3(0.7829)	5(0.1729)
	DAX 30 futures			
CTS-ARMA-GARCH-ln [V(t)/V($t-1$)]	1(0.2660)	1(0.2660)	2(0.7190)	8(0.0061)
CTS-ARMA-GARCH-ln [V($t-1$)/V($t-2$)]	3(0.7829)	4(0.3995)	5(0.1729)	3(0.7829)
	Nikkei 225 spot			
CTS-ARMA-GARCH-ln [V(t)/V($t-1$)]	3(0.7829)	5(0.1729)	7(0.0211)	8(0.0061)
CTS-ARMA-GARCH-ln [V($t-1$)/V($t-2$)]	3(0.7829)	3(0.7829)	4(0.3995)	5(0.1729)
	Nikkei 225 futures			
CTS-ARMA-GARCH-ln [V(t)/V($t-1$)]	1(0.2660)	1(0.2660)	3(0.7829)	11(0.0001)
CTS-ARMA-GARCH-ln [V($t-1$)/V($t-2$)]	4(0.3995)	4(0.3995)	6(0.0646)	6(0.0646)

The number of violations (N) and p-values of Kupiec's proportion of failures test for the S&P 500, DAX 30, and Nikkei 225 spot and futures indices has been shown. CTS-ARMA-GARCH with the relative change of trading volume and CTS-ARMA-GARCH with lagged relative change of trading volume compared

500 futures, it is rejected at the significance level of 5 % for the third period (2006–2007). Second, the model with lagged relative change of trading volume is not rejected for any stock index or market. It seems to be more robust than contemporaneous trading volume (although, in general, there are fewer violations when using it).

Our results suggest that it is not worth making an effort to predict contemporaneous trading volume because the forecasts will be flawed and two variables would have to be predicted (VaR and contemporaneous trading volume). Equivalently, the lagged trading volume relative change appears to be more robust because it is not rejected in any case, although it provides a poor improvement to the model.

48.5 Conclusions

Based on an empirical analysis of spot and futures trading for the S&P 500, DAX 30, and Nikkei 225 stock indices, in this paper we provide empirical evidence about the usefulness of using classical tempered stable distributions for predicting 1-day-ahead VaR. Unlike prior studies that investigated CTS models in the cash equity markets, we analyzed their suitability for both spot markets and futures markets. We find in both markets the CTS models perform better in forecasting 1-day-ahead VaR than models that assume innovations follow the normal law.

Second, we introduced trading volume into the CTS model. Our empirical evidence suggests that lagged trading volume relative change provides a slightly more conservative model (i.e., reduces the number of violations) to predict 1-day-ahead VaR for stock index futures contracts. We cannot state the same for the cash market because the results are mixed depending on the index. After that, we introduced contemporaneous trading volume to try to improve the forecasting ability of the model, but in the end, it did not seem to be worth the effort. That is, trading volume appeared not to offer enough information to improve forecasts.

Finally, we compared the number of violations of the estimated VaR in the spot and futures equity markets. For the CTS model without volume, in general, we find fewer violations in the spot indices than in the equivalent futures contracts. In contrast, our results suggest that the number of violations in futures markets is less in the case of the CTS model with trading volume in comparison to the CTS model that ignores trading volume. But if we contrast spot and futures equity markets, violations are still greater for futures than in spot markets. A possible reason is that futures markets demonstrate extra volatility or an overreaction when the market falls with respect to their corresponding spot markets.

Appendix: VaR on the CTS Random Variable

Let X be a CTS random variable. Since the CTS random variable is continuous and infinitely divisible, we obtain $\mathrm{VaR}_\eta(X) = -F_X(\eta)$, where the cumulative distribution function F_X of X is provided by the following proposition.

Proposition Let X be an infinitely divisible random variable and $\phi_x(u)$ be the characteristic function of X. If there is a $\rho > 0$ such that $|\phi_x(Z)| < \infty$ for all the complex z with $\Im(z) = \rho$, then

$$F_X(x) = \frac{e^{x\rho}}{\pi} \Re\left(\int_0^{\infty} e^{-ixu} \frac{\phi_X(u+i\rho)}{\rho - ui} du\right), \text{for } x \in \mathbb{R} \tag{48.5}$$

where $\Re(z)$ is the real part of a complex number z.

Proof By the definition of the cumulative density function, we have

$$F_X(x) = P(X \leq x) = \int_{-\infty}^{x} fX(t)dt,$$

where $f_X(x)$ is the density function of X. The probability density function $f_X(t)$ can be obtained from the characteristic function ϕ_X by the complex inverse formula (see Doetsch 1970); that is,

$$fX(t) = \frac{1}{2\pi} \int_{-\infty+i\rho}^{\infty+i\rho} e^{-itz} \phi_X(z)dz,$$

and we have

$$\begin{aligned} F_X(x) &= \int_{-\infty}^{x} \frac{1}{2\pi} \int_{-\infty+ip}^{\infty+ip} e^{-itz} \phi_X(z)dzdt \\ &= \frac{1}{2\pi} \int_{-\infty+ip}^{\infty+ip} \int_{-\infty}^{x} e^{-itz} dt\phi_X(z)dz. \end{aligned}$$

Note that if $\rho > 0$, then

$$\lim_{t\to-\infty} \left|e^{-it(a+i\rho)}\right| = \lim_{t\to\infty} \left|e^{it(a+i\rho)}\right| = \lim_{t\to\infty} e^{-\rho t} = 0, \quad a \in \mathbb{R},$$

and hence

$$\int_{-\infty}^{x} e^{-itz} dt = -\frac{1}{iz}\left[e^{-itz}\right]_{-\infty}^{x} = -\frac{1}{iz} e^{-ixz}$$

where $z \in \mathbb{C}$ with $\Im(z) = \rho$ Thus, we have

$$\begin{aligned} F_X(x) &= -\frac{1}{2\pi} \int_{-\infty+ip}^{\infty+ip} \frac{1}{iz} e^{-ixz} \phi_X(z)dz \\ &= -\frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{1}{i(u+i\rho)} e^{-ix(u+i\rho)} \phi_x(u+i\rho)du \\ &= \frac{e^{x\rho}}{2\pi} \int_{-\infty}^{\infty} e^{-ixu} \frac{\phi_X(u+i\rho)}{\rho - iu} du. \end{aligned}$$

Let

$$g_\rho(u) = \frac{\phi_X(u + i\rho)}{\rho - iu}.$$

Then we can show that $g_\rho(-u) = \overline{g_\rho(u)}$ with $u \in \mathbb{R}$ and hence we have

$$\int_{-\infty}^{\infty} e^{-ixu} g_\rho(u) du = 2\Re\left(\int_0^{\infty} e^{-ixu} g_\rho(u) du\right).$$

Therefore, we obtain Eq. 48.5.

References

Asai, M., & McAleer, M. (2009). The structure of dynamic correlations in multivariate stochastic volatility models. *Journal of Econometrics, 150*, 182–192.

Belsley, D., Kuh, E., & Welsch, R. (1980). *Regression diagnostics: Identifying influential data and sources of collinearity*. New York: Wiley.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics, 31*, 307–327.

Carchano, O., & Pardo, A. (2009). Rolling over stock index futures contracts. *Journal of Futures Markets, 29*, 684–694.

Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica, 41*, 135–155.

Conrad, J. S., Hameed, A., & Niden, C. (1994). Volume and autocovariances in short-horizon individual security returns. *Journal of Finance, 49*, 1305–1329.

Copeland, T. E. (1976). A model of asset trading under the assumption of sequential information arrival. *Journal of Finance, 34*, 1149–1168.

Doetsch, G. (1970). *Introduction to the theory and application of the Laplace transformation*. New York: Springer.

Engle, R. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance with estimates of the variance of United Kingdom inflation. *Econometrica, 50*, 987–1007.

Engle, R., Focardi, S. M., & Fabozzi, F. J. (2008). ARCH/GARCH Models in applied financial econometrics. In F. J. Fabozzi (Ed.), *Handbook of finance*. Hoboken: Wiley. vol III.

Epps, T. W., & Epps, L. M. (1976). The stochastic dependence of security price changes and transaction volumes: Implications for the mixture-of-distribution hypothesis. *Econometrica, 44*, 305–321.

Foster, F. J. (1995). Volume-volatility relationships for crude oil futures markets. *Journal of Futures Markets, 15*, 929–951.

Fung, H. G., & Patterson, G. A. (1999). The dynamic relationship of volatility, volume, and market depth in currency futures markets. *Journal of International Financial Markets Institutions and Money, 17*, 33–59.

Garcia, P., Leuthold, M. R., & Zapata, H. (1986). Lead-lag relationships between trading volume and price variability: New evidence. *Journal of Futures Markets, 6*, 1–10.

Gwilym, O., MacMillan, D., & Speight, A. (1999). The intraday relationship between volume and volatility in LIFFE futures markets. *Applied Financial Economics, 9*, 593–604.

Huang, C. Y., & Lin, J. B. (2004). Value-at-risk analysis for Taiwan stock index futures: Fat tails and conditional asymmetries in return innovations. *Review of Quantitative Finance and Accounting, 22*, 79–95.

Karpoff, J. M. (1987). The relation between price changes and trading volume: A survey. *Journal of Financial and Quantitative Analysis, 22*, 109–126.

Kim, Y. S., Rachev, S. T., Chung, D. M., & Bianchi, M. L. (2009). The modified tempered stable distribution, GARCH-models and option pricing. *Probability and Mathematical Statistics, 29*, 91–117.

Kim, Y. S., Rachev, S. T., Bianchi, M. L., & Fabozzi, F. J. (2010). Tempered stable and tempered infinitely divisible GARCH models. *Journal of Banking and Finance, 34*, 2096–2109.

Kim, Y. S., Rachev, S. T., Bianchi, M. L., Mitov, I., & Fabozzi, F. J. (2011). Time series analysis for financial market meltdowns. *Journal of Banking and Finance, 35*, 1879–1891.

Kupiec, P. (1995). Techniques for verifying the accuracy of risk measurement models. *Journal of Derivatives, 6*, 6–24.

Ragunathan, V., & Peker, A. (1997). Price variability, trading volume and market depth: Evidence from the Australian futures market. *Applied Financial Economics, 7*, 447–454.

Schwert, G. W. (1989). Why does stock market volatility change over time? *Journal of Finance, 44*, 1115–1153.

Smirlock, M., & Starks, L. (1985). A further examination of stock price changes and transactions volume. *Journal of Financial Research, 8*, 217–225.

Tang, L. T., & Shieh, J. S. (2006). Long memory in stock index futures markets: A value-at-risk approach. *Physica A: Statistical Mechanics and its Applications, 366*, 437–448.

Tauchen, G., & Pitts, M. (1983). The price variability-volume relationship on speculative markets. *Econometrica, 51*, 485–505.

Verbeek, M. (2004). *A guide to modern econometrics*. Hoboken: Wiley.

Wang, G., & Yau, J. (2000). Trading volume, bid-ask spread, and price volatility in futures markets. *Journal of Futures Markets, 20*, 943–970.

Computer Technology for Financial Service

49

Fang-Pang Lin, Cheng-Few Lee, and Huimin Chung

Contents

F.-P. Lin (✉)
National Center for High Performance Computing, Hsinchu, Taiwan
e-mail: fplin@nchc.narl.org.tw

C.-F. Lee
Department of Finance and Economics, Rutgers Business School, Rutgers, The State University of New Jersey, Piscataway, NJ, USA

Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
e-mail: cflee@business.rutgers.edu

H. Chung
Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
e-mail: chunghui@mail.nctu.edu.tw

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_49,

Abstract

Securities trading is one of the few business activities where a few seconds processing delay can cost a company big fortune. The growing competition in the market exacerbates the situation and pushes further towards instantaneous trading even in split second. The key lies on the performance of the underlying information system. Following the computing evolution in financial services, it was a centralized process to begin with and gradually decentralized into a distribution of actual application logic across service networks. Financial services have tradition of doing most of its heavy-duty financial analysis in overnight batch cycles. However, in securities trading it cannot satisfy the need due to its ad hoc nature and requirement of fast response. New computing paradigms, such grid and cloud computing, aiming at scalable and virtually standardized distributed computing resources, are well suited to the challenge posed by the capital market practices. Both consolidate computing resources by introducing a layer of middleware to orchestrate the use of geographically distributed powerful computers and large storages via fast networks. It is nontrivial to harvest the most of the resources from this kind of architecture. Wiener process plays a central role in modern financial modeling. Its scaled random walk feature, in essence, allows millions of financial simulation to be conducted simultaneously. The sheer scale can only be tackled via grid or cloud computing. In this study the core computing competence for financial services is examined. Grid and cloud computing will be briefly described. How the underlying algorithm for financial analysis can take advantage of grid environment is chosen and presented. One of the most popular practiced algorithms Monte Carlo simulation is used in our case study for option pricing and risk management. The various distributed computational platforms are carefully chosen to demonstrate the performance issue for financial services.

Keywords

Financial service • Grid and cloud computing • Monte Carlo simulation • Option pricing • Risk management • Cyberinfrastructure • Random number generation • High-end comptuing • Financial simulation • Information technology

49.1 Introduction

49.1.1 Information Technology (IT) for Financial Services

The finance services industry involves a broad range of organizations such as banks, credit card companies, insurance companies, consumer finance companies, stock brokerages, investment funds, and some government-sponsored enterprises. The industry represents a significant share of the global market. Information technology (IT) in the financial service industry is considered as an indispensable tool for productivity as well as competitiveness in the market. The IT spending in financial service industry grows constantly across different industry verticals (banking, insurance, and securities and investments). The impact directly from the use of advanced IT brings on financial services industry on the rise.

The structure of the industry has changed significantly in the last two decades as companies, which are not traditionally viewed as financial service providers, have taken advantage of opportunities created by technology to enter the market. New technology-based services keep emerging. These changes are direct result of the interaction of technology with the industrial environment, such as economic atmosphere, societal pressures, and the legal/regulatory environment in which the financial service industry operates. The effects of IT on the internal operations, the structure, and the types of services offered by the financial service industry have been particularly profound (Phillips et al. 1984; Hauswald and Marquez 2003; Griffiths and Remenyi 2003). IT technology has been and continues to be both a motivator and facilitator of change in the financial service industry, which ultimately leads to competitiveness of the industry. The change is in particular radical after 1991 when the World Wide Web was invented by Tim Berners-Lee and his group for information sharing in the community of high energy physics. It was later introduced to the rest of the world, which subsequently changed the face of how people doing business today.

Informational considerations have long been recognized to determine not only the degree of competition but also the pricing and profitability of financial services and instruments. Recent technological progress has dramatically affected the production and availability of information, thereby changing the nature of competition in such informationally sensitive markets. Hauswald and Marquez (2003) investigate how advances in information technology (IT) affect competition in the financial services industry, particularly credit, insurance, and securities markets. Two aspects of improvement in IT are focused: better processing and easier dissemination of information. In other words, two dimensions of technology progress that affect competition in financial services can be defined as advances in the ability to process and evaluate information and in the ease of obtaining information generated by competitors. While better technology may result in improved information processing, it might also lead to low cost or even free access to information through, for example, informational spillovers. They show that in the context of credit screening, better access to information decreases interest rates and the returns from screening. On the other hand, an improved ability to process information increases interest rates and bank profits. Hence predictions regarding financial claims' pricing hinge on the overall effect ascribed to technological progress. Their results conclude that in general financial market informational asymmetries drive profitability.

The viewpoint of Hauswald and Marquez is adopted in this work. Assuming competitors in the dynamics of financial market possess similar capacity, the informational asymmetries can be created sometimes only between seconds and now are possible to be achieved through the outperformance of underlying IT platforms.

49.1.2 Competitiveness Through IT Performance

Following the computing evolution in financial services, it was a centralized process to begin with and gradually decentralized into a distribution of practical

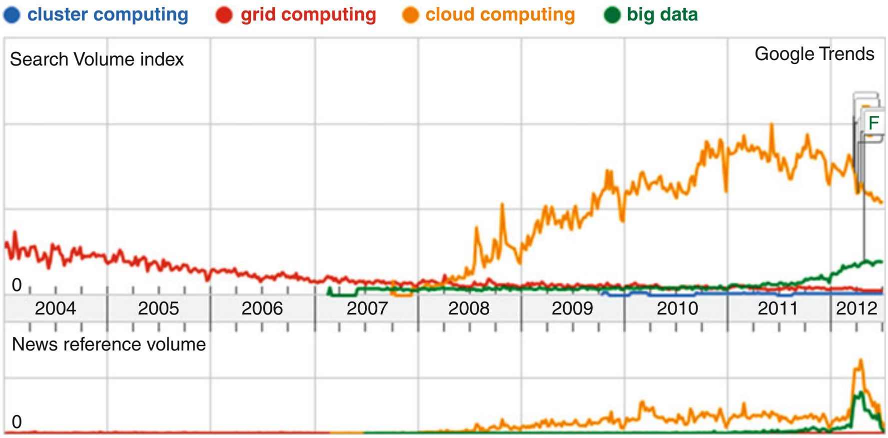


Fig. 49.1 The trend history from Google Trend according to global Search Volume and global News Reference Volume, in which the alphabetic letters represent the specific events that relate to each curve

trading application logic across service networks. Financial services have tradition of doing most of its heavy lifting financial analysis in overnight batch cycles. However, in securities trading it cannot satisfy the need due to its ad hoc nature and requirement of fast response.

New computing paradigms, grid computing and cloud computing were subsequently emerged in the last decade. The grid computing was initially incorporated into the core context of a well-referenced Atkins' report of National Science Board of the United States, namely, "Revolutionizing Science and Engineering Through Cyberinfrastructure" (Atkins et al. 2003), which lays down a visionary path for future IT platform development of the world. One may observe this trend from statistics from Google Trend regarding the global Search Volume and global News Reference Volume of key phrases of "cluster computing," "grid computing," "cloud computing," and "Big Data" (Fig. 49.1), which represents four main stream computing paradigms in high-end quantitative analysis.

Cluster computing is a group of coupled computers that work closely together so that in many respects they can be viewed as though they are a single computer. They are connected with high-speed local area networks and the purpose is usually to gain more compute cycles with better cost performance and higher availability. The grid computing aims at virtualizing scalable geographically distributed computing and observatory resources to maximize compute cycles and data transaction rates with minimum cost. Cloud computing is more of recent development owing to the similar technology used in global information services providers, such as Google and Amazon. The cloud is referred to as a subset of Internet if to be explained in a simplest fashion. Within the cloud the computers also talk with servers instead of communicating with each other similarly to that of peer-to-peer computing (Milojicic et al. 2002). There are no definitive definitions for the above terminology. However, people tend to view clusters as one of foundational components of grids, or grids as a meta-cluster on wide area networks. This is also known as horizontal integration. The cloud virtualizes further the compute, store, and network in a utility sense and provides an interface between users and grids. We refer to Foster et al. (2008) and Yelick et al. (2011) for a comparison. This perspective considers grids as a backbone of cyberinfrastructure to support the clouds. Similarly, in early days of development of grids, there is a so-called "@HOME" style PC grids (Korpela et al. 2001), which are exactly working on at least ten of thousands of PCs, in which owners of PCs donate their CPU times when their machines are in idle. The PC grids can be specifically categorized as clouds.

Figure 49.1 shows that there is a gradual drop in the curve of search volume for grid computing and cluster computing, and many surges, grows but a recent quick drop on cloud computing since its introduction in mid-2007. The new rising technology is Big Data (Bughin et al. 2010), which implies a paradigm shift from compute centric, network centric gradually to data-centric computing. However, the size of the search volume strongly relates to the degree of maturity of each computing paradigm. This is obvious in cluster computing. Clusters are the major market products, either in supercomputers from big vendors, such as IBM, HP, SGI, and NEC, or from aggregation of PCs in university research laboratories. Figure 49.1 also implies constant market need for high-end computing. The performance and security issues are fundamental to general distributed and parallel computing, which also remain as a challenge to cluster, grid, cloud, and Big Data (Lauret et al. 2010; Ghoshal et al. 2011; Ramakrishnan et al. 2011). Performance models in compute-based grid, which is also cloud-like, environment, are adopted in this work. The general definitions of grid and cloud computing will be introduced and briefly compared. To tackle the core performance issue, grids are chosen to demonstrate how fundamental financial calculations can be improved, hence leverage the financial service.

Grid computing, following by cloud computing as shown in Fig. 49.1, has been matured to serve as a production environment for finance services in recent years. Grid computing is well suited to the challenge posed by the capital market practices. In this study the core computing competence for financial services will be examined and how underlying algorithms for financial analysis can take advantage of grid environment scrutinized. One of the most popular practiced algorithms is Monte Carlo simulation (MCS), and it will be specifically used in our case study for calculations of option pricing and for value at risk (VaR) in risk management.

Three grid platforms are carefully chosen to exploit the performance issue for financial services. The first one is traditional grid platform with heterogeneous and distributed resources. Usually digital packets are connected via optical fibers. For long distance, depending on network traffics, it will produce approximately 150-300 microseconds (mm) latency across the Pacific Ocean. This is the physical constrain of light speed when traveling through the fiber channels. Therefore, even in split-second packets can still travel to anywhere in the world. The Pacific Rim Applications and Grid Middleware Assembly (PRAGMA) grid is a typical example, which linked with 14 countries and 36 sites. The system is highly heterogeneous. The computer nodes mounted to PRAGMA grid range from usual PC clusters to high-end supercomputers. The second one is a special Linux, or DRBL, PC cluster. It converts system into a homogenous Linux system and exploits the compute cycles of the cluster. The intention is to provide dynamic and flexible resources to cope better with uncertainty of the traders' cycle demand. Finally, PC grid is chosen to demonstrate finance services that can be effectively conducted through a cloud-based computing. The usefulness of PC grid is based on the fact that 90 % of CPUs time of PCs were in idled status.

49.2 Performance Enhancement

In this section two types of grid systems, compute intensive and data intensive, respectively, are introduced. The classification of the types is based on various grid applications. Traditionally, the grid systems provide a general platform to harvest or to scavenge, if used only in idle status, compute cycles for a collection of resources across boundaries of institutional administration. In real world most applications are in fact data centric. For example, in a trading center, it collects tick-by-tick volume data from all related financial markets and is driven by informational flows, hence typical data centric. However, as noted in Sect. 49.3.2.1, the core competence still lies on the performance enhancement of the IT system. The following two subsections will give more details of compute intensive as well as data-intensive grid systems by a survey of current development of grids specifically for financial services. In some cases, e.g., high-frequency data with real-time analysis, two systems have to work together to get better performance. Our emphasis will be more on compute intensive grid system.

49.2.1 High-End Computing Technology

49.2.1.1 Definitions of High-End Computing

Grid was coined by Ian Foster (Foster and Kessleman 2004) who gave the essence of the definitions as below:

> The sharing that we are concerned with is not primarily file exchange but rather direct access to computers, software, data, and other resources, as is required by a range of collaborative problem solving and resource-brokering strategies emerging in industry, science, and engineering. This sharing is, necessarily, highly controlled, with resource providers and consumers defining clearly and carefully just what is shared, who is allowed to share, and the conditions under which sharing occurs. A set of individuals and/or institutions defined by such sharing rules form what we call a *virtual organization*.

The definition is centered on the concept of virtual organization, but it is not explicit enough to explain what the grid is. Foster then provides additional checklist as below to safeguard the possible logic pitfalls of the definition. Hereby, grid is a system that:

1. Coordinates resources that are not subject to centralized control.

   A grid integrates and coordinates resources and users that live within different control domains – for example, the user's desktop vs. central computing, different administrative units of the same company, or different companies – and addresses the issues of security, policy, payment, membership, and so forth that arise in these settings. Otherwise, we are dealing with a local management system.

2. Using standard, open, general-purpose protocols and interfaces.

   A grid is built from multipurpose protocols and interfaces that address such fundamental issues as authentication, authorization, resource discovery, and resource access. As I discuss further below, it is important that these protocols and interfaces be *standard* and *open*. Otherwise, we are dealing with an application specific system.

3. *To deliver nontrivial qualities of service*.

   A grid allows its constituent resources to be used in a coordinated fashion to deliver various qualities of service, relating, for example, to response time, throughput, availability, and security, and/or co-allocation of multiple resource types to meet complex user demands, so that the utility of the combined system is significantly greater than that of the sum of its parts.

The definition of grid thus far is well accepted and has been stably used up to now. The virtual organization (VO) has strong implication of community driven and collaborative sharing of distributed resources. The advance of development of optical fiber network in recent years plays a critical role of why grids can be a reality. It is also the reason why now the computing paradigm shifts to distributed/grid computing.

Additionally, perhaps the most generally useful definition is that *a grid consists of shared heterogeneous computing and data resources networked across administrative boundaries*. Given such a definition, a grid can be thought of as both an access method and a platform, with grid middleware being the critical software that enables grid operation and ease of use.

The term "cloud computing" has been used to refer to different concepts, models, and services over the last few years. The definition for cloud computing provided by the National Institute of Standards and Technology (NIST) is well received in the IT community, which defines cloud computing as *a model for enabling convenient, on-demand network access to a shared pool of configurable computing resources (*e.g., *networks, servers, storage, applications, and services) that can be rapidly provisioned and released with minimal management effort service provider interaction* (Mell and Grance 2011). The model gains popularity in the industry for its emphasis on pay-as-you-go and elasticity, the ability to quickly expand and collapse the utilized service as demand requires. Thus new approaches to distributed computing and data analysis have also emerged in conjunction with the growth of cloud computing. These include models like MapReduce (Dean and Ghemawat 2004) and scalable key-value stores like Big Table (Chang et al. 2006).

From the high-end computing perspective, cloud computing technology allows users to have the ability to get on-demand access to resources to replace or supplement existing systems, as well as the ability to control the software environment. Yet the core competence still lies on the performance of financial calculation and further of the transactions of financial processes. This work will focus on the core competence in financial calculation based on grid environment. Grid computing technology will be used to explain how the core financial calculations can be significantly accelerated in various distributed and parallel computing environments. The calculation models in this work can be easily migrated to pure cloud environments.

49.2.1.2 Essence of IT Technology

To realize the above goal, it needs to handle technically interoperability of middleware that is capable of communicating between heterogeneous computer systems across institutional boundaries. The movement of grid began in 1996 by Ian Foster and Kessleman (2004). Before their development, another branch of high-performance computing that focuses on connecting geographically distributed supercomputers to achieve one single grand task had been developed by Smarr and Catlett (1992). They coined such a methodology as metacomputing and their query has been how we can have infinite computing power under the physical limit, such as Moore's Law. However, it remains to be less useful because its limit goal on pursuing top performance without noticing practical use in real world. The idea lives on and generates many tools dedicated to high-performance/high-throughput computing, such as Condor (Litzkow et al. 1988), Legion (Grimshaw and Wulf 1997), and UNICORE (Almond and Snelling 1999). Condor, as suggested by the name of the project, is devised to scavenge a large cluster of idle workstations. Legion is closer to the development of worldwide virtual computer. The goal of UNICORE is even much simpler and practical. It was developed when Germany government decided to consolidate their five national supercomputer centers into a virtual one to reduce the management cost and needed a software tool to integrate them, hence the UNICORE. These tools were successful under their development scope. However they fail to meet the first and the second items in Foster's checklist in the previous section.

The emergency of grids follows the similar path as that of Condor and Legion at the first place, in which its development aims at resources sharing in high-performance computing. However, its vision in open standards and the concept of virtual organization allows its development go far beyond merely cluster supercomputers together. It gives a broader view of resources sharing, in which it is not only limited to the sizable computing cycles and storage space to be shared but also extended virtually to calculable machines that are able to hook up to the Internet, such as sensors and sensor loggers, storage servers, and computers. Since 1996, Foster and his team have been developing software tools to achieve the purpose. Their software Globus Toolkit (Foster and Kessleman 2004) is now a de facto middleware for grids. However, the ambitious development is still considered insufficient to meet the ever-growing complexity of grid systems.

As mentioned earlier that grid is based on open specifications and standards, they allow all stakeholders within the virtual organization/grid to communicate with each other with ease and enable ones more to focus on integrated value creation activities. The open specifications and standards are made by the community of Open Grid Forum (OGF), which plays as a standard body and made, discussed, and announced new standards during regular OGF meetings. Grid Specifications and Standards include architecture, scheduling, resource management, system configuration, data, data movement, security, and grid security infrastructure. In 2004, OGF announced Globus Toolkit version, which adopts both the open standard of grid, Open Grid Services Architecture (OGSA), and the more widely adopted World Wide Web standard, web services resource framework (WSRF), which ultimately enable grids to tackle issues of both scalability and complexity of very large grid systems.

49.2.2 Compute Intensive IT Systems

The recent development of computational finance based on grids is hereby scrutinized and remarks given. Our major interest is to see if the split-second performance is well justified under the grid architecture. Also, real-time issue with real market parametric data should be used as input for practical simulation. In addition, issues of intersystem, interdisciplinary and geographically distribution of resources, and the degree of virtualization are crucial to the success of such a grid. The chosen projects are reviewed and discussed as follows:

1. PicsouGrid

   This is a French grid project for financial service. It provides a general framework for computation finance and targets on applications of option trading, option pricing, Monte Carlo simulation, aggregation of statistics, etc. (Stokes-Rees et al. 2007). The key for this development is the implementation of the middleware ProActive. ProActive is an in-house Java library for distributed computing developed by INRIA Sophia Antipolis, France. It provides transparent asynchronous distributed method calls and is implemented on top of Java RMI. It is also used in commercial applications. It also provides fault tolerance mechanism. The architecture is shown in Fig. 49.2, which is very similar to most of grid applications apart from the software stack used. The option pricing was tested in an approximately 894 CPUs. The underlying computer systems are heterogeneous. The system is used for metacomputing. As a result, the system has to specifically design to orchestrate and to synchronize and re-synchronize the whole distributed processes for one calculation. Once the grid system requires synchronization between processes, which implies stronger coupling of algorithm of interest, the performance will be seriously affected. There is no software treatment to solve such problems and should be tackled by physical infrastructure, e.g., optical fiber network with Layer 2 light path.

2. FinGrid

   FinGrid stands for Financial Information Grid. Its study includes components of bootstrapping, sentimental analysis, and multi-scale analysis, which focuses on information integration and analysis, e.g., data mining. It takes advantage of the huge collection of numerical and textual data simultaneously to emphasize the study of societal issues (Amad et al. 2004; Ahmad et al. 2005; Gillam et al. 2005). The architecture of FinGrid is shown in Fig. 49.3. It is a typical 3-tier system, in which the first tier facilitates the client in sending a request to one of the services: Text Processing Service or Time Series Service; the second tier

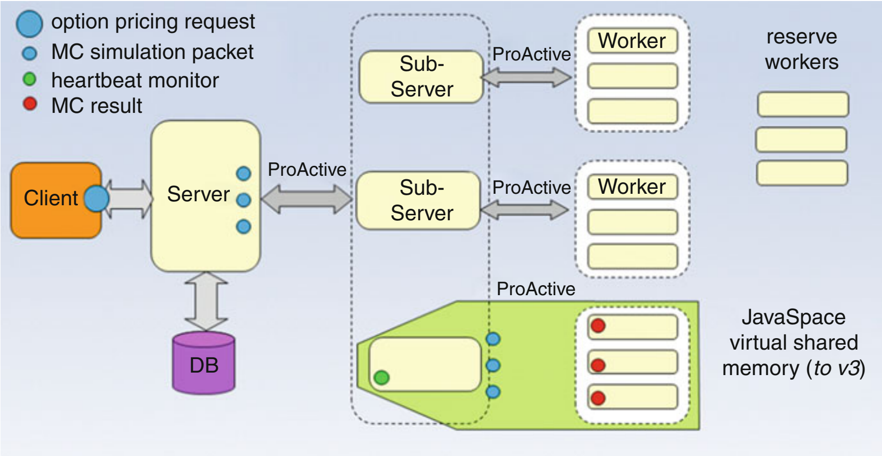


Fig. 49.2 Architecture of PicsouGrid for option pricing based on Monte Carlo simulation (Stokes-Rees et al. 2007)

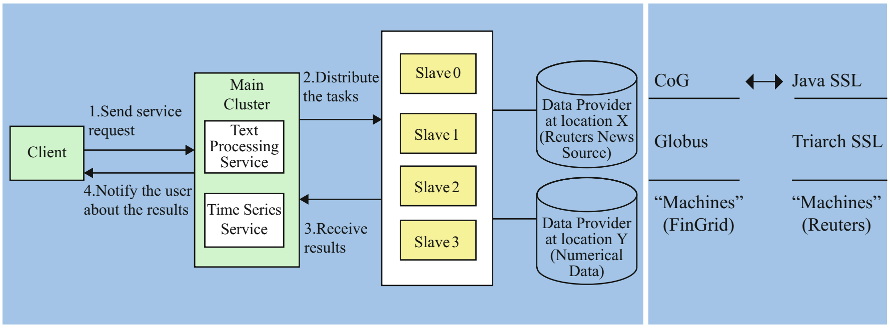


Fig. 49.3 The architecture of Financial Information Grid (FinGrid)

facilitates the execution of parallel tasks in the main cluster and is distributed to a set of slave machines (nodes), and the third tier comprises the connection of the slave machines to the data providers. This work focuses on small scale and dedicated grid system. It pumps in real and live numerical and textual data from say Reuters and performs real-time sophisticated data mining analysis. This is a good prototype for financial grid. However, it will encounter similar problem as that of PicsouGrid if it is to scale up. The model is more successful in automatically combining real data and the analysis.

3. IBM Japan collaborates with life insurance company and adopts PC grids concept to scavenge more compute cycles (Tanaka 2003):

   In this work an integrated risk management system (see Fig. 49.4) is modified, in which the future scenarios of red circle of Fig. 49.4 are send via grid middleware to a cluster of PCs. According to the size of the given PCs, the

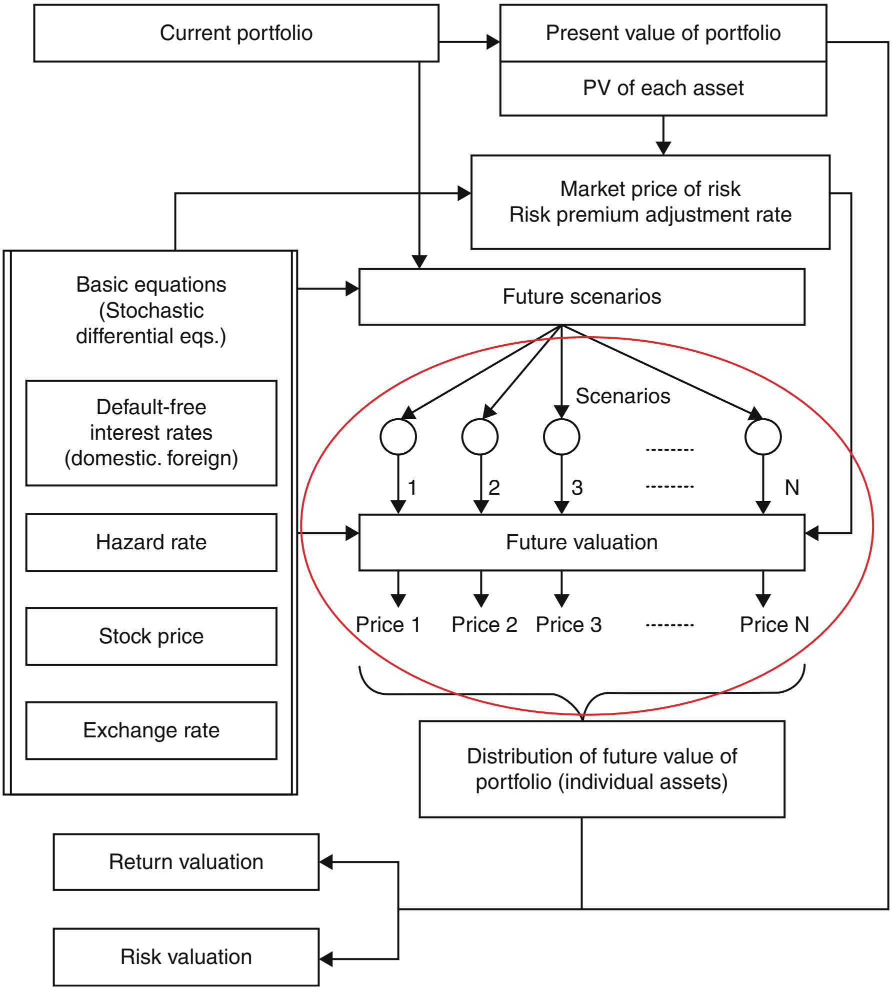


Fig. 49.4 Architecture of Integrated Risk Management System (Tanaka 2003)

number scenarios are then divided in a work balanced manner for each PC. This is the most typical use of compute intensive grid systems and a good practice for production system. However, the key issues that discussed in the above two cases cannot be answered in this study. Similar architecture can also be found in EGrid (Leto et al. 2005).

4. UK e-Science developed a grid service discovery in the financial market sector focusing on integration of different knowledge flows (Bell and Ludwig 2005).

From application's viewpoint, business and technical architecture of financial service applications may be segmented by product, process, or geographic concerns. Segmented inventories make inter silo reuse difficult. The service integration model is adopted and a loosely coupled inventory – containing differing explicit

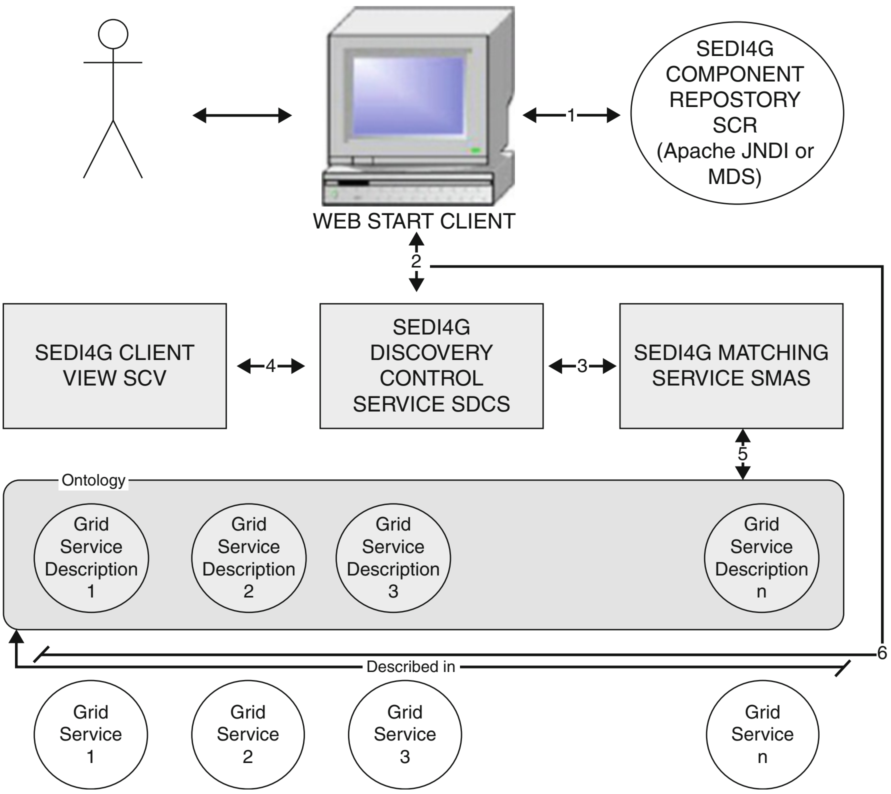


Fig. 49.5 The semantic discovery for Grid Services Architecture (SEDI4G) (Bell and Ludwig 2005)

capability knowledge. Three use cases were specifically chosen in this work to explore the use of semantic searching:

Use case 1 – Searching for trades executed with a particular counterparty

Use case 2 – Valuing a portfolio of interest rate derivative products

Use case 3 – Valuing an option-based product

The use cases were chosen to provide examples of three distinct patterns of use – aggregation, standard selection, and multiple selection. The architecture (see Fig. 49.5) is bound specifically with the user cases. The advantage for grid in this case is that it can be easily tailored into specific user need to integrate different applications, which is a crucial strength of using grid.

49.2.3 Data-Intensive IT Systems

Grid in financial services from the perspective of web services towards financial services industry. The perspective is more on transactional side. Once the bottleneck of compute cycle is solved, the data-centric nature will play the key role again.

The knowledge flows back to the customized business logic should provide the best path for users to access the live data of interest. There is no strong focus of development on this data-intensive grid system. Even in FinGrid, which claims in streaming live data for real-time analysis, the data issue remains part of compute grids. However, the need for dynamic data management is obvious as mentioned in Amad et al. (2004). Hereby, we like to introduce and implement a dynamic data management software Ring Buffer Network Bus (RBNB) DataTurbine to serve such a purpose.

RBNB DataTurbine was used recently to support global environmental observatory network, which involves linking with ten of thousand of sensors and is able to obtain the observed data online. It meets grid/cyberinfrastructure (CI) requirements with regard to data acquisition, instrument management, and state-of-health monitoring including reliable data capture and transport, persistent monitoring of numerous data channels, automated processing, event detection and analysis, integration across heterogeneous resources and systems, real-time tasking and remote operations, and secure access to system resources. To that end, streaming data middleware provides the framework for application development and integration.

Use cases of RBNB DataTurbine include adaptive sampling rates, failure detection and correction, quality assurance, and simple observation (see Tilak et al. 2007). Real-time data access can be used to generate interest and buy-in from various stakeholders. Real-time streaming data is a natural model for many applications in observing systems, in particular event detection and pattern recognition. Many of these applications involve filters over data values, or more generally, functions over sliding temporal windows. The RBNB DataTurbine middleware provides a modular, scalable, robust environment while providing security, configuration management, routing, and data archival services. The RBNB DataTurbine system acts as an intermediary between dissimilar data monitoring and analysis devices and applications. As shown in Fig. 49.6, a modular architecture is used, in which a source or "feeder" program is a Java application that acquires data from an external live data sources and feeds it into the RBNB server. Additional modules display and manipulate data fetched from the RBNB server. This allows flexible configuration where RBNB serves as a coupling between relatively simple and "single purpose" suppliers of data and consumers of data, both of which are presented a logical grouping of physical data sources. RBNB supports the modular addition of new sources and sinks with a clear separation of design, coding, and testing (ref. Fig. 49.6). From the perspective of distributed systems, the RBNB DataTurbine is a "black box" from which applications and devices send data and receive data. RBNB DataTurbine handles all data management operations between data sources and sinks, including reliable transport, routing, scheduling, and security. RBNB accomplishes this through the innovative use of memory and file-based ring buffers combined with flexible network objects. Ring buffers are a programmer-configurable mixture of memory and disk, allowing system tuning to meet application-dependent data management requirements. Network bus elements perform data stream multiplexing and routing. These elements

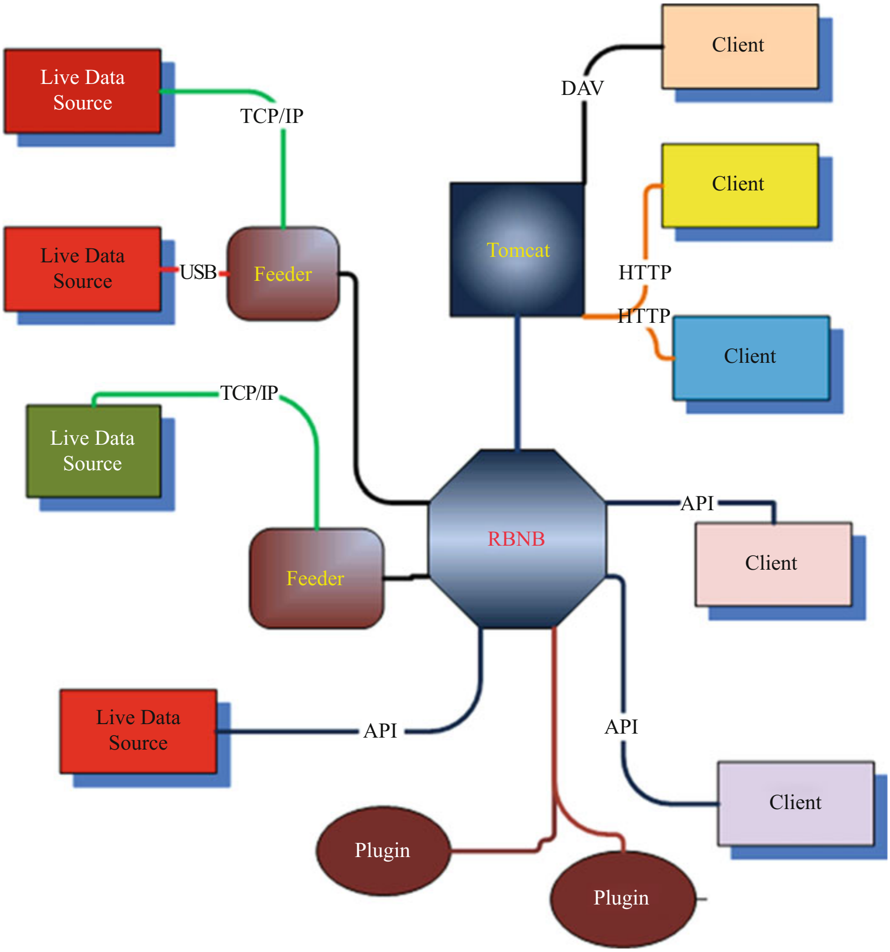


Fig. 49.6 RBNB DataTurbine use scenario for collaborative applications

combine to support seamless real-time data archiving and distribution over existing local and wide area networks. Ring buffers also connect directly to client applications to provide streaming-related services including data stream subscription, capture, rewind, and replay. This presents clients with a simple, uniform interface to real-time and historical (playback) data.

49.3 Distributed and Parallel Financial Simulation

In the previous sections, we address issues of incorporating IT technology for financial competitiveness and derive that the core lies on the performance of IT platform, providing the competitors in the market have similar capacity and are equally informed. Grid technology, as the leading IT development in high-performance computing, is introduced as the cutting-edge IT platform to meet our goal. Many companies have adopted similar technology of grids with success as mentioned in Sect. 49.1. There are also increasing research interests, which result in the work discussed in Sect. 49.2.3. Better performance, however, cannot be achieved by merely using a single architecture as observed in the cases of Sect. 49.2.3. The architecture obviously has to be specifically chosen for the analysis of interest. Simultaneously, the analysis procedures have to be tailored into the chosen architecture for performance fine-tune.

In this section, we will introduce and discuss analysis procedures of financial simulation and how to tailor the analysis procedures into grid architectures by distribution and parallelism. The popular calculations for option pricing and for value at risk (VaR) in trading practice are used to serve the purpose. The calculation is based on Monte Carlo simulation, which is chosen not only because it is a well-received approach due to the absence of straightforward closed-form solutions for many financial models but also a numerical method intrinsically suited to mass distribution and mass parallelism. The success of Monte Carlo simulation lies on the quality of random number generator, which will be discussed in details at the end of the section.

49.3.1 Financial Simulation

There are wide variety of sophisticated financial models developed, to name a few, ranging from analysis in time series, fractals, nonlinear dynamics, and agent-based modeling to applications in optional pricing, portfolio management, and market risk measure, etc. (Schmidt 2005), in which option pricing and VaR calculations of market risk measure can be considered crucial and one of the most practiced activities in market trading.

49.3.1.1 Option Pricing

An option is an agreement between two parties to buy or sell an asset at a certain time in the future for a certain price. There are two types of options:

Call Option: A call option is a contract that gives the right to its holder (i.e., buyer) without creating an obligation, to buy a prespecified underlying asset at a predetermined price. Usually this right is created for a specific time period, e.g., 6 months or more. If the option can be exercised only at its expiration (i.e., the underlying asset can be purchased only at the end of the life of the option), the option is referred to as a European-style Call Option (or European Call). If it can be exercised any date before its maturity, the option is referred to as an American-style Call Option (or American Call).

Put Option: A put option is a contract that gives its holder the right without creating the obligation to sell a prespecified underlying asset at a predetermined price. If the option can be exercised only at its expiration (i.e., the underlying asset can be sold only at the end of the life of the option), the option is referred to as a European-style Put Option (or European Put). If it can be exercised any date before its maturity, the option is referred to as an American-style Put Option (or American Put).

To price options in computational finance, we use the following notation: K is the strike price; T is the time to maturity of the option; S_t is the stock price at time t; r is the risk-free interest rate; μ is the drift rate of the underlying asset (a measure of the average rate of growth of the asset price); σ is the volatility of the stock; and V denotes the option value. Here is an example to illustrate the concept of option pricing. Suppose an investor enters into a call option contract to buy a stock at price K after 3 months. After 3 months, the stock price is S_t. If $S_t > K$ then one can exercise one's option by buying the stock at price K and by immediately selling in the market to make a profit of $S_T - K$. On the other hand, if he $S_T - K$ to be buy the stock. Hence, we see that a call option to buy the stock at time T at price K will get payoff $(S_T - K)^+$, where $(S_T - K)^+ \equiv \max(S_T - K, 0)$ (Schmidt 2005; Hull 2003).

49.3.1.2 Market Risk Measurement Based on VaR

Market risks are the prospect of financial losses or gains, due to unexpected changes in market prices and rates. Evaluating the exposure to such risks is nowadays of primary concern to risk managers in financial institutions. Until the late 1980s market risk was estimated through gap and duration analysis (interest rates), portfolio theory (securities), sensitivity analysis (derivatives), or scenarios analysis. However, all these methods could be either applied only to very specific assets or relied on subjective reasoning.

Since the early 1990s a commonly used market risk estimation methodology has been the value at risk (VaR). A VaR measure is the highest possible loss L incurred from holding the current portfolio over a certain period of time at a given confidence level (Dowd 2002):

$$\mathrm{P}(L > VaR) \leq 1 - c \tag{49.1}$$

where c is the confidence level, typically 95 %, 97.5 %, or 99 %, and P is cumulative distribution function. By convention, $L = -\Delta X(\tau)$, where $\Delta X(\tau)$ is the relative change (return) in portfolio value over the time horizon τ. Hence, large values of L correspond to large losses (or large negative returns).

The VaR figure has two important characteristics: (1) it provides a common consistent measure of risk across different positions and risk factors and (2) it takes into account the correlations or dependencies between different risk factors. Because of its intuitive appeal and simplicity, it is no surprise that in a few years value at risk has become the standard risk measure used around the world. However, VaR has a few deficiencies, among them the non-subadditivity – a sum of VaR's two portfolios can be smaller than the VaR of the combined portfolio. To cope with these shortcomings, Artzner et al. proposed an alternative measure that satisfies the assumptions of a coherent risk measure. The expected shortfall (ES), also called expected tail loss (ETL) or conditional VaR, is the expected value of the losses in excess of VaR:

$$\mathrm{ES} = \mathrm{E}(L|L > VaR) \tag{49.2}$$

It is interesting to note that although new to the finance industry – expected shortfall has been familiar to insurance practitioners for a long time. It is very similar to the mean excess function which is used to characterize claim size distribution; see (Cizek et al. 2011).

The essence of the VaR and ES computations is estimation of low quantiles in the portfolio return distributions. Hence, the performance of market risk measurement methods depends on the quality of distribution assumptions on the underlying risk factors. Many of the concepts in theoretical and empirical finance developed over the past decades, including the classical portfolio theory, the Black-Scholes-Merton option pricing model, and even the RiskMetrics variance-covariance approach to VaR rest upon the assumption that asset returns follow a normal distribution. The assumption is not justified by real market data. Our interest is more on the calculation side. For interested readers we refer further to (Weron 2004).

49.3.2 Monte Carlo Simulations

49.3.2.1 Monte Carlo and Quasi-Monte Carlo Methods

In general, Monte Carlo (MC) and quasi-Monte Carlo (QMC) methods are applied to estimate the integral of function $f(x)$ over $[0, 1]^d$ unit hypercube where d is the dimension of the hypercube:

$$I = \int_{[0,1]^d} f(x)dx \tag{49.3}$$

In MC methods, I is estimated by evaluating $f(x)$ at N independent points randomly chosen from a uniform random distribution over $[0, 1]^d$ and then evaluating average

$$\hat{I} = \frac{1}{N}\sum_{i=1}^{N} f(x_i) \tag{49.4}$$

From the law of large numbers, $\hat{I} \to I$ as $N \to \infty$. The standard deviation is

$$\sqrt{\frac{1}{N-1}\sum_{i=1}^{\mathrm{N}} (f(x_i) - I)^2} \tag{49.5}$$

Therefore, the error of MC methods is proportional to $N^{-1/2}$.

QMC methods compute the above integral based on low-discrepancy (LD) sequences. The elements in a LD sequence are "uniformly" chosen from $[0, 1]^d$ rather than "randomly." The discrepancy is a measure to evaluate the uniformity of points over $[0, 1]^d$. Let $\{q_n\}$ be a sequence in $[0, 1]^d$; the discrepancy D_N^* of q_n is defined as follows, using Niederreiter's notation (Niederreiter 1992):

$$D_N^*(q_n) = \sup_{B \in [0,1)^d} \left| \frac{A(B, q_n)}{N} - v_d(B) \right| \tag{49.6}$$

where B is a subcube of $[0, 1]^d$ containing the origin, $A(B, q_n)$ is the number of points in q_n that fall into B, and $V_d(B)$ is the d-dimensional Lebesgue measure of B. The elements of q_n are said uniformly distributed if its discrepancy $D_N^* \to 0$ as $N \to \infty$. From the theory of uniform distribution sequences (Kuipers and Niederreiter 1974), the estimate of the integral using a uniformly distributed sequence $\{q_n\}$ is $\hat{I} = \frac{1}{N}\sum_{n=1}^{N} f(q_n), as\, N \to \infty$ then $\hat{I} \to I$. The integration error bound is given by the Koksman-Hlawka inequality:

$$\left| I - \frac{1}{N}\sum_{n=1}^{N} f(q_n) \right| \le V(f) D_N^*(q_n) \tag{49.7}$$

where $V(f)$ is the variation of the function in the sense of Hardy and Krause (see Kuipers and Niederreiter 1974), which is assumed to be finite.

The inequality suggests a smaller error can be obtained by using sequences with smaller discrepancy. The discrepancy of many uniformly distributed sequences satisfies $O((\log N)^d/N)$. These sequences are called low-discrepancy (LD) sequences (Chen et al. 2006). Inequality (49.7) shows that the estimates using a LD sequence satisfy the deterministic error bound $O((\log N)^d/N)$.

49.3.2.2 Monte Carlo Simulations for Option Pricing

Under the risk-neutral measure, the price of a fairly valued European call option is the expectation of the payoff $E[e^{-rT}(S_T - K)^+]$. In order to compute the expectation, Black and Scholes (1973) modeled the stochastic process generating the price of a non-dividend-paying stock as geometric Brownian motion:

$$dS_t = \mu S_t dt + \sigma S_t dW_t \tag{49.8}$$

where W is a standard Wiener process, also known as Brownian motion. Under the risk-neutral measure, the drift μ is set to $\mu = r$.

To simulate the path followed by S, suppose the life of the option has been divided into n short intervals of length $\Delta t (\Delta t = T/n)$, the updating of the stock price at $t + \Delta t$ from t is (Hull 2003):

$$S_{t+\Delta t} - S_t = r S_t \Delta t + \sigma S_t Z \sqrt{\Delta t} \tag{49.9}$$

where Z is a standard random variable, i.e., $Z{\sim}(0,1)$. This enables the value of $S_{\Delta t}$ to be calculated from initial value S_t at time Δt, the value at time $2\Delta t$ to be calculated from $S_{\Delta t}$, and so on. Hence, a completed path for S has been constructed.

In practice, in order to avoid discretization errors, it is usual to simulate ln$S$ rather than S. From It ô's lemma, the process followed by of Eq. 49.9 is (Bratley and Fox 1988)

$$d\ln S = \left(r - \frac{\sigma^2}{2}\right)dt + \sigma dz \tag{49.10}$$

so that

$$\ln S_{t+\Delta t} - \ln S_t = \left(r - \frac{\sigma^2}{2}\right)dt + \sigma Z\sqrt{\Delta t} \tag{49.11}$$

or equivalently

$$S_{t+\Delta t} = S_t \exp\left[\left(r - \frac{\sigma^2}{2}\right)dt + \sigma Z\sqrt{\Delta t}\right] \tag{49.12}$$

Substituting independent samples $Z_i, \cdots, Z_n$ from the normal distribution into (Eq. 49.12) yields independent samples ${S_T}^{(i)}$, $i = 1, \cdots, n$, of the stock price at expiry time T. Hence, the option value is given by

$$V = \frac{1}{n}\sum_{i=1}^{n} V_i = \frac{1}{n}\sum_{i=1}^{n} e^{-rT}\max\left[{S_T}^{(i)} - K, 0\right] \tag{49.13}$$

The QMC simulations follow the same steps as the MC simulations, except that the pseudorandom numbers are replaced by LD sequences. The basic LD sequences known in literature are Halton (1960), Sobol (1967), and Faure (1982). Niederreiter (1992) proposed a general principle of generating LD sequences. In finance, several examples have shown that the Sobol sequence is superior to others. For example, Galanti and Jung (1997) observed that the Sobol sequence outperforms the Faure sequence, and the Faure marginally outperforms the Halton sequence. In this research, we use Sobol sequence in our experiments. The generator used for generating the Sobol sequence comes from the modified algorithm 659 of Joe and Kuo (2003).

49.3.2.3 Monte Carlo Bootstrap for VaR

Monte Carlo simulation is applicable with virtually any model of changes in risk factors and any mechanism for determining a portfolio's value in each market scenario. But revaluing a portfolio in each scenario can present a substantial computational burden, and this motivates research into ways of improving the efficiency of Monte Carlo methods for VaR.

The bootstrap (Efon 1981; Efron and Tibshirani 1986) is a simple and straightforward method for calculating approximated biases, standard deviations, confidence intervals, and so forth, in almost any nonparametric estimation problem. Method is a keyword here, since little is known about the bootstrap's theoretical basis, except that (a) it is closely related to the jackknife in statistic inferring; (b) under reasonable condition, it gives asymptotically correct results; and (c) for some simple problems which can be analyzed completely, for example, ordinary linear regression, the bootstrap automatically produces standard solutions.

The bootstrap method is straightforward. Suppose we observe returns $X_i = x_i$, $i = 1, 2, \cdots, n$, where the X_i are independent and identically distributed (iid) according to some unknown probability distribution F. The X_i may be real valued and two-dimensional or take values in a more complicated space. A given parameter $\theta(F)$, perhaps the mean, median, correlation, and so forth, is to be estimated, and we agree to use the estimate $\hat{\theta} = \theta\left(\hat{F}\right)$, where $\hat{F}$ is the empirical distribution function putting mass $1/n$ at each observed value x_i. We wish to assign some measure of accuracy to $\hat{\theta}$.

Let $\sigma(F)$ be some measure of accuracy that we would use if F were known, for example, $\sigma(F) = \mathrm{SD}_F\left(\hat{\theta}\right)$, the standard deviation of $\hat{\theta}$ when $X_1, X_2, \cdots, X_n \sim F^{(idd)}$. The bootstrap estimate of accuracy $\hat{\sigma} = \sigma\left(\hat{F}\right)$ is the nonparametric maximum likelihood estimate of $\sigma(F)$. In order to calculate $\hat{\sigma}$ it is usually necessary to employ numerical methods. (a) A bootstrap sample $X_1^*, X_2^*, \cdots, X_n^*$ is drawn from $\hat{F}$, in which each X_i^* independently takes value x_j with probability $1/n$, $j = 1, 2, \cdots, n$. In other words, $X_1^*, X_2^*, \cdots, X_n^*$ is an independent sample of size n drawn with replacement from the set of observations $\{x_1,x_2, \cdots, x_n\}$. (b) This gives a bootstrap empirical distribution function $\hat{F}^*$, the empirical distribution of the n values $X_1^*, X_2^*, \cdots, X_n^*$, and a corresponding bootstrap value $\hat{\theta}^* = \theta\left(\hat{F}^*\right)$. (c) Steps (a) and (b) are repeated, independently, in a large number of times, say N, giving bootstrap values $\hat{\theta}^{*1}, \hat{\theta}^{*2}, \cdots, \hat{\theta}^{*N}$. (d) The value of $\hat{\sigma}$ is approximated, in the case where $\sigma(F)$ is the standard deviation by the sample standard deviation of the $\hat{\theta}^*$ values, where

$$\frac{\hat{\mu} = \sum_{j=1}^{n} \hat{\theta}^{*j}}{N} \tag{49.14}$$

and

$$\hat{\sigma}^2 = \frac{\sum_{j=1}^{n} \left(\hat{\theta}^{*j} - \hat{\mu}\right)^2}{N - 1} \tag{49.15}$$

49.3.3 Distribution and Parallelism Based on Random Number Generation

Financial variables, such as prices and returns, are random time-dependent variables. Wiener process plays the central role in modeling. As shown in

Eqs. 49.8 and 49.9 for approximating the underlying prices $S_{t+\Delta t}$, or the bootstrap samples of return X_i^*, the solution methods involve basic market parameters, drift μ, volatility σ, and risk-free interest rate r, current underlying price S or return X, strike price K, and Wiener process, which is related to time to maturity Δt and standard random variable Z, i.e., $\Delta W = Z\sqrt{\Delta t}$. Monte Carlo methods simulate this nature of the Brownian motion directly. It follows Wiener process and approximates the standard random variable Z by introducing pseudo *iid* random number into each Wiener process. When the simulation number is large enough, e.g., if n in Eq. 49.13 is large enough, the mean value will approach the exact solution. The large number for n also implies the performance problems are the key problems for Monte Carlo methods. One the other hand, the *iid* property of the random number Z shows possible solution to tackle the performance problem through mass distribution and/or parallelism. The solution method centers on the random number generation.

The techniques of random number generation can be developed in a simple form through the approximation of a d-dimensional integral, e.g., (Eq. 49.3). Mass distribution and parallelism required solutions of for large dimension. However, most modern techniques in random number generation have limitations. In this study, both tradition pseudorandom number generation and high-dimensional low-discrepancy random number generator are considered.

Following Sect. 49.3.2.1 better solution can be achieved by making use of Sobol sequences, which were proposed by Sobol (1967). A computer implementation in Fortran 77 was subsequently given by Bratley and Fox (1988) as Algorithm 659. Other implementations are available as C, Fortran 77, or Fortran 90 routines in the popular Numerical Recipes collection of software. However, as given, all these implementations have a fairly heavy restriction on the maximum value of d allowed. For Algorithm 659, Sobol sequences may be generated to approximate integrals in up to 40 dimensions, while the Numerical Recipes routines allow the generation of Sobol sequences to approximate integrals in up to six dimensions only. The FinDer software of Paskov and Traub (1995) provides an implementation of Sobol sequences up to 370 dimensions, but it is licensed software. As computers become more powerful, there is an expectation that it should be possible to approximate integrals in higher and higher dimensions. Integrals in hundreds of variables arise in applications such as mathematical finance (e.g., see Paskov and Traub (1995)). Also, as new methods become available for these integrals, one might wish to compare these new methods with Sobol sequences. Thus, it would be desirable to extend these existing implementations such as Algorithm 659 so they may be used for higher-dimensional integrals. We remark that Sobol sequences are now considered to be examples of (t, d)-sequences in base 2. The general theory of these low-discrepancy (t, d)-sequences in base b is discussed in detail in Niederreiter (1992). The generation of Sobol sequences is clearly explained in Bratley and Fox (1988). We review the main points so as to show what extra data would be required to allow Algorithm 659 to generate Sobol sequences to approximate integrals in more than 40 dimensions. To generate the j th component of the points in a Sobol sequence, we need to choose a primitive polynomial of some degree *sj* in the field $\mathbb{Z}_2$, that is, a polynomial of the form

$$x^{s_j} + a_{1,j}x^{s_j-1} + \cdots + a_{s_j-1,j}x + 1, \tag{49.16}$$

where the coefficients $a_{1,j}\_\cdots\_a_{s_j}-1,j$ are either 0 or 1.

We use these coefficients to define a sequence $\{\mathrm{m}_{1,\ j},\ \mathrm{m}_{2,\ j,}\ \cdots\}$ of positive integers by the recurrence relation

$$\begin{aligned} m_{k,j} = a_{1,j}m_{k-1,j} \oplus 2^2 a_{2,j}m_{k-2,j} \oplus \cdots \\ \oplus 2^{s_j-1}a_{s_j-1,j}m_{k-s_j+1,j} \oplus 2^{s_j}a_{s_j,j}m_{k-s_j,j} \oplus m_{k-s_j,j} \end{aligned} \tag{49.17}$$

for $k \geq s_j + 1$, where $\oplus$ is the bit-by-bit exclusive-OR operator. The initial values $m_{1,j}, m_{2,j}, \cdots, m_{s_j,j}$ can be chosen freely provided that each $M_{k,j}$, $1 \leq k \leq s_j$ is odd and less than 2^k. The "direction numbers" $\{v_{1,j}, v_{2,j}, \cdots\}$ are defined by

$$v_{1,j} \equiv \frac{m_{k,j}}{2^k} \tag{49.18}$$

Then $x_{i,j}$, the j th component of the ith point in a Sobol sequence, is given by

$$x_{i,j} = b_1 v_{1,j} \oplus b_2 v_{2,j} \oplus \cdots \tag{49.19}$$

Where b_l is the lth bit from the right when i is written in binary, that is, $(\cdots b_2 b_1)_2$ is the binary representation of i. In practice, a more efficient Gray code implementation proposed by Antonov and Saleev (1979) is used; see Bratley and Fox (1988) for details. We then see that the implementation in Bratley and Fox (1988) may be used to generate Sobol sequences to approximate integrals in more than 40 dimensions by providing more data in the form of primitive polynomials and direction numbers (or equivalently, values of $m_{1,j}, m_{2,j}, \cdots, m_{s_j,j}$). When generating such Sobol sequences, we need to ensure that the primitive polynomials used to generate each component are different and that the initial values of the $m_{k,j}$'s are chosen differently for any two primitive polynomials of the same degree. The error bounds for Sobol sequences given in Sobol (1967) indicate we should use primitive polynomials of as low a degree as possible. We discuss how additional primitive polynomials may be obtained in the next section. After these primitive polynomials have been found, we need to decide upon the initial values of the $m_{k,j}$ for $1 \leq k \leq s_j$. As explained above, all we require is that they be odd and that $m_{k,j} < 2^k$. Thus, we could just choose them randomly, subject to these two constraints. However, Sobol and Levitan (1976) introduced an extra uniformity condition known as Property A. Geometrically, if the cube $[0, 1]^d$ is divided up by the planes $x_j = 1/2$ into 2^d equally sized subcubes, then a sequence of points belonging to $[0, 1]^d$ possesses Property A if, after dividing the sequence into consecutive blocks of 2^d points, each one of the points in any block belongs to a different subcube. Property A is not that useful to have for large d because of the computational time required to approximate an integral using 2^d points. Also, Property A is not enough to ensure that there are no bad correlations between pairs of dimensions. Nevertheless, Property A would seem a reasonable criterion to use in deciding upon a choice of the initial $m_{k,j}$. The numerical results for Sobol sequences given in Sect. 49.4 suggest that the direction numbers obtained here are indeed reasonable. Other ways of obtaining the direction numbers are also possible. For example, in Cheng and Druzdzel (2000), the initial direction numbers are obtained by an interesting technique of minimizing a measure of uniformity in two dimensions. This technique may alleviate the problem of bad correlations between pairs of dimensions that was mentioned above. Sobol (1967) showed that a Sobol sequence used to approximate a d-dimensional integral possesses Property A if and only if

$$\det(V_d) = 1(\mathrm{mod}2), \tag{49.20}$$

where V_d is the $d \times d$ binary matrix defined by

$$V_d = \begin{bmatrix} v_{1,1,1} & v_{2,1,1} & \cdots & v_{d,1,1} \\ v_{1,2,1} & v_{2,2,1} & & v_{d,2,1} \\ \vdots & & \ddots & \vdots \\ v_{1,d,1} & v_{2,d,1} & \cdots & v_{d,d,1} \end{bmatrix} \tag{49.21}$$

With $v_{k,j,1}$ denoting the first bit after the binary point of $v_{k,j,}$. The primitive polynomials and direction numbers used in Algorithm 659 are taken from Sobol and Levitan (1976), and a subset of this data may be found in Sobol (1967). Though it is mentioned in Sobol (1967) that Property A is satisfied for $d \leq 16$, that is, $\det(V_d) = 1$ (mod 2) for all $d \leq 16$, our calculations showed that Property A is actually satisfied for $d \leq 16$. As a result, we change the values of the $m_{k,j}$ for $21 \leq j \leq 40$, but keep the primitive polynomials. For $j \geq 41$, we obtain additional primitive polynomials. The number of primitive polynomials of degree s is $\phi(2^s - 1)/s$, where ϕ is Euler's totient function. Including the special case for $j = 1$ when all the $M_{k,j}$ are 1, this allows us to approximate integrals in up to dimension $d = 1{,}111$ if we use all the primitive polynomials of degree 13 or less. We then choose values of the $M_{k,j}$ so that we can generate Sobol sequences satisfying Property A in dimensions d up to 1,111. This is done by generating some values randomly, but these are subsequently modified so that the condition $\det(V_d) = 1$ (mod 2) is satisfied for all d up to 1,111. This process involves evaluating values of the $v_{k,j,1}$'s to obtain the matrix V_d and then evaluating the determinant of V_d. A more detailed discussion of this strategy is given in the next section. It is not difficult to produce values to generate Sobol's points for approximating integrals in even higher dimensions.

The following figures are the two-dimensional plots of high-dimensional Sobol sequences of Joe and Kuo with $d = 1{,}000$. It is compared with pseudorandom number

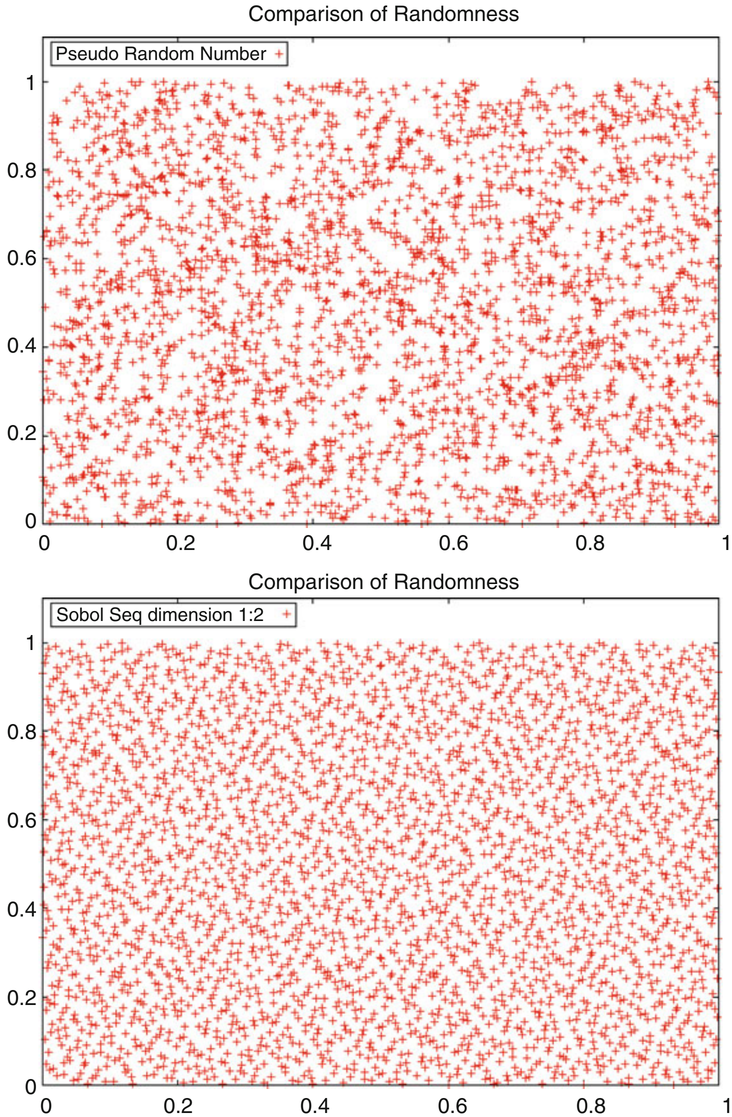


Fig. 49.7 Pseudorandom number plot comparing with quasi-random number of Sobol for dimensions 1 and 2

generation. The number of sampling points is 3,000. In Fig. 49.7 pseudorandom number is plotted in comparison with that of quasi-random number of Sobol. The leading dimensions 1 and 2 of Sobol sequences are used. The improvement is immense. In order to understand more of the nature of Sobol sequences, we chose prime dimensional numbers 499, 503, 991, and 997, respectively, as suggested by Joe

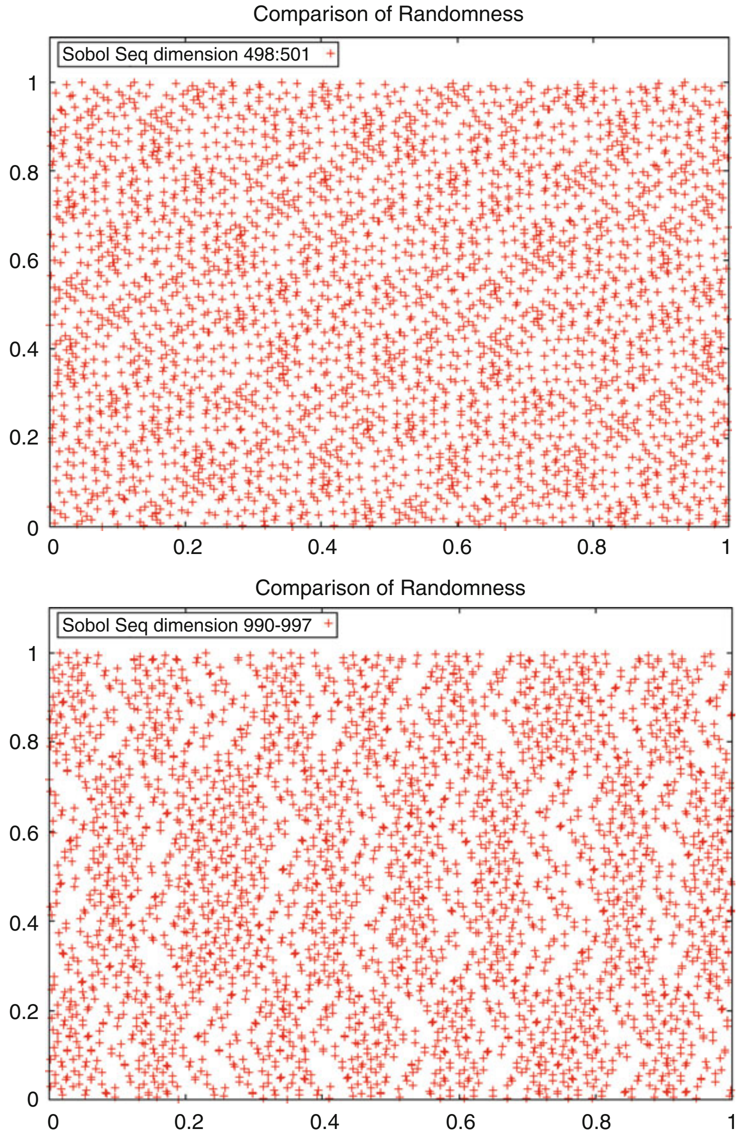


Fig. 49.8 Comparison of adjacent dimensions in quasi-random number Sobol sequence. The dimensions are chosen according to prime numbers. There is a high discrepancy found in higher dimensions of Sobol sequence modified by Joe and Kuo (2003)

and Kuo. The results are plotted in Figs. 49.8 and 49.9. It is found that there are stronger correlations between Sobol sequences of nonadjacent dimensions in the fashion of the dimensional comparison of their randomness. Larger numbers of sampling points, e.g., 10,000, are also tested and the patterns persist. It implied the

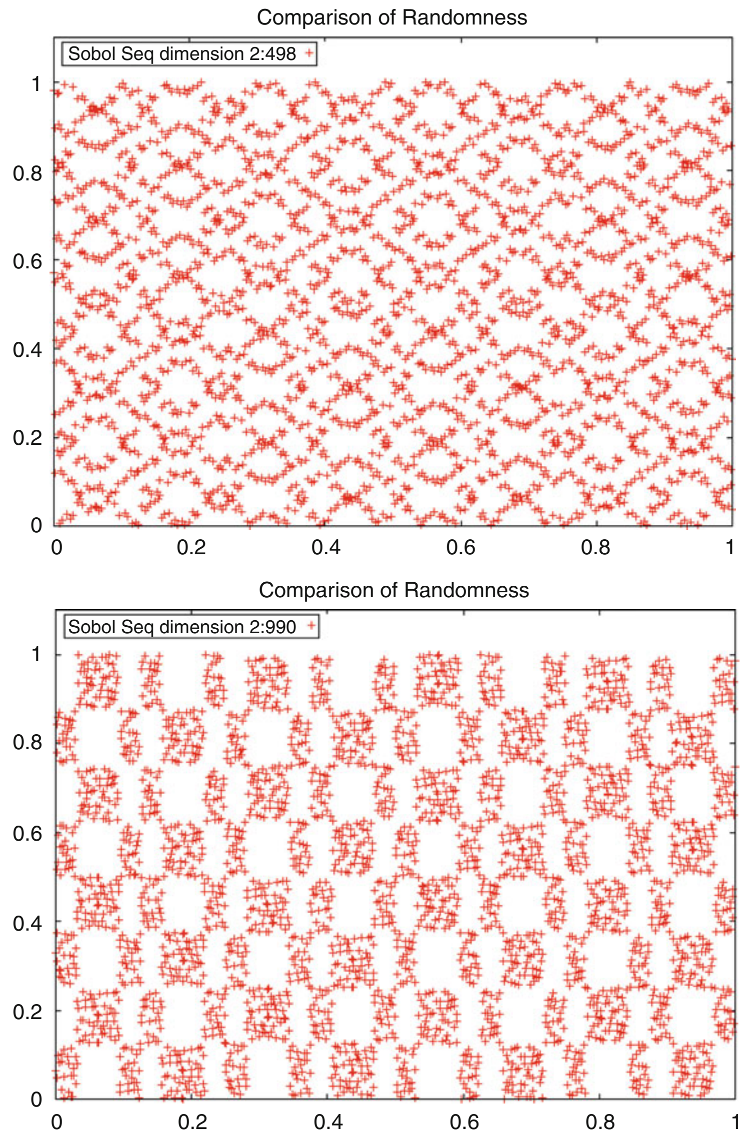


Fig. 49.9 Comparison of nonadjacent dimensions. High discrepancy is found in their correlations and forms clusters of islands in the distribution

violation of *idd* assumption and may incur problems in mass distribution and parallelism, in which each process the random number is generated independently without knowing what other processes are doing. The dependency may deteriorate the quality of randomness. Nevertheless, in our numerical experiments, there are no significant differences found thus far (Figs. 49.10 and 49.11).

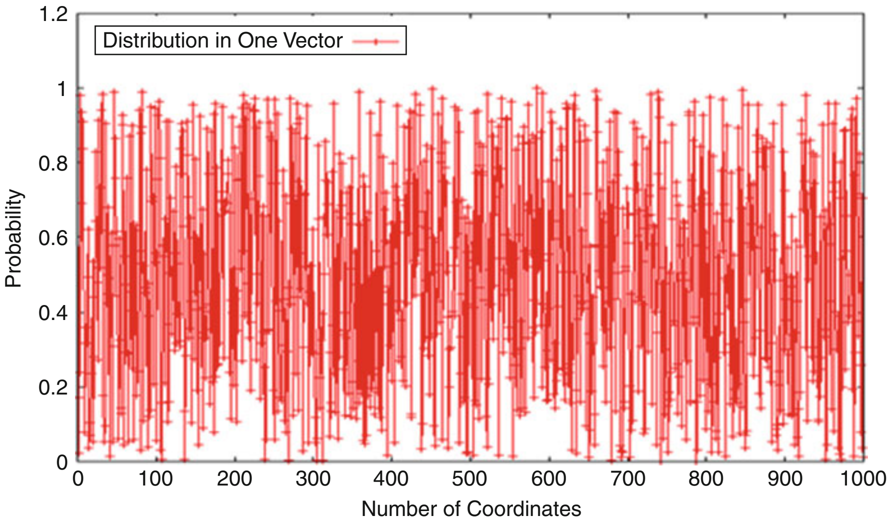


Fig. 49.10 The distribution of probability in directional vector $v_{i,j}$ of Sobol sequences at $i = 3{,}000$ with $j \in \{1, \cdots, 1000\}$. The mean of the distribution is 0.491357, which approaches the mean of the normal distribution 0.5

49.4 Case Study and Discussions

49.4.1 Case Study

49.4.1.1 Asian Options and Rainbow Options

To demonstrate what the grid computing can contribute to the financial service in a significant manner, two kinds of popular options, Asian options and rainbow options, are chosen for Monte Carlo pricing model. Asian options have payoffs that depend on the average price of the underlying asset such as stocks, commodities, or financial indices. However, there is no exact closed-form formula existed for these popular options. Rainbow options, also known as basket options, are referred to as an entire class of options which consist of more than one underlying asset. Rainbow options usually call or put on the best or worst of the underlying assets, or options which pay the best or worst of the assets. They are excellent tools for hedging risk of multiple assets. The rainbow options are therefore used for our bootstrap calculations of VaR.

49.4.1.2 Parallelization, Distribution, and Message Passing Interface (MPI)

MPI is a library specification for message passing, proposed and developed as a standard by a broadly based committee of vendors, developers, and users (Snir et al. 1996). MPI was designed for high performance on both massively parallel

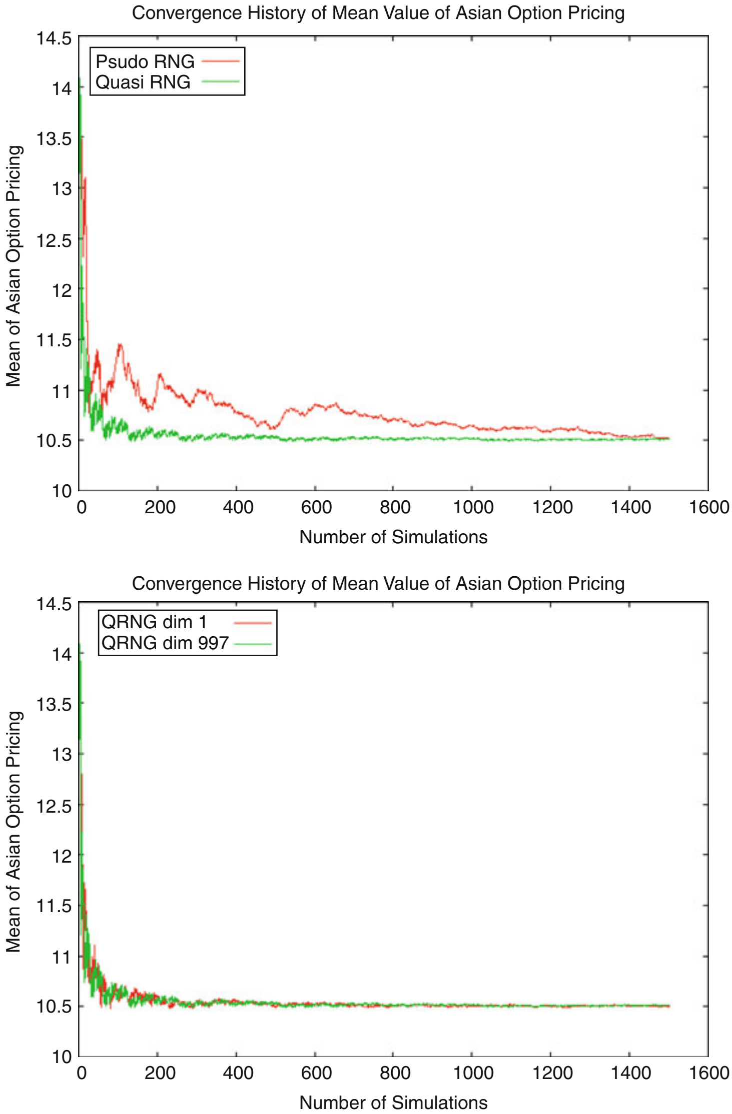


Fig. 49.11 The convergence history of the mean value of Asian option pricing with risk-free interest rate $r = 0.1$, underlying asset spot price $S = 100$, strike price $X = 100$, duration to maturity $T = 1$, and volatility $\sigma = 0.3$: The comparison is based on a single dimension of the extended high-dimensional Sobol sequences. The quasi-random number generator (QRNG) outperforms pseudorandom number generator. The test also is conducted to compare the convergence history between different dimensions in Sobol sequences and found that all perform consistently as shown in the *right* figure, in which the low dimension and high dimension are chosen for the comparison

machines and on distributed clusters. The MPI standard is nowadays widely accepted and used in the community of high-performance computing.

The basic MPI functions are point-to-point pair-wise message passing for send and for receive. Collective communications are also provided for ease of use as well as better performance. These communication methods, when used in supercomputers, do facilitate the parallelization of numerical methods that require both heavy compute cycles and stronger dependency between parallel processes.

Recent development of supercomputer, affected by the popularity of cluster computing in PCs, tends to be designed hierarchically scalable. Further extension of clusters of supercomputers can be regarded as initial concept of grids (see Sect. 49.2). The use of MPI is straightforward in this kind of hardware architecture and interlink of networks. There is always an obvious physical limitation in this architecture, which is also proportional to the limitation of the investment of governmental research funding. People tend to use mass distribution of computers, mostly PCs, which linked loosely in the Internet cloud. The terminology cloud is often used in networking community to show that in the Internet there is no specific network path from one computer to another. MPI working in such an environment is expected to be inefficient and unstable, e.g., high network latency induced packet lost in long-distance real-time communication. In the following sections three specific platforms, including local clusters, geographically distributed large clusters, and PC grids with ten of thousand of PCs connected in the cloud, will be used for the financial calculations to demonstrate benefits in using grids.

49.4.1.3 Empirical Study for Data Grid System

In order to demonstrate the usefulness of grid system, in particular in data-intensive application, the real market data are used, including daily from iShares (Morgan Stanley Capital International) MSCI Taiwan Index (ETF) and Taiwan Stock Exchange Center (TSEC) weighted index, extracted specifically from May 31, 2005, to May 31, 2008, and 30 days tick-by-tick trading data from Taiwan Futures Exchange Center (TAIFEX).

49.4.2 Grid Platforms Tests

The various grid platforms are carefully chosen to demonstrate the performance issue in finance services, which include a small diskless remote boot Linux (DBRL) PC clusters, large-scale and geographically widely distributed test-bed the Pacific Rim Applications and Grid Middleware Assembly (PRAGMA) compute grid, and a densely distributed at-home style PC grid, which resembles the clouding computing.

49.4.2.1 Diskless Remote Boot Linux (DRBL) Cluster

DRBL is an in-house program of National Center for High-Performance Computing (NCHC) and was an original software product developed by Steven Shiau and his group under the auspice of National Knowledge Innovation Grid (KING) of

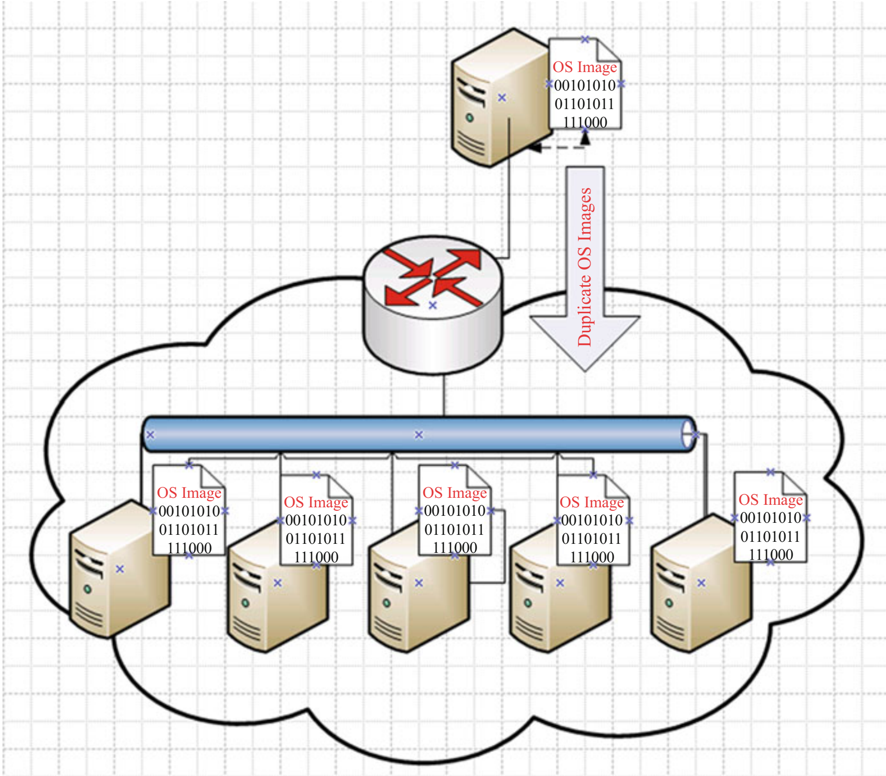


Fig. 49.12 The schematic of DRBL system: DRBL duplicates image files of operational system via network to the clients, in which the clients' original operational systems are untouched. Therefore, the clients are temporarily turned into dedicated compute resources, which also provide additional security to financial data

Taiwan. It was developed initially as a centralized system management tool aiming at small and median size of PC clusters. It is nowadays recognized internationally as one of the most advanced mass backup solutions. Here DRBL is used as an alternative scavenger for compute cycles of spare clusters. When needed, it converts systems of PC clusters into an aggregated and homogenous Linux system, simultaneously with a mass backup of the original systems, and recovered back the original systems once the need was satisfied. The most popular use is to convert a PC classroom into a compute Linux cluster. In such a case, compute cycles of the clusters can be fully exploited. In a grid environment, this is a perfect case to resources scaleup when in contingent need and once the situation relieved resources will be released correspondingly. Such a dynamic feature can be beneficial for the financial services.

The schematic of DRBL system can be shown in Fig. 49.12, where DRBL duplicates image files of an operational system, e.g., Linux kernel, via network to

Table 49.1 Comparison of performance between DRBL-based PC platform with 32 nodes, FORMOSA II of NCHC with a batch job of 32 nodes and IBM Cluster 1350 with a batch job of 32 nodes. The PCs are 20 XEON 2.6 GHz and 4GB RAM

	DRBL cluster	FORMOSA II	IBM cluster 1350
Asian option pricing (AOP)	81	49	24
Rainbow option pricing (ROP)	329	197	98
VaR calculation (based on ROP with bootstrap)	322	194	97

Unit: micro-secs per Monte Carlo path. Averaged from 1000,000 Monte Carlo paths

Table 49.2 Comparison of speedup ratios based on the calculations in Table 49.1

	DRBL cluster	FORMOSA II	IBM cluster 1350
Asian option pricing (AOP)	28.68	29.46	30.16
Rainbow option pricing (ROP)	27.31	28.59	29.34
VaR calculation (based on ROP with bootstrap)	27.20	28.10	29.10

Unit: speedup ratio: CPU(nonparallel single node)/CPU(parallel single node)

the clients. The clients' original operational systems are not used. The clients are therefore temporarily turned into dedicated compute resources, which also provide additional security to financial data.

The test case here involves Monte Carlo simulations on Asian option pricing, on rainbow option pricing and on a bootstrap VaR calculation, respectively (see Sect. 49.3.2). The market parameters are given as risk-free interest rate $r = 0.1$, underlying asset spot price $S = 100$, strike price $X = 100$, duration to maturity $T = 125$, and volatility $\sigma = 0.3$. For the rainbow options a linear weighted combination of 4 underlying assets is assumed with underlying prices of $S_i = 100, 110, 120,$ and 130 and the corresponding weightings of volatility $\sigma_i = 0.3,$ 0.4, 0.5, and 0.6. The correlation matrix is taken to be

$$\rho_{ij} = \begin{pmatrix} 0.5 & 0.4 & 0.5 \\ 0.4 & 0.3 & 0.4 \\ 0.5 & 0.4 & 0.6 \end{pmatrix}$$

The calculation of the VaR uses the same 4-dimensional rainbow options with additional expectation of return 0.07, 0.08, 0.09, and 0.10, respectively. They are calculated in DRBL cluster as well as benchmarked with two cluster-based supercomputers in NCHC. The results are shown in Tables 49.1 and 49.2.

In Table 49.1, instead of giving a total wall clock time of the calculation with some given numbers of Monte Carlo simulations or paths, a more useful averaged single Monte Carlo simulation based on 1,000,000 simulations is used to demonstrate the performance when different system architectures are used. The results show that the traditional big irons, i.e., supercomputers, still outperform the cluster.

Yet, considering there is no extra cost invested in the computing resources and still obtains compute cycle in a sizable manner, the approach is appealing to be further developed in to a fully operational production system.

49.4.2.2 Pacific Rim Applications and Grid Middleware Assembly (PRAGMA) Grid

The Pacific Rim Applications and Grid Middleware Assembly (PRAGMA), founded in 2002, is an open international organization that focuses on a variety of practical issues of building international scientific collaborations in a number of application areas. PRAGMA grid is established by the resources and data computing working group as a global grid test-bed for benchmarking the interoperability of grid middleware and the usability and productivity of grids. The PRAGMA grid consists of physical resources as well as system administration supports from 29 institutions across 5 continents and 14 countries. It is an instantiation of a useful, interoperable, and consistently available grid system that is neither dictated by the needs of a single science domain nor funded by a single national agency. It does not have uniform yet robust infrastructure management and supports a wide range of scientific applications. The software stack of the system is shown in (Fig. 49.13).

The PRAGMA grid successfully tackles the issues of distance and time zone differences among sites, lack of infrastructure tools for heterogeneous global grid, nonuniform system and network environments, and diverse application requirements. For more technical details both in theory and practice we refer to (Abramson et al. 2006)

Following the similar test case in Sect. 49.4.2.2, but extended the platform with a collection of clusters across institute boundaries, the common job submission is executed via a homogeneous middleware Globus Toolkit. We demonstrate the usefulness of the platform by grouping compute resources across national boundaries and still achieve good performance. The results are shown in Tables 49.3 and 49.4.

49.4.2.3 At-Home Style PC Grid

With the continued penetration of personal computers and the remarkable improvement of CPU processing speed, 80–90 % of most PCs' processing power is untapped, according to a study. This does not mean that many PCs remain turned off, but that the capacity of the CPU, the brain of the PC, is not fully utilized. In case the CPU is more extensively used when a task requiring an enormous number of operations, such as three-dimensional graphical processing, is assigned, it sits idle most of the time during word processing and Internet browsing because CPU processing speeds are much faster than the speeds of input from the keyboard or the communication line.

This fact led to the idea of virtually gathering the power of idle CPUs to use as a computer resource. In other words, this means networking numerous computers to make them work like a single high-performance computer and assigning complex processing tasks to it. The assigned task will be divided into a myriad of small tasks

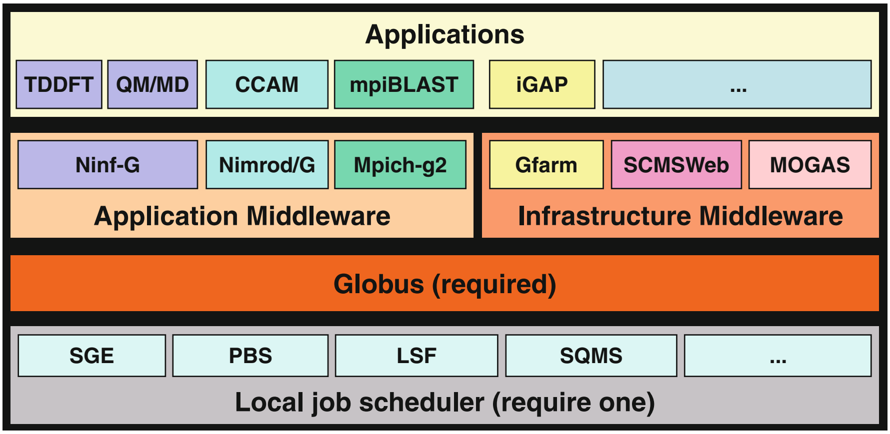


Fig. 49.13 Software stack developed in the PRAGMA grid

Table 49.3 Comparison of performance between Group A, which consists of 13 nodes from NCHC and 15 from UCSD, and Group B, which consists of 122 nodes collectively from UCSD, AIST, NCHC, and Osaka University. The details of resources are referred to (http://pragma-goc.rocksclusters.org/pragma-doc/resources.html)

	Group A	Group B
Asian option pricing (AOP)	68	56
Rainbow option pricing (ROP)	278	234
VaR calculation (based on ROP with bootstrap)	271	229

Unit: seconds/per Monte Carlo path. Averaging from 1000,000 Monte Carlo paths

Table 49.4 Comparison of speedup ratios based on the calculations in Table 49.3

	Group A	Group B
Asian option pricing (AOP)	25.24	107.58
Rainbow option pricing (ROP)	25.81	110.34
VaR calculation (based on ROP with bootstrap)	25.63	110.09

Unit: seconds/per Monte Carlo path. CPU(nonparallel single node)/CPU(parallel single node)

and allocated to individual computers on the grid-like network. Even with increasingly faster CPUs, the power of PCs is not comparable to those of supercomputers, but in a networked environment where individual PCs simply process complex task in parallel, PCs can deliver surprisingly high performance. This is the core concept of CP grid computing, and it came into a reality several years ago (Table 49.5).

The commoditization and the increased processing speed of PCs lead the growth of idle CPU power. This will facilitate the construction of PC grid computing system along with improvement of communication environment by broadband connectivity (Chen et al. 2006). In practice, a PC grid platform Korea@Home (or K@H) is used in our study. Its architecture is shown in Fig. 49.14. It is based on MS Windows. Asian option pricing is used to demonstrate the performance of the

Table 49.5 Summary of the case for the PC grid calculations

Asian option pricing	Statistics
Number of Monte Carlo path	1,000,000 × 10.000
The running period (1) (wall clock time)	28 h 51 m *(104,911 s)*
Number of jobs	10,000
CPU time per job	28 ~ 30 s
Total running time (2) (wall clock time)	4 days 14 h 31 m (397,919 s)
Speedup ratio $\frac{(2)}{(1)}$	3.79

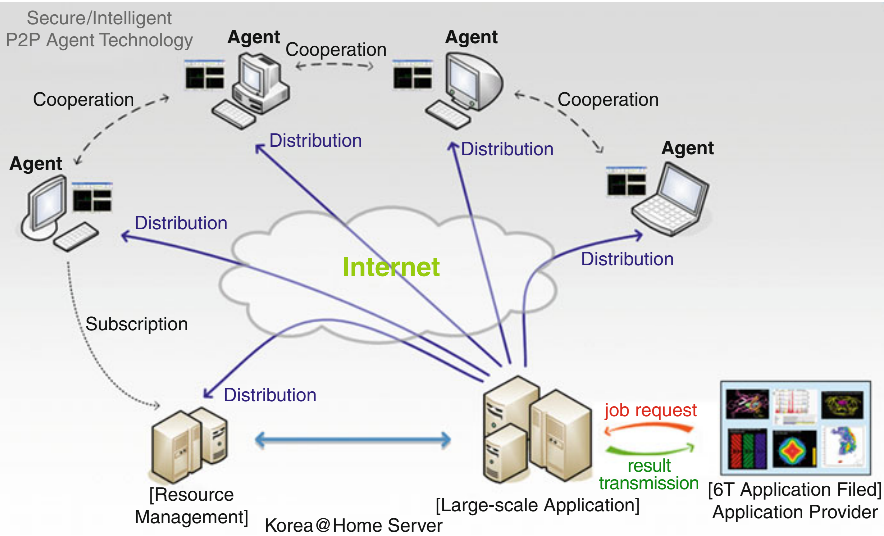


Fig. 49.14 The architecture of Korea@Home, a specific @Home style PC grid used in our case study (Jun-Weon Yoon 2008)

system, in which the number of iterations is taken to be 1,000,000. The duration to maturity is further divided by 10,000 periods, which push the system to run on the mass parallel system of K@H. To tackle this scale, or even larger scale for all kinds of possible scenarios in real trading practice, an off-line distributed and parallel approach is adopted. The total number of Monte Carlo simulation is 1,000,000 × 10,000. It was divided into 10,000 jobs and each job consists of 1,000,000 Monte Carlo simulations. The market parameters are given as above.

The results demonstrated here are not as good as expected (see Table 49.1). It shows that the speedup ratio is only 3.79. In this test, K@H further divides the jobs into ten groups. Each group was send and run in a sequential fashion, which causes the low speedup ratio. However, if one looks into the executed CPU time for each job, our assumption is still valid. We simulated the result in a small cluster with the same scenario and Monte Carlo paths. The result shows 90 % speedup can be easily achieved.

49.4.2.4 RBNB Data Grid

RBNB DataTurbine in market data streaming is implemented here (see Fig. 49.15), in which the real data from iShares MSCI Taiwan Index (ETF) and TSEC weighted index from May 31, 2005, to May 31, 2008, and 30 days tick-by-tick trading data from Taiwan Futures Exchange Center (TAIFEX) are used and open in different file-based online channels from the RBNB DataTurbine. The purpose is to dynamically manage the high-frequency market data and connect the data with analysis applications on the fly. The system implemented up-to-date large-scale dynamic and static data management.

49.5 Conclusions

Securities trading is one of the few business activities where a few seconds processing delay can cost a company big fortune. The growing competitive in the market exacerbates the situation and pushes further towards instantaneous trading even in split second. The key lies on the performance of the underlying information system. Following the computing evolution in financial services, it was a centralized process to begin with and gradually decentralized into a distribution of actual application logic across service networks. Financial services have tradition of doing most of its heavy lifting financial analysis in overnight batch cycles. However, in securities trading it cannot satisfy the need due to its ad hoc nature and requirement of immediate response. A new computing paradigm, grid computing, aiming at virtualizing scale-up distributed computing resources, is well suited to the challenge posed by the capital market practices.

In this study we revisit the theoretical background of how performance will affect the market competition. The core concept lies on information asymmetry. Due to the advance of IT, even in split second, it will be a matter of win or lose in real market practice. After establishing the motivation, we review recent grid development specifically used for finance service. Monte Carlo simulations are chosen not only because of its popularity in real world but also because of its nature so-called “fine grain” or mass parallelism approach. The success of Monte Carlo simulations lies on better random number generators. The well-recognized Sobol sequences as a quasi-random number generator are carefully studied to ensure the quality of Monte Carlo simulations when employed for mass parallelism. Then some popular basic option pricing models, collectively Asian option pricing, rainbow option pricing, and VaR calculation with constant market parameters, are introduced as drivers to introduce more details of grids for better finance service. Finally, we test various grid platforms, based on the methodology of mass parallelism and mass distribution, with the drivers. The real market data are also used, but at this stage they are only used to demonstrate the dynamic data management, in which grids can offer better.

During this study, we encountered system architect Koschnick from Zürcher Kantonalbank of Switzerland (Koschnick 2008). Coincidentally, the system they plan to migrate from big irons is the similar system to that of DRBL with additional

Fig. 49.15 RBNB DataTurbine streaming open for data channels of iShares MSCI Taiwan Index (ETF) and TSEC weighted index from May 31, 2005, to May 31, 2008, and 30 days tick-by-tick trading data from Taiwan Futures Exchange Center (TAIFEX) (Real data plot in collaboration with Strandell et al. 2007)

virtual local area networks (VLAN) for security. The system is used for overnight batch job as well as real-time trading practice. It is confirmed that for the years to come, financial services providers will adopt more grid or grid-based technology to enhance their competitiveness.

Our future work will be following the current work, continuously using the current grid platforms and extending them to the use high-frequency real market data. Along the track of this development, we will also develop sophisticated Monte Carlo-based option pricing and risk management based on tick-by-tick daily market information.

References

Abramson, D., et al. (2006). Deploying scientific applications to the PRAGMA grid testbed: strategies and lessons. In *Sixth IEEE international symposium on cluster computing and the grid*, Singapore, pp. 241–248.

Ahmad, K., Gillam, L., & Cheng, D. (2005). Textual and quantitative analysis: Towards a new, e-mediated social science. In *Proceedings of the 1st international conference on e-Social science*, Manchester, pp. 22–24.

Almond, J., & Snelling, D. (1999). UNICORE: Uniform access to supercomputing as an element of electronic commerce. *Future Generation Computer Systems, 15*(5), 539–548.

Amad, K., et al. (2004). Financial information grid –an ESRC e-Social social science pilot. In *Proceedings of the third UK e-Science programme all-hands meeting*, Nottingham.

Antonov, I. A., & Saleev, V.M. (1979). An economic method of computing LP[tau]-sequences. USSR Computational Mathematics and Mathematical Physics, 19(1), 252–256.

Atkins, D.E., et al. (2003). *Revolutionizing science and engineering through cyberinfrastructure*. Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure. National Science Foundation.

Bell, D., & Ludwig, A. S. (2005). Grid service discovery in the financial markets sector. In *7th international conference on information technology interfaces*, Cavtat, Croatia.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy, 81*, 637–654.

Bratley, P., & Fox, B. L. (1988). Algorithm 659: Implementing Sobol's quasirandom sequence generator. *ACM Transactions on Mathematical Software, 14*, 88–100.

Bughin, J., Chui, M., & Manyika, J. (2010). Cloud, big data and smart assets: Ten tech-enabled business trends to watch. *McKinsey Quarterly August*, 7–8.

Chang, F., Dean, J., Ghemawat, S., Hsieh, W.C., Wallach, D.A., Burrows, M., Chandra, M., Fikes, A., Gruber, R.E. (2006). Big table: A distributed storage system for structured data. In *OSDI 06: Proceedings of the 7th USENIX Symposium on Operating Systems Design and Implementation*, Seattle, WA, p. 15-15.

Cheng, J., & Druzdzel, M. J. (2000). *AIS-BN: An Adaptive importance sampling algorithm for evidential reasoning in large bayesian networks*. Journal of Artificial Intelligence Research, 13, 155–188.

Chen, G., Thulasiraman, P., & Thulasiram, R.K. (2006). *Distributed quasi-monte carlo algorithm for option pricing on HNOWs using mpC*. In *Proceedings of the 39th annual simulation symposium (ANSS'06)*, Huntsville.

Cizek, P., Härdle, W., & Weron, R. (2011). *Statistical tools for finance and insurance*. Berlin: Springer.

Dean, J., & Ghemawat, S. (2004). Map reduce: Simpli_ed data processing on large clusters. In *OSDI 04*, San Francisco, CA, pp. 137–150.

Dowd, K. (2002). *Measuring market risk*. Chichester/England/New York: Wiley.

Efon, B. (1981). Censored data and the bootstrap. *Journal of the American Statistical Association, 76*(374), 312–319.

Efron, B., & Tibshirani, R. (1986). Bootstrap methods for standard errors, confidence intervals, and other measures of statistical accuracy. *Statistical Science, 1*(1), 54–75.

Faure, H. (1982). Discrepance de Suites Associees a un Systeme de Numeration (en Dimension s). *Acta Arithmetica, 41*, 337–351.

Foster, I., & Kessleman, C. (2004). *The grid: Blueprint for a new computing infrastructure* (2nd ed.). Amsterdam: Morgan Kaufmann.

Foster, I., Zhao, Y., Raicu, I., & Lu, S. (2008). Cloud computing and grid computing 360-degree compared. In *Grid computing environments workshop*, Austin, TX, pp. 1–10.

Galanti, S., & Jung, A. (1997). Low-discrepancy sequences: Monte Carlo simulation of option prices. *Journal of Derivatives, 5*(1), 63–83.

Ghoshal, D., Canon, R. S., & Ramakrishnan, L. (2011). I/O performance of virtualized cloud environments. In *The second international workshop on data intensive computing in the clouds*, Seattle, WA, pp 71–80.

Gillam, L., Ahmad, K., & Dear, G. (2005). Grid-enabling social scientists: The Fingrid infrastructure. In *Proceedings of the 1st international conference on e-social science*, Manchester, pp. 22–24.

Griffiths, P., & Remenyi, D. (2003). Information technology in financial services: A model for value creation. *Electronic Journal of Information Systems Evaluation, 6*(2), 107–116.

Grimshaw, A. S., & Wulf, W. A. (1997). The legion vision of a worldwide virtual computer. *Communications of the ACM, 40*(1), 39–45.

Halton, J. H. (1960). On the efficiency of certain quasirandom sequences of points in evaluating multidimensional integrals. *Numerische Mathematik, 2*, 84–90.

Hauswald, R., & Marquez, R. (2003). Information technology and financial services competition. *The Review of Financial Studies, 16*(3), 921–948.

Hull, J. C. (2003). *Options, futures and other derivatives* (5th ed.). Upper Saddle River: Prentice Hall.

Joe, S., & Kuo, F. Y. (2003). Remark on algorithm 659: Implementing Sobol's quasirandom sequence generator. *ACM Transactions on Mathematical Software, 29*(1), 49–57.

Korpela, E., et al. (2001). SETI@home-massively distributed computing for SETI. *Computing in Science and Engineering, 3*(1), 78–83.

Koschnick, T. (2008). Cluster & security: Separating application environments. In *Int'l Supercomputing conference and exhibition*, Dresden, pp. 17–20.

Kuipers, L., & Niederreiter, H. (1974). *Uniform distribution of sequence*. New York: Wiley.

Lauret, J., Walker, M., Goasguen, M., & Hajdu, L. (2010). From grid to cloud, the STAR experience. In *SciDAC 2010 proceedings*, Chattanooga, TN.

Leto, A., et al. (2005). EGRID Project: Experience report on the implementation of a grid infrastructure for the analysis of financial data. In *Proceedings of the international conference on information technology: Coding and computing*, Las Vegas, NV.

Litzkow, M. J., Livny, M., & Mutka, M. W. (1988). Condor-a hunter of idle workstations. In *8th international conference on distributed computing systems, proceeding*, San Jose, CA, pp. 104–111.

Mell, P., Grance, T. (2011). *The NIST definition of cloud computing* (pp. 800-145). Special Publication.

Milojicic, S. D., et al. (2002). *Peer-to-peer computing* (HL laboratories research report).

Niederreiter, H. (1992). Random number generation and Quasi-Monte Carlo methods. In *CBMS-NSF regional conference series in appl. math.* Philadelphia, PA.

Paskov, S. H., & Traub, J. F. (1995). Faster valuation of financial derivatives. *Journal of Portfolio Management, 22*(1), 113–120.

Phillips, A., et al. (1984). *Effects of information technology on financial services systems*. Washington, D. C.: U.S. Congress, Office of Technology Assessment, OTA-CIT-202.

Ramakrishnan, L., Canon, R. S., Muriki, K., Sakrejda, I., & Wright, N. J. (2011). Evaluating interconnect and virtualization performance for high performance computing. In *Proceedings of 2nd international workshop on performance modeling, benchmarking and simulation of high performance computing systems*, Seattle, WA, pp. 1–2.

Schmidt, A. B. (2005). *Quantitative finance for physicists: An introduction*. Academic Press Advanced Finance Series, Elsevier, Oxford.

Smarr, L., & Catlett, C. (1992). Metacomputing. *Communications of the ACM archive, 35*(6), 44–52.

Snir, M., et al. (1996). *MPI: The complete reference*. The MIT Press, Cambridge, Massachusetts.

Sobol, I. M. (1967). On the distribution of points in a cube and the approximate evaluation of integers. U.S.S.R. *Computational Mathematics and Mathematical Physics, 7*(4), 86–112.

Sobol, I. M., & Levitan, Y. L. (1976). *The production of points uniformly distributed in a multidimensional cube*. Tech. Rep. 40, Institute of Applied Mathematics, USSR Academy of Sciences (in Russian).

Stokes-Rees, I., et al. (2007). *Multi-cluster parallel job submission: Experiences with Monte Carlo simulations for computational finance on Grid5000* (IRIS PicsouGrid Project Report).

Strandell, E., Tilak, S., Chou, H. M., Wang, Y. T., Lin, F. P., Arzberger, P. Fountain, T. (2007). Data management at kenting's underwater ecological observatory. In *Proceedings of the third International Conference on Intelligent Sensors, Sensor Networks and Information Processing*, Melbourne, Australia.

Tanaka, S. (2003). *Joint development project applies grid computing technology to financial risk management*. Nikkei Computer, special section on grid computing.

Tilak, S., Hubbard, P. Miller, M., & Fountain, T. (2007). The ring buffer network bus (RBNB) data turbine streaming data middleware for environmental observing systems. In *Proceedings of third IEEE international conference on e-science and grid computing*, Bangalore, India, pp. 125–133.

Weron, R. (2004). *Handbook of computational statistics*. Berlin: Springer.

Yelick, K., Coghlan, S., Draney, B., Canon, R. S. (2011). *The magellan report on cloud computing for science*. ASCR, DOE. Washington, DC.

Yoon, J. (2008) *Korea@Home (PC Grid project in Korea)*. GridAsia 2008: Singapore.

Long-Run Stock Return and the Statistical Inference

50

Yanzhi Wang

Contents

Abstract

This article introduces the long-run stock return methodologies and their statistical inference. The long-run stock return is usually computed by using a holding strategy more than 1 year but up to 5 years. Two categories of long-run return methods are illustrated in this article: the event-time approach and calendar-time approach. The event-time approach includes cumulative abnormal return, buy-and-hold abnormal return, and abnormal returns around earnings announcements. In former two methods, it is recommended to apply the empirical distribution (from the bootstrapping method) to examine the statistical inference, whereas the last one uses classical *t*-test. In addition, the benchmark selections in the long-run return literature are introduced. Moreover, the calendar-time approach contains mean monthly abnormal return, factor models, and Ibbotson's

Y. Wang
Yuan Ze University, Taiwan

Department of Finance, College of Management, National Taiwan University, Taipei, Taiwan
e-mail: yzwang@ntu.edu.tw

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_50,

RATS, which could be tested by time-series volatility. Generally, calendar-time approach is more prevailing due to its robustness, yet event-time method is still popular for its ease of implementation in the real world.

Keywords
Long-run stock return • Buy-and-hold return • Factor model • Event time • Calendar time • Cumulative abnormal • Return • Ibbotson's RATS • Conditional market model • Bootstrap • Zero-investment portfolio

50.1 Introduction

The long-run stock return has been an important facet of the stock performance of firms since the 1990s. Ritter (1991) starts the line of long-run return studies by investigating the initial public offering (IPO) cases. He finds that the post-IPO stock performance is poor in the long run. His paper is then followed by many scholars who study other important corporate events and asset pricing anomalies.[1] Fama (1998) reviews these papers which engage in long-run event studies. Hence, the long-run return method is nowadays a standard way for the stock performance of firms.

I review the methodologies about the long-run stock abnormal return in this article. Since the mid-1990s, some papers started to be aware of the properties of long-term stock performance and suggested various ways to calculate the long-run abnormal return. Therefore, Barber and Lyon (1997), Kothari and Warner (1997), and Lyon et al. (1999) review and compare long-run return methods in the late 1990s. Based on these methodology papers, I update recent developments about the new methods that are not mentioned in these two papers, such as the earnings announcement abnormal return applications (La Porta et al. 1997), conditional market model method (Eberhart et al. 2004; Petkova and Zhang 2005), Ibbotson's RATS (Ibbotson 1975; Agrawal et al. 1992; Peyer and Vermaelen 2009), and zero-investment portfolio method (Daniel and Titman 1997; Eberhart et al. 2004). I also make a clearer categorization on these long-run stock return methodologies given that the long-run return methodologies nowadays are much more mature than in the late 1990s.

Two main categories of the long-run stock performance illustrated in this paper are event-time and calendar-time approaches. The event-time approach includes cumulative abnormal return (CAR), buy-and-hold abnormal return (BHAR), rebalanced buy-and-hold abnormal return (RBHAR), and abnormal returns around

[1]For corporate events, papers have studied seasoned equity offerings, mergers, dividend initiations and omissions, quarterly earnings announcements, share repurchases, proxy flights, stock splits and spinoffs, and other corporate events for their long-run stock performance. For asset pricing anomalies, papers investigate value premium, momentum profit, research and development profit, accrual effect, asset growth, net share issuance, and other anomalies in terms of the long-term impact.

earnings announcements. The calendar-time approach contains mean monthly abnormal return (MMAR), Ibbotson's RATS, and factor models. The factor model has various specifications, in which Fama and French (1993) three-factor model and Carhart (1997) four-factor model are most prevailing two.

In addition, to estimate the *abnormal* return, the benchmark is vital because a misspecification of the matching procedure yields an incorrect statistical inference and leads to the statistical type I or type II error. For event-time approach, either a matching firm or a matching portfolio can control for the firm characteristic effect (Daniel and Titman 1997). Yet, the more relevant problems are how we should select the weighting scheme and what the matching criterion should be. On one hand, the selection of the weighting scheme determines the degree of the abnormal return and involves in different magnitude of transaction costs from the rebalancing problem. On the other hand, different matching methods may result in a superior or an inferior statistical inference, particularly the skewness of the long-run return estimation.

For MMAR of the calendar-time approach, the benchmark selection problem is similar to what we may face in the event-time approach. Yet, the family of factor models has a simpler benchmark problem because the only issue is the factor model specification (e.g., a three-factor or a four-factor model). The same situation applies the Ibbotson's RATS that we only pay attention to the market model setting, namely, adding more independent variables in the market model. Nevertheless, few complicated modifications for the factor models are employed in the long-run return studies, such as conditional market model and the zero-investment portfolio method in factor model. The former deals with the time-varying systematic risk loadings in the market model, whereas the later is to combine the factor model method with the matching methods in the event-time approach. In general, the robustness of calendar-time approach raises its popularity in recent finance studies.

The rest of this paper is organized as follows. Section 50.2 introduces the long-run return estimation in event-time approach. Section 50.3 states the calendar-time method applied in the long-run stock return estimation. Finally, Sect. 50.4 concludes this review paper.

50.2 Long-Run Return Estimation in Event-Time Approach

To estimate the long-run stock return, it is nature to identify an event to track the stock performance of a firm. Figure 50.1a shows a general time line for the long-run return estimation. For any specific event and its event day noted by day 0, we can start a holding strategy by purchasing the stock from event day 0 to 251th day (for 1-year return), to 503th day (for 2-year return), 755th day (for 3-year return), to 1,007th day (for 4-year return), or to 1,259th day (for 5-year return). Papers usually compute the long-run stock return up to 5 years (Ritter 1991). As indicated in Fig. 50.1a, an accounting reporting lag is required between the event date and the previous fiscal year-end with at least 4 months.

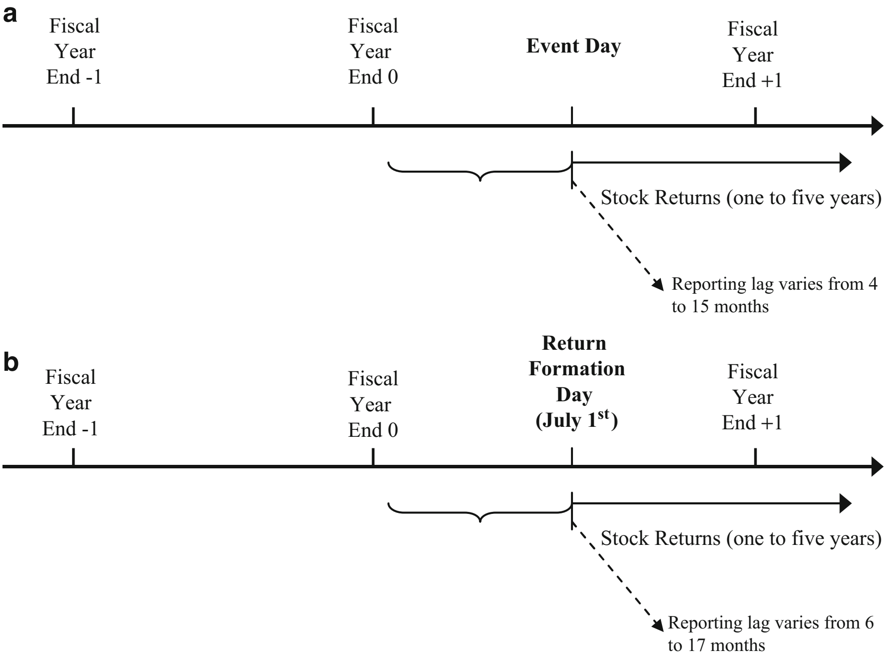


Fig. 50.1 (**a**) time line of general event study. (**b**) time line of asset pricing study

For the asset pricing study, we usually do not have a specific event day for tracking the compounding daily returns. As suggested in Fama and French (1992, 1993), a general investment strategy starts the return formation from 1 July of each year by exerting the accounting information in the previous fiscal year-end with at least 6 months reporting lag. Figure 50.1b shows the time line for asset pricing studies. Because asset pricing papers start from 1st July, the long-run returns are estimated based on compounded monthly returns. Basically, 1-year return includes 12 monthly returns, while 5-year return is computed upon 60 monthly returns.

50.2.1 Return Estimations: CAR, BHAR, and Earnings Announcement Returns

The cumulative abnormal return (CAR) is estimated as follows. For a given benchmark $E(R_{i,t})$, the abnormal return $AR_{i,t} = R_{i,t} - E(R_{i,t})$. The average of estimations for CAR is

$$\overline{CAR} = \sum_{t=1}^{T} \frac{\sum_{i=1}^{N} AR_{i,t}}{N}, \tag{50.1}$$

in which the N_t is for number of observations at time t and T for the holding period (252 for a year if the return is computed based on daily returns; 12 for a year if the return is computed based on monthly returns).

The buy-and-hold abnormal return is defined as

$$\overline{BHAR} = \sum_{i=1}^{N} \frac{\left[\prod_{t=1}^{T} \left(1 + R_{i,t}\right)\right] - 1}{N} - \sum_{i=1}^{N} \frac{\left[\prod_{t=1}^{T} \left(1 + E\left(R_{i,t}\right)\right)\right] - 1}{N}. \tag{50.2}$$

Thus, we can use classical t-test or other methods (will be illustrated in later section) to carry out the statistical inference for CAR and BHAR. One modification on BHAR is the rebalanced buy-and-hold abnormal return (RBHAR), which comes from the combination of rebalanced return (R^{reb}) and BHAR. Given the rebalanced return as

$$R^{\text{reb}} = \prod_{t=1}^{T} \left(1 + \frac{\sum_{i=1}^{N} R_{i,t}}{N} \right) - 1, \tag{50.3}$$

the RBHAR is to calculate buy-and-hold return for a certain period (e.g., a year) and rebalance the portfolio equally for every specific period. Taking the 1-year rebalanced BHAR as the example, we should compute 1-year BHAR for each event year (i.e., first to fifth event year) and obtain the average BHAR for every event year. Finally, we obtain compounding return for this average rebalanced BHAR with yearly rebalancing. Because BHAR has inflated compounding return, which results in many outliers in the long-run return estimation, RBHAR with rebalancing every year is able to reduce the impact from extreme values (Ikenberry et al. 1995; Chan et al. 2010). In general, RBHAR could be described as

$$\begin{aligned} \overline{RBHAR} = &\prod_{\text{year1}}^{\text{year5}} \left\{ 1 + \sum_{i=1}^{N} \frac{\left[\prod_{j=1}^{T} \left(1 + R_{i,j}\right)\right] - 1}{N} \right\} \\ &- \prod_{\text{year1}}^{\text{year5}} \left\{ 1 + \sum_{i=1}^{N} \frac{\left[\prod_{j=1}^{T} \left(1 + E\left(R_{i,j}\right)\right)\right] - 1}{N} \right\} \end{aligned} \tag{50.4}$$

Note that T stands for 252 days as an event year, and year 1 to year 5 represent the first event year to the fifth event year.

Distinct from CAR and BHAR, the quarterly earnings announcement return is the estimate of the long-term stock performance by successive short-term quarterly announcement returns (La Porta et al. 1997; Denis and Sarin 2001; Chan et al. 2004). In other words, we can use the 2-day or 3-day abnormal returns around quarterly announcement dates, and the successive earnings surprises should be aligned with the long-run stock abnormal return. The benchmark selection determines the long-run abnormal return estimation because of its statistical property (will be discussed in next section), so the results may be changed when benchmark settings are altered. This sensitive outcome driven from the benchmark problem leads to an inconclusive long-run abnormal return. Yet, the short-run return is not sensitive to the selection of benchmarks. For example, given 10 % and 20 % of market index returns, the 3-day market index returns are expected to be 0.12 % and 0.24 % only. When an event occurs, the 3-day announcement return could be generally above 1 %, which significantly exceeds any selected benchmark returns. Therefore, if a corporate event is followed by profitability improvements that are not observed by investors, then there should be successive earnings surprises following the event date, and the short-term announcement abnormal returns around the earnings announcement dates should be positive. To capture the long-run abnormal return, papers usually study 12–20 quarters (for 3–5 years) for the quarterly earnings announcement abnormal return. Generally, the quarterly earnings announcement abnormal returns capture about 25–40 % of the long-run stock return of a firm (Bernard and Thomson 1989; Sloan 1996).

50.2.2 Benchmark Problem

Long-run stock return has some issues regarding the calculation and testing when we select the benchmarks for the expected return of the firm. Namely, the conventional methodologies for the long-run stock return might be biased if the chosen matching procedure is inappropriate. First, some long-term return measures have the rebalancing and new-listing biases problems. Second, long-run stock return is positive skewed, thus the traditional t-test is inappropriate for the long-run return. Although I will review the skewness-adjusted t-statistics and empirical p-value in the next section, matching firm method is another way to alleviate the skewness concern in the long-run return studies.

The first benchmark for the expected return is the CRSP equal-weighted index return, which includes whole stocks in CRSP database and computes the simple average of stock returns. However, this approach involves rebalancing problem, which ignores the transaction costs from broker's fee and tax to the government. To maintain the weights on stocks equally, investors must sell profitable stocks and buy stocks with loss. This rebalancing leads to huge transaction costs that are not considered in the CRSP equal-weighted index return, making the abnormal return underestimated. In fact, CAR per se also has this rebalancing problem because of its average in cross section.

The second benchmark is the CRSP value-weighted index return, which also comes from CRSP database. This index return uses firm value (equal to the price timing shares outstanding) as the weighted scheme to compute the weighted average of stock returns. Most importantly, it is a rebalance-free benchmark and accordingly has no transaction cost during the holding period (except the beginning and end). Therefore, recent papers tend to use CRSP value-weighted index return instead of CRSP equal-weighted index return as the benchmark return.

The third benchmark is the reference portfolio return. Before constructing the reference portfolio, we may need matching criterions. As the important pricing factors, size and book-to-market (BM) ratio are the two most important in determining the expected return (Fama and French 1992, 1993, 1996; Lakonishok et al. 1994; Daniel and Titman 1997). In particular, we have to determine the matching pool, which includes stocks that are irrelevant to the sample firm. We form 50 size and book-to-market portfolios (ten size portfolios and five book-to-market portfolios in each size decile) where the size and book-to-market cutoff points are obtained from stock in NYSE exchange. We then are able to compute either equal-weighted or value-weighted portfolio return as the expected return.

The last one is the matching firm method. Because the long-run stock return is positive skewed, we may take the long-run return of the matching firm as the expected return. The skewed return of the sample firm and skewed return of the matching firm offset each other and make the abnormal return symmetric (Barber and Lyon,1997). In addition, matching firm method avoids the new-listing bias and rebalancing problem. In general, the matching criterions of the matching firm are similar to the reference portfolio. Within each reference portfolio, a matching firm could be selected by minimizing the book-to-market difference between the sample firm and the matching firm (Ikenberry et al. 1995). Sometimes, papers use few matching firms but not single matching firm as the benchmark to avoid few outlier impacts (e.g., Lee 1997) or use different matching variable (e.g., Ikenberry and Ramnath 2002; Eberhart et al. 2004). Generally, various matching methods under size and book-to-market effect controls do not largely change the results.

50.2.3 Statistical Inference

The most important statistical problem for the long-run abnormal return is the skewness. The minimum loss of a long-term stock investment is -100 % while the maximum potential gain approaches infinite. Thus, the distribution of the long-run stock return is positive skewed. If we test the long-run stock return by a standard normal distribution, then we tend to reject the null hypothesis (that suggests no abnormal return) for negative returns and accept the null for positive returns. This misspecification leads to a type I error in the distributional left tail but causes a type II error in the distributional right tail.

To solve the skewness problem, Barber and Lyon (1997) suggest the matching firm method because abnormal stock return is the return difference between the sample firm and matching firm, making the skewness from matching firm and sample firm offset by each other. In addition, there are two ways to alleviate the testing problem under the skewness, one is the skewness-adjusted t-statistic and the other is the empirical p-value, suggested by Ikenberry et al. (1995) and Lyon et al. (1999).

Given an abnormal return AR_i, the skewness-adjusted t-statistic illustrated by Lyon et al. (1999) is tested as follows:

$$t_{sa} = \sqrt{N}(S) + \frac{1}{3}\hat{\gamma}S^2 + \frac{1}{6N}\hat{\gamma},$$

where

$$S = \frac{\overline{AR}}{\sigma(AR_i)}, \text{and } \hat{\gamma} = \frac{\sum_{i=1}^{N}\left(AR_i - \overline{AR}\right)^3}{N\sigma(AR_i)^3}. \tag{50.5}$$

Note that the S is the conventional t-statistic and $\hat{\gamma}$ is the coefficient for the skewness adjustment.

The second suggested statistical inference method is the empirical p-value. This approach uses bootstrapping method to construct an empirical distribution with general long-term return features. We use bootstrapping method to select the pseudo-sample firm. Thus, we are able to compare sample firm and pseudo-sample firm as the base of the empirical p-value. In fact, it is possible that we may face moment conditions (higher than third moment condition) in the long-term return estimation, and the skewness-adjusted t-statistic is not enough to capture the return characteristic. Also, the strong cross-sectional correlations among sample observations in BHARs can lead to poorly specified test statistics (Fama 1998; Lyon et al. 1999; Brav 2000). Under the empirical distribution, we can examine the statistical inference without a parametric distribution but are able to capture more unknown statistical features.

As mentioned above, the empirical p-value is generated from the empirical distribution from bootstrapping, and the empirical distribution well controls the skewness and time-dependent properties of the long-run stock returns (Ikenberry, et al. 1995; and Lyon et al. 1999; Chan et al. 2004). In addition, this empirical p-value also solves the statistical inference problem in RBHAR because we may have too few observations in times series for computing standard deviation.[2] To construct the empirical distribution, we need to find pseudo-sample firms that share similar firm characteristics as the sample firm but do not have the interested corporate events. Next, we construct 25 size and book-to-market portfolios (five size portfolios and five book-to-market portfolios in each size quintile) from all

[2]For example, if we compute a 5-year RBHAR, then we have five averages of BHARs in five event years only.

nonevent firms. The nonevent firm selection criterions are similar to the matching firm selection. Then, we randomly sample one pseudo-sample firm out of the corresponding portfolio for each sample firm. For example, if an event firm is with the second size quintile and third book-to-market quintile, then we randomly choose a pseudo firm that is also with the second size quintile and third book-to-market quintile. Hence, we are able to obtain N pseudo firms for N sample firms. Upon these pseudo-sample firms, we calculate the long-run return by CAR, BHAR, or RBHAR method. Finally, we repeat the sampling and long-run return estimation for 1,000 times and obtain 1,000 averages of long-run returns for pseudo-sample firms. The empirical distribution is plotted according to the frequency distribution diagram of those 1,000 average returns of pseudo-sample firms. If the long-run return is larger than the $(1-\alpha)$ percentile of the 1,000 average pseudo-sample firm returns, then we obtain the p-value as α for testing the *positive* average abnormal return. Similarly, if the long-run return is smaller than α percentile of the 1,000 average pseudo-sample firm returns, then we obtain the p-value as α for testing the *negative* average abnormal return.

Figures 50.2a–50.2d are empirical distributions for 1-year to 4-year long-run returns upon size/book-to-market controlled pseudo-sample firms of US repurchase firms during 1980–2008. I collect the repurchase data from SDC database as the example for the empirical distribution construction. For those empirical distributions, it is obvious that 4-year return has more outliers than 1-year return. Moreover, the 4-year return figure has lower kurtosis and is more positive skewed. Obviously, the shape of long-run returns does not obey normal distribution, and the empirical p-value is more relevant to the long-run abnormal return testing.

I also show the specification (statistical size) for different long-run return methods in Table 50.1, which is obtained from Table 5 of Barber and Lyon (1997) and Table 3 of Lyon et al. (1999). They show the percentage of 1,000 random samplings from 200 firms rejecting the null hypothesis that suggests no abnormal return in terms of CAR, BHAR, and RBHAR with different benchmark and testing methods. First, the CAR with matching portfolio as the benchmark has type I error in left tail when measuring the abnormal return in 5 years. Second, the matching firm method yields good specification no matter how we focus on size control, BM control, or a combination control for both size and BM ratio. Third, skewness-adjusted t-statistics has type I error, implying that skewness is not the only statistical feature that we should address. Forth, empirical p-value method performs well even when adopting the reference portfolio as the benchmark, at least for 10 % significance level.

Figure 50.3 shows the testing power of alternative tests by using BHAR as the primary method, and this figure is originally plotted in Fig. 1 of Lyon et al. (1999). The empirical distribution performs better in testing power than classical t-test, no matter what the standard empirical distribution or the bootstrapped skewness-adjusted t-statistic is employed.

In sum, in the event-time approach, the BHAR is suggested. Matching firm is a better matching method than other approaches. In statistical testing, the empirical p-value could be the best way due to its well statistical size control and testing power.

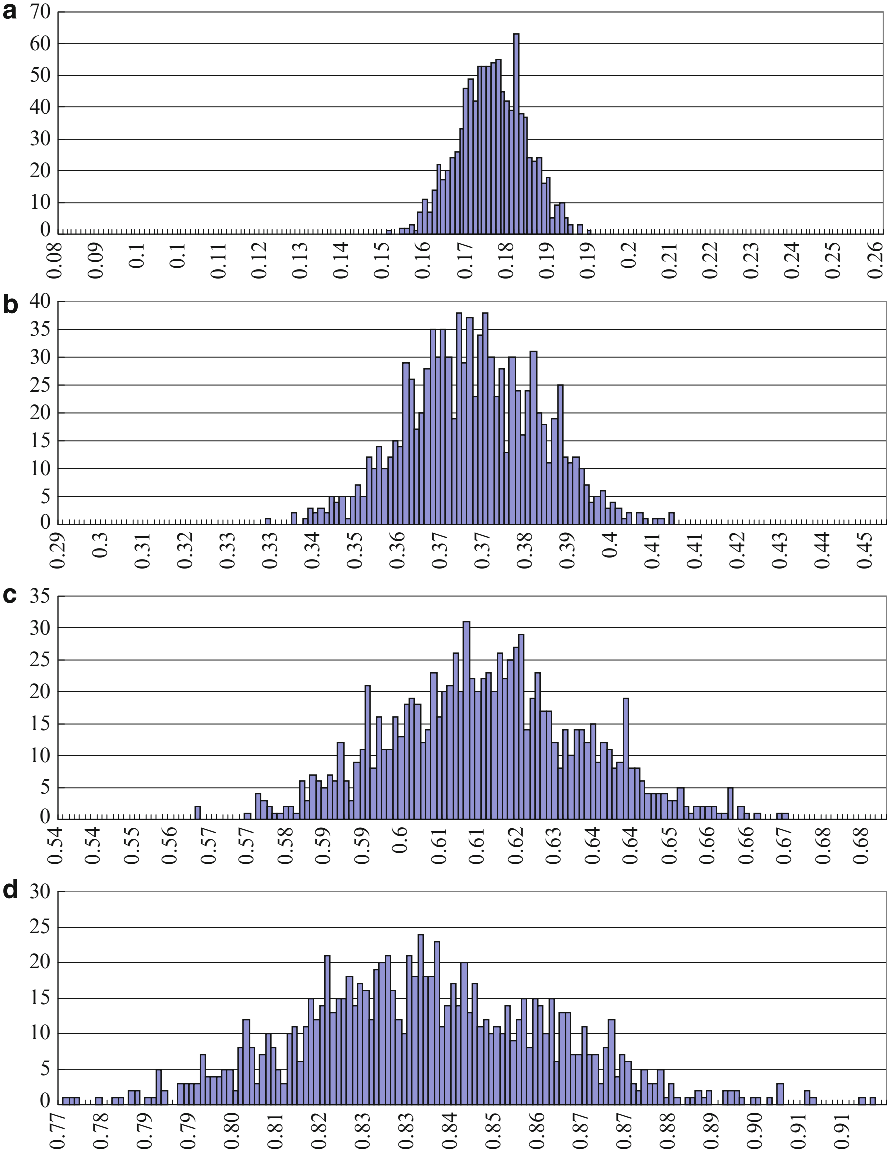


Fig. 50.2 (**a**) One-year return distribution of size/BM controlled pseudo-sample firms. (**b**) Two-year return distribution of size/BM controlled pseudo-sample firms. (**c**) Three-year return distribution of size/BM controlled pseudo-sample firms. (**d**) Four-year return distribution of size/BM controlled pseudo-sample firms

Table 50.1 Specification (size) for alternative statistical tests

A: Specification (size) for CAR with different benchmarks								
Two-tailed theoretical significant level (%)	1		5		10			
Theoretical cumulative density function (%)	0.5	99.5	2.5	97.5	5.0	95.0		
Description of return benchmark							Mean	Skew
Panel C: 60-month CARs								
Size deciles	0	2.4*	0.6	8.0*	1.2	14.7*	3.45	1.11
Book-to-market deciles	0.1	0.7	1.9	4.4*	2.6	7.6*	1.47	1.24
Fifty size/book-to-market portfolio	0.2	1.3*	0.9	5.5*	2.2	10.0*	2.10	1.21
Equally weighted market index	0.0	5.5*	0.2	17.3*	0.5	25.1*	6.27	1.11
Size-matched control firm	0.6	0.3	2.1	2.2	5.2	4.3	−0.59	−0.14
Book-to-market-matched control firm	0.4	0.8	2.9	3.1	5.2	5.4	0.00	−0.01
Size-/book-to-market-matched control firm	0.2	0.4	2.4	2.3	4.3	4.3	−0.63	0.07
Fama-French three-factor model α	0.5	0.3	2.1	2.3	4.9	5.1	−0.94	−1.76

B: Specification (size) for BHAR with different benchmarks							
Statistic	Benchmark	Two-tailed theoretical significance level					
		1 %		5 %		10 %	
		Theoretical cumulative density function (%)					
		0.5	99.5	2.5	97.5	5	95
t-Statistic	Rebalanced size/book-to-market portfolio	11.7*	0.0	23.7*	0.0	33.2*	0.2
t-Statistic	Buy-and-hold size/book-to-market portfolio	2.4*	0.0	6.1*	0.5	10.5*	1.6
Skewness-adjusted t-statistic	Buy-and-hold size/book-to-market portfolio	1.4*	0.4	4.4*	1.7	8.2*	4.8
t-Statistic	Size/book-to-market control firm	0.1	0.1	3.0	1.9	5.4	3.9
Bootstrapped skewness-adjusted t-statistic	Buy-and-hold size/book-to-market portfolio	0.6	1.2*	2.2	3.1	5.0	5.7
Empirical p-value	Buy-and-hold size/book-to-market portfolio	0.2	1.5*	2.7	3.7*	4.9	6.3

This table is from Table 5 of Barber and Lyon (1997, p. 363) and Table 3 of Lyon et al. (1999, p. 179). The numbers presented represent the percentage of 1,000 random samples of 200 firms that reject the null hypothesis of 5-year CAR (Panel A) and buy-and-hold abnormal return (Panel B) at the theoretical significance levels of 1 %, 5 %, or 10 % in favor of the alternative hypothesis of a significantly negative abnormal return (i.e., calculated p-value is less than 0.5 % at the 1 % significance level) or a significantly positive abnormal return (calculated p-value is greater than 99.5 % at the 1 % significance level). The alternative statistics and benchmarks are described in detail in the main text. * indicates the percentage is significantly different from the theoretical significance level at the 5 % (Panel A) and 1 % level (Panel B), one-sided binomial test statistic

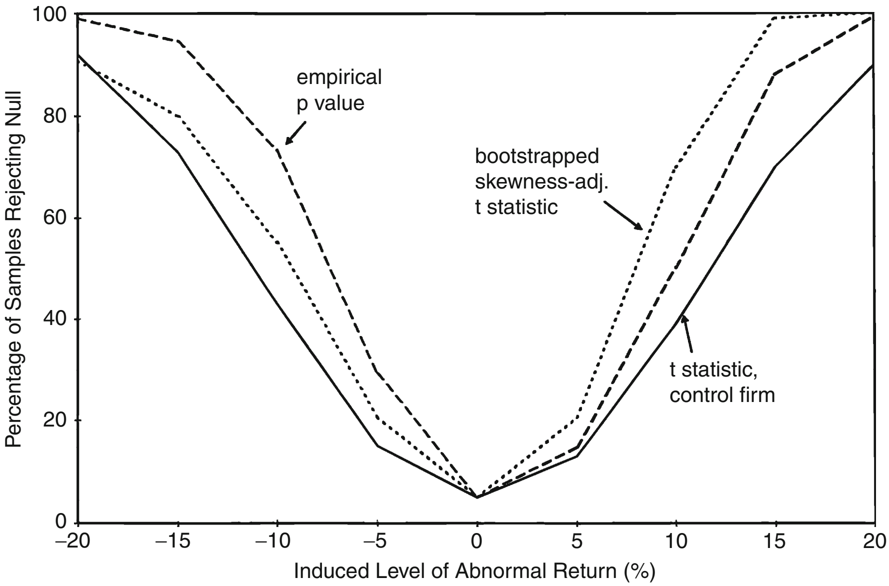


Fig. 50.3 Power of alternative tests in random ample. This figure is from Fig. 1 of Lyon et al. (1999, p. 180). The percentage of 1,000 random samples of 200 firms rejecting the null hypothesis of no annual buy-and-hold abnormal return at various induced levels of abnormal return (horizontal axis) based on control firm method, bootstrapped skewness-adjusted *t*-statistic, and empirical *p*-values

50.3 Long-Run Return Estimation in Calendar-Time Approach

A potential problem in the event-time return estimations is the cross-sectional dependence. For example, buy-and-hold return is computed over a long horizon; it is possible that many sample firms' returns overlap with each other, making strong cross-sectional correlations among long-horizon returns. This cross-sectional dependence is even more profound when we have repeating events by the same firm, such as repurchase, SEO, merger, and stock splits. In addition, buy-and-hold returns also enlarge the long-run abnormal return because of the inflated returns stemming from the compounding effect. Therefore, the long-run return results might disappear if we apply other methodologies to compute the long-run abnormal return (Mitchell and Stafford 2000). Fama (1998) also documents that the long-run stock return should be examined by the value-weighted factor model since the buy-and-hold return usually uses an equal-weighted scheme that is related to the ignored transaction costs. Although Loughran and Ritter (2000) suggest that the value-weighted factor model is the least powerful test for long-run returns, the calendar-time method could be always a robust check for our long-run return estimation.

To estimate the abnormal return in calendar time, we need to form a monthly portfolio for each calendar month. The portfolio return could be generated by an equal-weighted, value-weighted, or a log-value-weighted method (Ikenberry et al. 2000). Upon these monthly returns in calendar time, we are able to carry out the mean monthly abnormal return or factor model analysis.

50.3.1 Return Estimations: Mean Monthly Return and Factor Models

The first method in the calendar-time approach is the mean monthly abnormal return (MMAR). We can start from a 5-year holding strategy on a corporate event. For any given calendar month (e.g., June 2002), we need to include a stock if it had the corporate event in the past 60 months (e.g., looking backward at a period of May 2002–June 1997). We then need to form a monthly portfolio for this specific calendar month, and the portfolio return could be computed upon equal-weighted, value-weighted or log-value-weighted scheme. Next, we repeat the abovementioned step for all calendar months throughout the sample period, and then the mean monthly return is the time-series average of monthly portfolio returns.

Similar to the benchmark problem in event-time approach, we need to select the CRSP market index return, reference portfolio, or the matching firm as the benchmark for the MMAR. For any benchmark $E(R_i)$, the MMAR is

$$MMAR = \frac{\sum_{t=1}^{T} \left(R_t^P - R_t^M\right)}{T},$$

where

$$R_t^P = \frac{\sum_{i=1}^{N_t} w_{i,t} R_{i,t}}{N_t}, \text{ and } R_t^M = \frac{\sum_{i=1}^{N_t} w_{i,t} E\left(R_{i,t}\right)}{N_t}. \tag{50.6}$$

T is for calendar month in this setting; R_t^P is the monthly portfolio return of sample firms; R_t^M is the monthly portfolio return of benchmarks; and N_t is the number of observations in each calendar month. The statistical inference can be either classical t-statistic or the Newey and West (1987) estimation.

As suggested by Fama (1998), Mitchell and Stafford (2000), and Schultz (2003), factor model is a robust method in estimating the long-run stock abnormal return. The standard Fama and French three-factor and Carhart four-factor models could be described as

$$R_t^P - r_{f,t} = \alpha + \beta\left(r_{m,t} - r_{f,t}\right) + sSMB_t + hHML_t + e_t, \tag{50.7}$$

$$R_t^P - r_{f,t} = \alpha + \beta\left(r_{m,t} - r_{f,t}\right) + sSMB_t + hHML_t + mMOMENTUM_t + e_t, \tag{50.8}$$

where R_t^P is the sample firm portfolio return for each calendar month, and could be obtained from the equal-weighted, value-weighted, or log-value-weighted average. This average return in calendar is similar to what we compute in the MMAR. r_f is the risk-free rate, usually the short-term Treasury bill rate; r_m is usually computed as CRSP value-weighted index return; *SMB* is small firm portfolio return minus big-firm portfolio return; *HML* is the high book-to-market portfolio return minus low book-to-market portfolio return; *MOMENTUM* is the winner portfolio return minus loser portfolio return where winner and loser portfolios are identified by past 1-year return. *SMB*, *HML*, and *MOMENTUM* are applied to control size and book-to-market and momentum effects, respectively (Fama and French 1992, 1993; Jegadeesh and Titman 1993; Lakonishok et al. 1994; Carhart 1997). The abnormal return is the regression intercept and can be tested based on the *t*-values or Newey and West (1987) estimator.

Next, I introduce a modification of the factor model analysis: the zero-investment portfolio method. Daniel and Titman (1997) and Eberhart et al. (2004) study the long-run return by using the zero-investment portfolio approach to control for both risk and firm characteristic effects. To form the factor model under a zero-investment portfolio strategy, we have to buy sample stocks and short-sell matching stocks. Taking Carhart (1997) four-factor model as the example, we have

$$\begin{aligned} R_t^P - R_t^M = {} & \left(\alpha^P - \alpha^M\right) + (\beta_P - \beta_M)\left(r_{m,t} - r_{f,t}\right) + (s_P - s_M)SMB_t \\ & + (h_P - h_M)HML_t + (m_P - m_M)MOMENTUM_t + e_t, \end{aligned} \tag{50.9}$$

and we use $(\alpha^P – \alpha^M)$ as the abnormal return controlling for both risks and firm characteristics. It is also the hedging portfolio return controlled for the common risk factors. The matching firm selection criterions can apply the steps in benchmark problem section. For other modifications of the factor model analysis, Eberhart et al. (2004) provide more examples in their Table 3.

50.3.2 Conditional Market Model and Ibbotson's RATS

One major challenge to the standard market model is that the risk loadings are assumed to be unchanged. To estimate the factor loadings, we usually need long time series to obtain the estimated risk loadings, and the fixed risk loading over time is naturally assumed in the OLS analysis. Yet, the magnitude of the risk of a firm could be changed; in particular many corporate events change the risk of the firm (e.g., R&D increases could be followed by risk increases, and share repurchase could be followed by risk decreases). Accordingly it is needed to introduce the conditional market model to address the time-varying market model.

There are at least two ways to address the time-varying risks in the regression model as the conditional market model. To simplify the problem, I use the CAPM as the first example. First, the systematic risk could change along with some firm characteristics and macroeconomic variables. Petkova and Zhang (2005) use the following regression analysis to estimate abnormal return:

$$R_t^P - r_{f,t} = \alpha + (b_0 + b_1 DIV_t + b_2 DEF_t + b_3 TERM_t + b_4 TB_t)\left(r_{m,t} - r_{f,t}\right) + e_t. \tag{50.10}$$

$(b_0 + b_1 DIV_t + b_2 DEF_t + b_3 TERM_t + b_4 TB_t)$ is the β_t that accommodates to the time-varying risk loading. They assume that the risk changes with the dividend yield (DIV), the default spread (DEF), the term spread (TERM), and the short-term Treasury bill rate (TB). Again, the abnormal return is the intercept α.

The second method is the rolling regression, as suggested by Petkova and Zhang (2005) and Eberhart et al. (2004). If we have a sample period from January 1990, then we can use the portfolio returns in the first 60 months (i.e., January 1990–December 1994) to carry out the Carhart (1997) four-factor regression. We substitute the estimated factor loadings from these 60 monthly returns into the equity premiums in 61th month (i.e., January 1995) and then obtain the expected portfolio return for 61th month. Thus, the abnormal return for 61th month is from the portfolio return of the sample firm minus the expected portfolio return. Next, we need to repeat the abovementioned steps for every month by rolling return windows. Finally, we estimate the abnormal return as the average of abnormal returns across time and use the time-series volatility to test the statistical significance.

The final method relating to the time-varying risk is the Ibbotson (1975) RATS though it is not under the family of the factor model analysis. The original setting of Ibbotson RATS is designed for the long-run return estimation, yet recent papers use this method combining the factor model analysis to measure the long-run abnormal return (e.g., Peyer and Vermaelen 2009). Based on Carhart (1997) four-factor model, we regress the security excess return on the Carhart (1997) four factors for each month in the event time. Given a 60-month holding strategy, we have to carry out this regression for 1st month to 60th month following the corporate event date. Then, the abnormal return for month τ is the intercept of this four-factor regression:

$$R_{i,t} - r_{f,t} = \alpha_\tau + \beta_\tau\left(r_{m,t} - r_{f,t}\right) + s_\tau SMB_t + h_\tau HML_t + m_\tau MOMENTUM_t + e_t. \tag{50.11}$$

The regression analysis is similar to what I introduce in Eq. 50.8; however, the regression is examined every event month τ. For every event month or event year, we can obtain the average abnormal return as the average of the intercepts (α_τ), which is obtained from a model allowing time-varying risks.

50.4 Conclusion

The long-run return studies have been investigated for many corporate events and asset pricing studies in the past two decades. I introduce the long-run stock return methodologies and their statistical inference adopted in recent papers. Two categories of long-run return methods are illustrated: the event-time approach and calendar-time approach. Under the event-time category, we have methods including cumulative abnormal return, buy-and-hold abnormal return, and abnormal returns around earnings announcements. Although the event-time approach is able to be implemented as an investment strategy in real world, it also raises more benchmark and statistical inference problems. Under the calendar-time category, we have mean monthly abnormal return, factor models, and Ibbotson's RATS. Generally, calendar-time approach is more popular due to its robustness and variety. For any long-run return study, I may suggest that combining works on those methodologies could be necessary.

References

Agrawal, A., Jaffe, J., & Mandelker, G. (1992). The post-merge performance of acquiring firms: A re-examination of an anomaly. *Journal of Finance, 47*, 1605–1621.

Barber, B., & Lyon, J. (1997). Detecting long-run abnormal stock returns: The empirical power and specification of test statistics. *Journal of Financial Economics, 43*, 341–372.

Bernard, V., & Thomas, J. (1989). Post-earnings-announcement drift: Delayed price response or risk premium. *Journal of Accounting Research, 27*, 1–36.

Brav, A. (2000). Inference in long-horizon event studies: A Bayesian approach with application to initial public offerings. *Journal of Finance, 55*, 1979–2016.

Carhart, M. (1997). On persistence in mutual fund performance. *Journal of Finance, 52*, 57–82.

Chan, K., Ikenberry, D., & Lee, I. (2004). Economic sources of gain in share repurchases. *Journal of Financial and Quantitative Analysis, 39*, 461–479.

Chan, K., Ikenberry, D., Lee, I., & Wang, Y. (2010). Share repurchases as a potential tool to mislead investors. *Journal of Corporate Finance, 16*, 137–158.

Daniel, K., & Titman, S. (1997). Evidence on the characteristics of cross sectional variation in stock returns. *Journal of Finance, 52*, 1–33.

Denis, D., & Sarin, A. (2001). Is the market surprised by poor earnings realizations following seasoned equity offerings? *Journal of Financial and Quantitative Analysis, 36*(2), 169–193.

Eberhart, A., Maxwell, W., & Siddique, A. (2004). An examination of long-term abnormal stock returns and operating performance following R&D increases. *Journal of Finance, 59*(2), 623–650.

Fama, E., & French, K. (1992). The cross-section of expected returns. *Journal of Finance, 47*, 427–466.

Fama, E., & French, K. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*, 3–56.

Fama, E., & French, K. (1996). Multifactor explanations of asset pricing anomalies. *Journal of Finance, 51*, 55–84.

Fama, E. F., (1998). Market efficient, long-term returns, and behavioral finance. *Journal of Financial Economics, 49*, 283–306.

Ibbotson, R. (1975). Price performance of common stock new issues. *Journal of Financial Economics, 2*, 235–272.

Ikenberry, D., & Ramnath, S. (2002). Underreaction to self-selected news events: The case of stock splits. *Review of Financial Studies, 15*, 489–526.

Ikenberry, D., Lakonishok, J., & Vermaelen, T. (1995). Market underreaction to open market share repurchases. *Journal of Financial Economics, 39*, 181–208.

Ikenberry, D., Lakonishok, J., & Vermaelen, T. (2000). Stock repurchases in Canada: Performance and strategic trading. *Journal of Finance, 55*, 2373–2397.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance, 48*, 65–91.

Kothari, S. P., & Warner, J. (1997). Measuring long-horizon security price performance. *Journal of Financial Economics, 43*, 301–339.

La Porta, R., Lakonishok, J., Shleifer, A., & Vishny, R. (1997). Good news for value stocks: Further evidence on market efficiency. *Journal of Finance, 52*(2), 859–874.

Lakonishok, J., Shleifer, A., & Vishny, R. (1994). Contrarian investment, extrapolation, and risk. *Journal of Finance, 49*, 1541–1578.

Lee, I. (1997). Do firms knowingly sell overvalued equity? *Journal of Finance, 52*, 1439–1466.

Loughran, T., & Ritter, J. (2000). Uniformly least powerful tests of market efficiency. *Journal of Financial Economics, 55*, 361–389.

Lyon, J., Barber, B., & Tsai, C. L. (1999). Improved methods for tests of long-run abnormal stock returns. *Journal of Finance, 54*, 165–201.

Mitchell, M., & Stafford, E. (2000). Managerial decisions and long-term stock price performance. *Journal of Business, 73*, 287–329.

Newey, W., & West, K. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica, 55*, 703–708.

Petkova, R., & Zhang, L. (2005). Is value riskier than growth. *Journal of Financial Economics, 78*, 187–202.

Peyer, U., & Vermaelen, T. (2009). The nature and persistence of buyback anomalies. *Review of Financial Studies, 22*, 1693–1745.

Ritter, J. (1991). The long-run performance of initial public offerings. *Journal of Finance, 46*, 1–27.

Schultz, P. (2003). Pseudo market timing and the long-run underperformance of IPOs. *Journal of Finance, 58*, 483–517.

Sloan, R. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? *Accounting Review, 71*, 289–315.

Value-at-Risk Estimation via a Semi-parametric Approach: Evidence from the Stock Markets

51

Cheng-Few Lee and Jung-Bin Su

Contents

C.-F. Lee
Department of Finance and Economics, Rutgers Business School, Rutgers, The State University of New Jersey, Piscataway, NJ, USA

Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
e-mail: lee@business.rutgers.edu; cflee@business.rutgers.edu

J.-B. Su (✉)
Department of Finance, China University of Science and Technology, Nankang, Taipei, Taiwan
e-mail: jungbinsu@cc.cust.edu.tw; jungbinsu@gmail.com

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_51,

Abstract

This study utilizes the parametric approach (GARCH-based models) and the semi-parametric approach of Hull and White (Journal of Risk 1: 5–19, 1998) (HW-based models) to estimate the Value-at-Risk (VaR) through the accuracy evaluation of accuracy for the eight stock indices in Europe and Asia stock markets. The measure of accuracy includes the unconditional coverage test by Kupiec (Journal of Derivatives 3: 73–84, 1995) as well as two loss functions, quadratic loss function, and unexpected loss. As to the parametric approach, the parameters of generalized autoregressive conditional heteroskedasticity (GARCH) model are estimated by the method of maximum likelihood and the quantiles of asymmetric distribution like skewed generalized student's t (SGT) can be solved by composite trapezoid rule. Sequentially, the VaR is evaluated by the framework proposed by Jorion (*Value at Risk: the new benchmark for managing financial risk*. New York: McGraw-Hill, 2000). Turning to the semi-parametric approach of Hull and White (Journal of Risk 1: 5–19, 1998), before performing the traditional historical simulation, the raw return series is scaled by a volatility ratio where the volatility is estimated by the same procedure of parametric approach. Empirical results show that the kind of VaR approaches is more influential than that of return distribution settings on VaR estimate. Moreover, under the same return distributional setting, the HW-based models have the better VaR forecasting performance as compared with the GARCH-based models. Furthermore, irrespective of whether the GARCH-based model or HW-based model is employed, the SGT has the best VaR forecasting performance followed by student's t, while the normal owns the worst VaR forecasting performance. In addition, all models tend to underestimate the real market risk in most cases, but the non-normal distributions (student's t and SGT) and the semi-parametric approach try to reverse the trend of underestimating.

Keywords

Value-at-Risk • Semi-parametric approach • Parametric approach • Generalized autoregressive conditional heteroskedasticity • Skewed generalized student's t • Composite trapezoid rule • Method of maximum likelihood • Unconditional coverage test • Loss function

51.1 Introduction

Over the last two decades, a number of global and national financial disasters have occurred due to failures in risk management procedures. For instance, US Savings and Loan crisis of 1989–1991, Japanese asset price bubble collapse of 1990, Black Wednesday of 1992–1993, 1994 economic crisis in Mexico, 1997 Asian Financial Crisis, 1998 Russian financial crisis, financial crisis of 2007–2010, followed by the late 2000s recession, and the 2010 European sovereign debt crisis. The crises caused many enterprises to be liquidated and many countries to face near depressions in their economies. These painful experiences once again underline the importance of accurately measuring financial risks and implementing sound risk management policies. Hence, Value-at-Risk (VaR) is a widely used risk measure of the risk of loss on a specific portfolio of financial assets because it is an attempt to summarize the total risk with a single number. For example, if a portfolio of stocks has a 1-day 99 % VaR of US$1,000, there is a 1 % probability that the portfolio will fall in value by more than US$1,000 over a 1-day period. In other words, we are 99 % certain that we will not lose more than US$1,000 in the next 1 day, where 1 day is the time horizon, 99 % is the confidence level, and the US$1,000 is the VaR of the portfolio.

VaR estimates are currently based on either of three main approaches: the historical simulation, the parametric method, and the Monte Carlo simulation. The Monte Carlo simulation is a class of computational algorithms that rely on repeated random sampling to compute their results. That is, this approach allows for an infinite number of possible scenarios you are exposing yourself to huge model risks in determining the likelihood of any given path. In addition, as you had more and more variables that could possibly alter your return paths, model complexity and model risks also increase in scale. Like historical simulation, however, this methodology removes any assumption of normality and thus, if modeled accurately, probably would give the most accurate measure of the portfolio's true VaR. Besides, little research such as Vlaar (2000) had applied this approach to estimate the VaR. The parametric method is also known as variance/covariance approach. This method is popular because the only variables you need to do the calculation are the mean and standard deviation of the portfolio, indicating the simplicity of the calculations. The parametric method assumes that the returns of the portfolios are normally distributed and serially independent. In practice, this assumption of return normality has proven to be extremely risky. Indeed, this was the biggest mistake that LTCM made gravely underestimating their portfolio risks. Another weakness with this method is the stability of the standard deviation through time as well as the stability of the variance/covariance matrix in your portfolio. However, it is easy to depict how correlations have changed over time particularly in emerging markets and through contagion in times of financial crisis. Additionally, numerous studies focused on the parametric approach of generalized autoregressive conditional heteroskedasticity (GARCH) family variance specifications (i.e., risk metrics, asymmetric power ARCH (APARCH), exponential GARCH (EGARCH), threshold GARCH (TGARCH), integrated GARCH (IGARCH), and fractional IGARCH (FIGARCH)) to estimate the VaR (see Vlaar (2000), Giot and Laurent (2003a, b), Gencay et al. (2003), Cabedo and Moya (2003), Angelidis et al. (2004), Huang and Lin (2004), Hartz et al. (2006), So and Yu (2006), Sadeghi and Shavvalpour (2006), Bali and Theodossiou (2007), Bhattacharyya et al. (2008), Lee et al. (2008), Lu et al. (2009), Lee and Su (2011), and so on). Lately, in the empirical study of parametric VaR approach, several researches have utilized the other type of volatility specifications besides GARCH family such as the ARJI-GARCH-based model (hereafter ARJI) of Chan and Maheu (2002) which combines the GARCH specification of volatility and autoregressive jump intensity (ARJI) in jump intensity (see Su and Hung (2011) Chang et al. (2011), and so on). Moreover, the other types of long memory volatility specifications such as fractional integrated APARCH (FIAPARCH) and hyperbolic

GARCH (HYGARCH) besides FIGARCH mentioned above are also used to estimate the VaR (see Aloui and Mabrouk (2010), Degiannakis et al. (2012)). As to distribution setting, some special distributions like the Weibull distribution (see Gebizlioglu et al. (2011)), the asymmetric Laplace distribution (see Chen et al. (2012)), and the Pearson type-IV distribution (see Stavroyiannis et al. (2012)) are also employed to estimate VaR.

The historical simulation assumes that the past will exactly replicate the future. The VaR calculation of this approach is literally ranking all of your past historical returns in terms of lowest to highest and computing with a predetermined confidence rate what your lowest return historically has been.[1] In addition, several studies such as Vlaar (2000), Gencay et al. (2003), Cabedo and Moya (2003), and Lu et al. (2009) had applied this approach to estimate the VaR. Even though it is relatively easy to implement, there is a couple of shortcomings of this approach, and first of all is that it imposes a restriction on the estimation assuming asset returns are independent and identically distributed (iid) which is not the case. From empirical evidence, it is known that asset returns are clearly not independent as they exhibit volatility clustering.[2] Therefore, it can be unrealistic to assume iid asset returns. Second restriction relates to time. Historical simulation applies equal weight to all returns of the whole period, and this is inconsistent with the nature where there is diminishing predictability of data that are further away from the present.

These two shortcomings of historical simulation lead this paper to use the approach proposed by Hull and White (1998) (hereafter, HW method) as a representative of the semi-parametric approach. This semi-parametric approach combines the abovementioned parametric approach of GARCH-based variance specification with the weighted historical simulation. The weighted historical simulation applies decreasing weights to returns that are further away from the present, which overcomes the inconsistency of historical simulation with diminishing predictability of data that are further away from the present. Hence, this study utilizes the parametric approach (GARCH-N, GARCH-T, and GARCH-SGT models) and the semi-parametric approach of Hull and White (1998) (HW-N, HW-T, and HW-SGT models), totaling six models, to estimate the VaR for the eight stock indices in Europe and Asia stock markets, then uses three accuracy measures: one likelihood ratio test (the unconditional coverage test (LR_{uc}) of Kupiec (1995)) and two loss functions (the average quadratic loss function (AQLF) of Lopez (1999) and the unexpected loss (UL)) to compare the forecasting ability of the aforementioned models in terms of VaR.

Our results show that the kind of VaR approaches is more influential than that of return distribution settings on VaR estimate. Moreover, under the same return distributional setting, the HW-based models have the better VaR forecasting

[1]This means if you had 200 past returns and you wanted to know with 99 % confidence what's the worst you can do, you would go to the 2nd data point on your ranked series and know that 99 % of the time you will do no worse than this amount.

[2]Large changes tend to be followed by large changes, of either sign, and small changes tend to be followed by small changes.

performance as compared with the GARCH-based models. Furthermore, irrespective of whether the GARCH-based model or HW-based model is employed, the skewed generalized student's t (SGT) has the best VaR forecasting performance followed by student's t, while the normal owns the worst VaR forecasting performance. In addition, all models tend to underestimate the real market risk in most cases, but the non-normal distributions (student's t and SGT) and the semi-parametric approach try to reverse the trend of underestimating.

The remainder of this paper is organized as follows. Section 51.2 describes the methodology of two dissimilar VaR approaches (the parametric and semi-parametric approaches) and the VaR calculations using these approaches. Section 51.3 provides criteria for evaluating risk management, and Sect. 51.4 reports on and analyzes the empirical results of the out-of-sample VaR forecasting performance. The final section makes some concluding remarks.

51.2 Empirical Methodology

In this paper, there are two approaches of calculating VaR to be introduced, that is, the parametric method and the semi-parametric approach. Here, we use the GARCH(1,1) model with three conditional distributions, namely, the normal, student's t, and SGT distributions, to estimate the corresponding volatility in terms of different stock indices then employ the framework of Jorion (2000) to evaluate the VaR of parametric approach whereas utilizing the weighting scheme of volatility proposed by Hull and White (1998) (hereafter, HW method) which is a straightforward extension of traditional historical simulation to calculate the VaR of semi-parametric VaR.

51.2.1 Parametric Method

Many time series data of financial assets appear to exhibit autocorrelated and volatility clustering. Bollerslev et al. (1992) showed that the GARCH(1,1) specification works well in most applied situations. Furthermore, the unconditional distribution of those returns displays leptokurtosis and a moderate amount of skewness. Hence, this study thus considers the applicability of the GARCH(1,1) model with three conditional distributions, namely, the normal, student's t, and SGT distributions, to estimate the corresponding volatility in terms of different stock indices and use the GARCH model as an official delegate of the VaR model.

51.2.1.1 GARCH Model with Normal Distribution

Let $r_t = (\ln P_t - \ln P_{t-1}) \times 100$, where P_t denotes the stock price and r_t denotes the continuously compounded daily returns of the underlying assets on time t. The GARCH (1,1) model with SGT distribution (GARCH-SGT) can be expressed as follows:

$$r_t = \mu + e_t, \quad e_t = \varepsilon_t \sigma_t, \quad \cdot\, \varepsilon_t \sim \text{IID} \;\; \text{SGT}(0, 1; \kappa, \lambda, n) \qquad (51.1)$$

$$\sigma_t^2 = \omega + \alpha e_{t-1}^2 + \beta\sigma_{t-1}^2 \tag{51.2}$$

where e_t is the current error and μ and $\sigma_t{}^2$ are the conditional mean and variance of return, respectively. Moreover, the variance parameters ω, α, and β are the parameters to be estimated and obey the constraints ω, α, $\beta > 0$ and $\alpha + \beta < 1$. IID denotes that the standardized errors ε_t are independent and identically distributed. Since ε_t is drawn from the standard normal distribution, the probability density function for ε_t is

$$f(\varepsilon_t) = \frac{1}{\sqrt{2\pi}}\exp\left(-\frac{\varepsilon_t^2}{2}\right), \tag{51.3}$$

and the log-likelihood function of GARCH-N model thus can be written as

$$L(\psi) = \ln f(r_t|\Omega_{t-1};\psi) = -0.5\left(\ln\ 2\pi + \ln\ \sigma_t^2 + (r_t - \mu)^2/\sigma_t^2\right) \tag{51.4}$$

where $\psi = [\mu, \omega, \alpha, \beta]$ is the vector of parameters to be estimated and Ω_{t-1} denotes the information set of all observed returns up to time $t - 1$. Under the framework of the parametric techniques (Jorion 2000), the 1-day-ahead VaR based on GARCH-N model can be calculated as

$$VaR_{t+1|t}^N = \mu + F_c(\varepsilon_t)\cdot\hat{\sigma}_{t+1|t} \tag{51.5}$$

where $F_c(\varepsilon_t)$ is the left-tailed quantile at c% for the standardized normal distribution. $\hat{\sigma}_{t+1|t}$ is the one-step-ahead forecasts of the standard deviation of the returns conditional on all information upon the time t.

51.2.1.2 GARCH Model with Student's t Distribution

Since the characteristics of many financial data are non-normal, the student's t distribution is most commonly employed to capture the fat-tailed properties of their empirical distributions. Moreover, Bollerslev (1986) argued that using the student's t distribution as the conditional distribution for GARCH model is more satisfactory since it exhibits thicker tail and larger kurtosis than normal distribution. Under the same specifications of mean and variance equation as the GARCH-N model, the probability density function for the standardized student's t distribution can be represented as follows:

$$f(\varepsilon_t) = \frac{\Gamma(0.5(n+1))}{\Gamma(0.5\ n)\sqrt{\pi(n-2)}}\left\{1 + \frac{\varepsilon_t^2}{n-2}\right\}^{-\frac{n+1}{2}} \tag{51.6}$$

where $\Gamma(\bullet)$ is the gamma function and n is the shape parameter. Hence, the log-likelihood function of the GARCH-T model can be expressed as

$$
\begin{aligned}
L(\psi) &= \ln f(r_t|\Omega_{t-1};\psi) = \ln\left[\frac{\Gamma(0.5(n+1))}{\Gamma(0.5n)\sqrt{\pi(n-2)}}\right] \\
&\quad -\ln\sigma_t - \frac{n+1}{2}\ln\left[1+\left(\frac{r_t-\mu}{\sigma_t}\right)^2(n-2)^{-1}\right]
\end{aligned}
\tag{51.7}
$$

where $\psi = [\mu, \omega, \alpha, \beta, n]$ is the vector of parameters to be estimated. The 1-day-ahead VaR based on GARCH-T model can be obtained as

$$
\mathrm{VaR}^{T}_{t+1|t} = \mu + F_c(\varepsilon_t; n)\cdot\hat{\sigma}_{t+1|t} \tag{51.8}
$$

where $F_c(\varepsilon_t;\ n)$ denotes the left-tailed quantile at c% for standardized student's t distribution with shape parameter n.

51.2.1.3 GARCH Model with Skewed Generalized Student's t Distribution

This study also employs the SGT distribution of Theodossiou (1998) which allows return innovation to follow a flexible treatment of both skewness and excess kurtosis in the conditional distribution of returns. Under the same specifications of mean and variance as the GARCH-N model, the probability density function for the standardized SGT distribution is derived by Lee and Su (2011) and can be represented as follows:

$$
f(\varepsilon_t) = C\left\{1+\frac{|\varepsilon_t+\delta|^{\kappa}}{[1+\mathrm{sign}\,(\varepsilon_t+\delta)\lambda]^{\kappa}\theta^{\kappa}}\right\}^{-\frac{n+1}{\kappa}} \tag{51.9}
$$

where $\theta = \frac{1}{S(\lambda)}B\left(\frac{1}{\kappa},\frac{n}{\kappa}\right)^{\frac{1}{2}}B\left(\frac{3}{\kappa},\frac{n-2}{\kappa}\right)^{-\frac{1}{2}}, S(\lambda) = \sqrt{1+3\lambda^2-4A^2\lambda^2},$

$$
A = B\left(\tfrac{2}{\kappa},\tfrac{n-1}{\kappa}\right)B\left(\tfrac{1}{\kappa},\tfrac{n}{\kappa}\right)^{-0.5}B\left(\tfrac{3}{\kappa},\tfrac{n-2}{\kappa}\right)^{-0.5},\ \delta = \tfrac{2\lambda A}{S(\lambda)},\ C = \tfrac{\kappa}{2\theta}B\left(\tfrac{1}{\kappa},\tfrac{n}{\kappa}\right)^{-1}
$$

where κ, n, and λ are scaling parameters and C and θ are normalizing constants ensuring that $f(\bullet)$ is a proper p.d.f. The parameters κ and n control the height and tails of density with constraints $\kappa > 0$ and $n > 2$, respectively. The skewness parameter λ controls the rate of descent of the density around the mode of ε_t with $-1 < \lambda < 1$. In the case of positive (resp. negative) skewness, the density function skews toward the right (resp. left). Sign is the sign function, and $B(\bullet)$ is the beta function. The parameter n has the degrees of freedom interpretation in case $\lambda = 0$ and $\kappa = 2$. The log-likelihood function of the GARCH-SGT model thus can be written as

$$
\begin{aligned}
L(\psi) = \ln f(r_t|\Omega_{t-1};\psi) &= \ln C - \ln\sigma_t - \frac{n+1}{\kappa}\ln\left\{1+\left|\frac{r_t-\mu}{\sigma_t}+\delta\right|^{\kappa}\right. \\
&\qquad \left.\left[1+\mathrm{sign}\left(\frac{r_t-\mu}{\sigma_t}+\delta\right)\lambda\right]^{-\kappa}\theta^{-\kappa}\right\}
\end{aligned}
\tag{51.10}
$$

where $\psi = [\mu, \omega, \alpha, \beta, \kappa, \lambda, n]$ is the vector of parameters to be estimated, and Ω_{t-1} denotes the information set of all observed returns up to time $t - 1$.

Under the framework of the parametric techniques (Jorion 2000), the 1-day-ahead VaR based on GARCH-SGT model can be calculated as

$$\mathrm{VaR}^{\mathrm{SGT}}_{t+1|t} = \mu + F_c(\varepsilon_t; \kappa, \lambda, n) \cdot \hat{\sigma}_{t+1|t} \tag{51.11}$$

where $F_c(\varepsilon_t; \kappa, \lambda, n)$ denotes the left-tailed quantile at c% for standardized SGT distribution with shape parameters κ, λ, and n and can be evaluated by a numerical integral method (composite trapezoid rule).[3] Particularly, the SGT distribution generates the student's t distribution for $\lambda = 0$ and $\kappa = 2$. Moreover, the SGT distribution generates the normal distribution for $\lambda = 0$, $\kappa = 2$, and $n = \infty$.

51.2.2 Semi-parametric Method

In this paper, we use the approach proposed by Hull and White (1998) (hereafter, HW method) as a representative of the semi-parametric approach. The HW method is a straightforward extension of traditional historical simulation. Instead of using the actual historical percentage changes in market variables for the purposes of calculating VaR, we use historical changes that have been adjusted to reflect the ratio of the current daily volatility to the daily volatility at the time of the observation and assume that the variance of each market variable during the period covered by the historical data is monitored using a GARCH model. The methodology is explained in the following three steps: ***First***, use a raw return series, $\{r_1, r_2, r_3, \ldots\ldots, r_{t=T}\}$, to fit the GARCH(1,1) models with alternative distributions expressed as in Sect. 51.2.1. Thus, a series of daily volatility estimates, $\{\sigma_1, \sigma_2, \sigma_3, \ldots\ldots, \sigma_{t=T}\}$, are obtained where T is the number of estimated samples. ***Second***, the modified return series are obtained by the raw return series multiplied by the ratio of the current daily volatility to the daily volatility at the time of the observation, σ_T/σ_i. That is, the modified return series are expressed as $\{r_1^*, r_2^*, r_3^*, \ldots\ldots, r_{t=T}^*\}$, where $r_i^* = r_i(\sigma_T/\sigma_i)$. ***Finally***, sort the returns ascendingly to achieve the empirical distribution. Thus, VaR is the percentile that corresponds to the specified confidence level.

The HW-GARCH-SGT model (simply called HW-SGT) implies that the standardized residual return of the GARCH-SGT model is applied by the HW approach to estimate the VaR so are HW-N and HW-T models.

51.3 Evaluation Methods of Model-Based VaR

Many financial institutions have been required to hold capital against their market risk exposure, while the market risk capital requirements are based on the VaR estimates

[3] See Faires and Burden (2003) for more details.

generated by the financial institutions' own risk management models. Explicitly, the accuracy of these VaR estimates is of concern to both financial institutions and their regulators. Hence, model accuracy is important to all VaR model users. To compare the forecasting ability of the aforementioned models in terms of VaR, this study considers three accuracy measures: the unconditional coverage test of Kupiec (1995) which are quite standard in the literatures. Moreover, the quadratic loss function and the unexpected loss are introduced and used for determining the accuracy of model-based VaR measurements.

51.3.1 Log-Likelihood Ratio Test

Before performance competition of alternative VaR models, with regard to two models, we can use the log-likelihood ratio test to compare which one model has the better matching ability of between actual data and empirical model. This can be regarded as the preliminary analysis. The log-likelihood ratio test is a statistical test used to compare the fit of two models, one of which, the null model, is a special case of the other, the alternative model. The test is based on the likelihood ratio, which expresses how many times more likely the data are under one model than the other. This log-likelihood ratio can then be used to compute a p-value, or compared to a critical value, to decide whether to reject the null model in favor of the alternative model.

The log-likelihood ratio test LR_N (LR_T), used to test the null hypothesis that log-returns are normally (student's t) distributed against the alternative hypothesis, is given by

$$LR_N \text{ or } LR_T = -2(LR_r - LR_u) \sim \chi^2(m) \tag{51.12}$$

where LR_r and LR_u are, respectively, the maximum value of the log-likelihood values under the null hypothesis of the restricted model and the alternative hypothesis of the unrestricted model and m is the number of the restricted parameters in the restricted model. For example, LR_N for GARCH-SGT model could be used to test the null hypothesis that log-returns are normally distributed against the alternative hypothesis that they are SGT distributed. The null hypothesis for testing normality is $H_0{:}\kappa = 2$, $\lambda = 0$ and $n \to \infty$, and the alternative hypothesis is $H_1{:}\kappa \in R^+$, $n > 2$ and $|\lambda| < 1$. Restate, $LR_N = -2(LR_r - LR_u) \sim \chi^2(3)$ where LR_r and LR_u are, respectively, the maximum value of the log-likelihood values under the null hypothesis of restricted model (GARCH-N model) and the alternative hypothesis of unrestricted model (GARCH-SGT model) and m is the number of the restricted parameters in the restricted model ($\kappa = 2$, $\lambda = 0$ and $n \to \infty$) and equal to 3 in this case. At the same inference, LR_N for GARCH-T model follows the $\chi^2(1)$ distribution with one degree of freedom. Moreover, LR_T for GARCH-SGT model follows the $\chi^2(2)$ distribution with two degrees of freedom.

51.3.2 Binary Loss Function or Failure Rate

If the predicted VaR is not able to cover the realized loss, this is termed a violation. A binary loss function (BLF) is merely the reflection of the LR test of unconditional coverage test and gives a penalty of one to each exception of the VaR. The BLF for long position can be defined as follows:

$$BL_{t+1} = \begin{cases} 1 & \text{if } r_{t+1} < VaR_{t+1|t}, \\ 0 & \text{if } r_{t+1} \geq VaR_{t+1|t}. \end{cases} \tag{51.13}$$

where BL_{t+1} represents the 1-day-ahead BLF for long position. If a VaR model truly provides the level of coverage defined by its confidence level, then the average binary loss function (ABLF) or the failure rate over the full sample will equal c for the $(1-c)$th percentile VaR.

51.3.3 Quadratic Loss Function

The quadratic loss function (QLF) of Lopez (1999) penalizes violations differently from the binary loss function and pays attention to the magnitude of the violation. The QLF for long position can be expressed as

$$QL_{t+1} = \begin{cases} 1 + \left(r_{t+1} - VaR_{t+1|t}\right)^2 & \text{if } r_{t+1} < VaR_{t+1|t}, \\ 0 & \text{if } r_{t+1} \geq VaR_{t+1|t}. \end{cases} \tag{51.14}$$

where QL_{t+1} represents the 1-day-ahead QLF for long position. The quadratic term in Eq. 51.14 ensures that large violations are penalized more than the small violations which provides a more powerful measure of model accuracy than the binary loss function.

51.3.4 The Unconditional Coverage Test

Kupiec (1995) proposes the unconditional coverage test which is a likelihood ratio test for testing the model accuracy which is identical to a test of the null hypothesis that the probability of failure for each trial $(\hat{\pi})$ equals the specified model probability (p). The likelihood ratio test statistics is given by

$$LR_{uc} = -2\ln\left(p^{n_1}(1-p)^{n_0}\hat{\pi}^{-n_1}(1-\hat{\pi})^{-n_0}\right) \sim \chi^2(1) \tag{51.15}$$

where $\hat{\pi} = \frac{n_1}{n_0+n_1}$ is the maximum likelihood estimate of p, n_1 denotes a Bernoulli random variable representing the total number of VaR violations, and $n_0 + n_1$

represents the full sample size. The LR_{uc} test can be employed to test whether the sample point estimate is statistically consistent with the VaR model's prescribed confidence level or not.

51.3.5 Unexpected Loss

The unexpected loss (UL) will equal the average magnitude of the violation over the full sample. The magnitude of the violation for long position is given by

$$L_{t+1} = \begin{cases} r_{t+1} - VaR_{t+1|t} & \text{if } r_{t+1} < VaR_{t+1|t}, \\ 0 & \text{if } r_{t+1} \geq VaR_{t+1|t}. \end{cases} \tag{51.16}$$

where L_{t+1} is the 1-day-ahead magnitude of the violation for long position.

51.4 Empirical Results

The study data comprises daily prices of the following eight stock indices: the Austria ATX (6/29/1999–8/10/2009), the Belgium Brussels (10/19/1999–8/10/2009), the France CAC40 (10/22/1999–8/10/2009) and the Switzerland Swiss (9/8/1999–8/10/2009) in Europe, the India Bombay (7/8/1999–8/10/2009), the Malaysia KLSE (6/23/1999–8/10/2009), the South Korea KOSPI (6/21/1999–8/10/2009), and the Singapore STRAITS (8/24/1999–8/10/2009) in Asia, where the numbers in parentheses are the start and end dates for our sample. Daily closing spot prices for the study period, totaling 2,500 observations, were obtained from http://finance.yahoo.com. The stock returns are defined as the first difference in the logarithms of daily stock prices then multiplied by 100.

51.4.1 Data Preliminary Analysis

Table 51.1 summarizes the basic statistical characteristics of return series for both the estimation and forecast periods. Notably, the average daily returns are all negative (resp. positive) for forecast (resp. estimation) period and very small compared with the variable standard deviation, indicating high volatility. Except the Brussels of estimation period and the CAC40, Swiss and Bombay of forecast period, all returns series almost exhibit negative skewness for both the estimation and forecast periods. The excess kurtosis all significantly exceeds zero at the 1 % level, indicating a leptokurtic characteristic. Furthermore, J-B normality test statistics are all significant at the 1 % level and thus reject the hypothesis of normality and confirm that neither return series is normally distributed. Moreover, the Ljung-Box $Q^2(20)$ statistics for the squared returns are all significant at the 1 %

Table 51.1 Descriptive statistics of daily return

	Mean	Std. dev.	Max.	Min.	Skewness	Kurtosis	J-B	$Q^2(20)$
Panel A. Estimation period (2,000 observations)								
ATX	0.0675	0.9680	4.6719	−7.7676	−0.6673[c]	4.4547[c]	1,802.17[c]	547.23[c]
Brussels	0.0179	1.1577	9.3339	−5.6102	0.2607[c]	6.0567[c]	3,079.66[c]	1,479.84[c]
CAC40	0.0090	1.4012	7.0022	−7.6780	−0.0987[a]	2.9924[c]	749.48[c]	2,270.73[c]
Swiss	0.0086	1.1602	6.4872	−5.7803	−0.0530	4.5084[c]	1,694.78[c]	1,985.51[c]
Bombay	0.0648	1.5379	7.9310	−11.8091	−0.5632[c]	4.2350[c]	1,600.43[c]	707.26[c]
KLSE	0.0243	0.9842	5.8504	−6.3422	−0.3765[c]	6.2537[c]	3,306.35[c]	5,54.16[c]
KOSPI	0.0397	1.8705	7.6971	−12.8046	−0.4671[c]	3.6010[c]	1,153.39[c]	365.29[c]
STRAITS	0.0239	1.1282	4.9052	−9.0949	−0.5864[c]	4.8254[c]	2,055.03[c]	239.53[c]
Panel B. Forecast period (500 observations)								
ATX	−0.1352	2.6532	12.0210	−10.2526	−0.0360	2.4225[c]	122.37[c]	735.66[c]
Brussels	−0.1195	1.9792	9.2212	−8.3192	−0.0888	3.1545[c]	207.97[c]	581.42[c]
CAC40	−0.0907	2.1564	10.5945	−9.4715	0.2209[b]	4.4068[c]	408.65[c]	353.15[c]
Swiss	−0.0704	1.8221	10.7876	−8.1077	0.2427[b]	4.4101[c]	410.10[c]	502.59[c]
Bombay	−0.0101	2.6043	15.9899	−11.6044	0.2529[b]	3.4248[c]	249.70[c]	57.24[c]
KLSE	−0.0166	1.2040	4.2586	−9.9785	−1.1163[c]	9.8218[c]	2,113.64[c]	17.87[c]
KOSPI	−0.0327	2.1622	11.2843	−11.1720	−0.4177[c]	4.3977[c]	417.47[c]	343.50[c]
STRAITS	−0.0594	2.0317	7.5305	−9.2155	−0.1183	2.3827[c]	119.45[c]	219.52[c]

Notes: 1. [a], [b], and [c] denote significantly at the 10 %, 5 %, and 1 % levels, respectively. 2. J-B statistics are based on Jarque and Bera (1987) and are asymptotically chi-squared distributed with 2 degrees of freedom. 3. $Q^2(20)$ statistics are asymptotically chi-squared distributed with 20 degrees of freedom

level and thus indicate that the return series exhibit linear dependence and strong ARCH effects. Therefore, the preliminary analysis of the data suggests the use of a GARCH model to capture the fat-tails and time-varying volatility found in these stock indices return series.

Descriptive graphs (levels of spot prices, density of the daily returns against normal distribution) for each stock index are illustrated in Fig. 51.1a–h. As shown in Fig. 51.1, all stock indices have experienced a severe slide in price levels and display pictures of volatile bear markets for forecast period. Moreover, comparing density graphs against the normal distribution shows that each return distribution of data employed exhibits non-normal characteristics. This provides evidence in favor of some of skewed, leptokurtic, and fat-tailed return distributions. These results are in line with those of Table 51.1.

51.4.2 Estimation Results for Alternate VaR Models

This section estimates the GARCH(1,1) model with alternative distributions (normal, student's t, and SGT) for performing VaR analysis. For each data series, three GARCH models are estimated with a sample of 2,000 daily returns, and the estimation period is then rolled forwards by adding one new day and dropping the

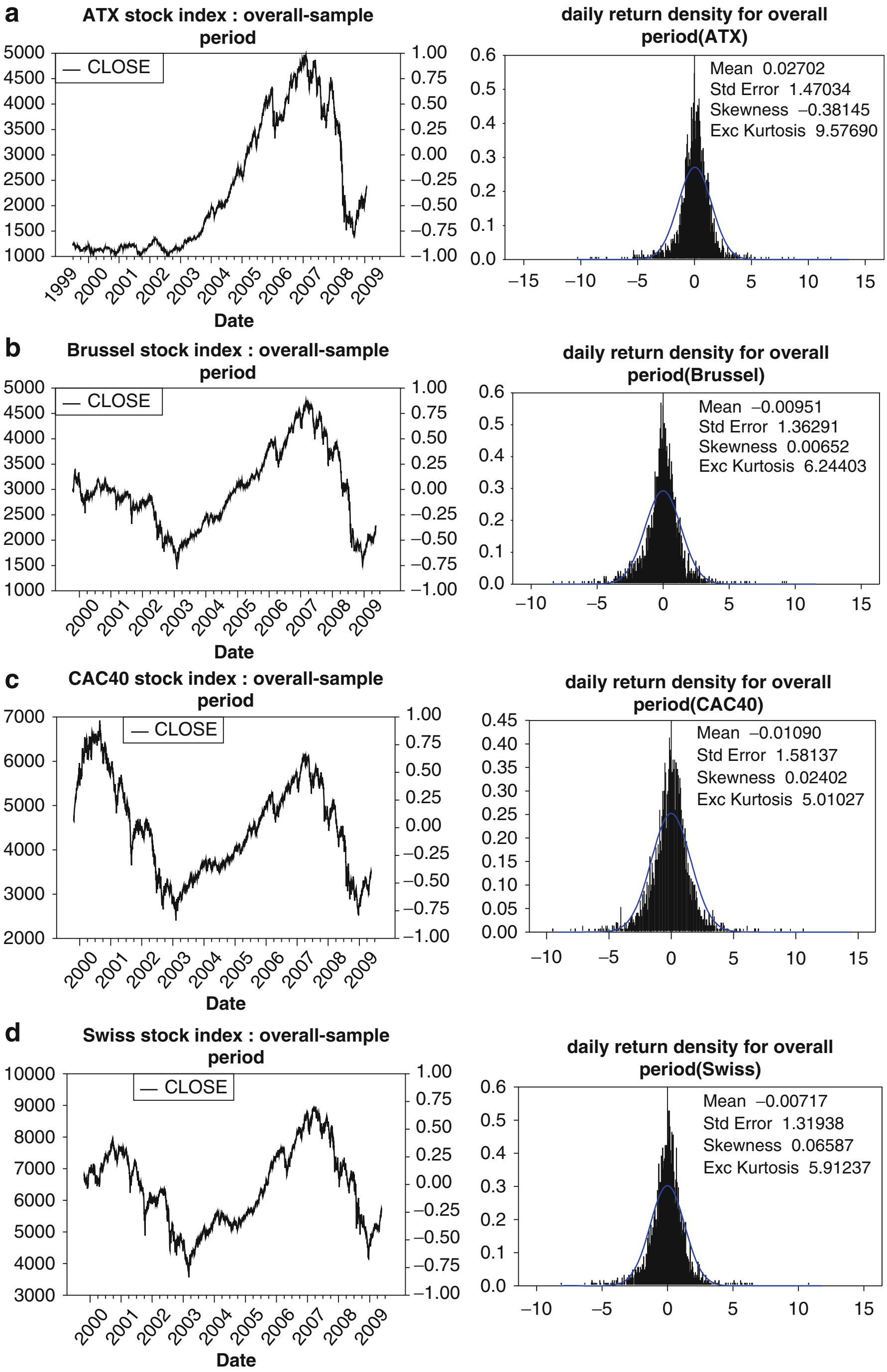


Fig. 51.1 (continued)

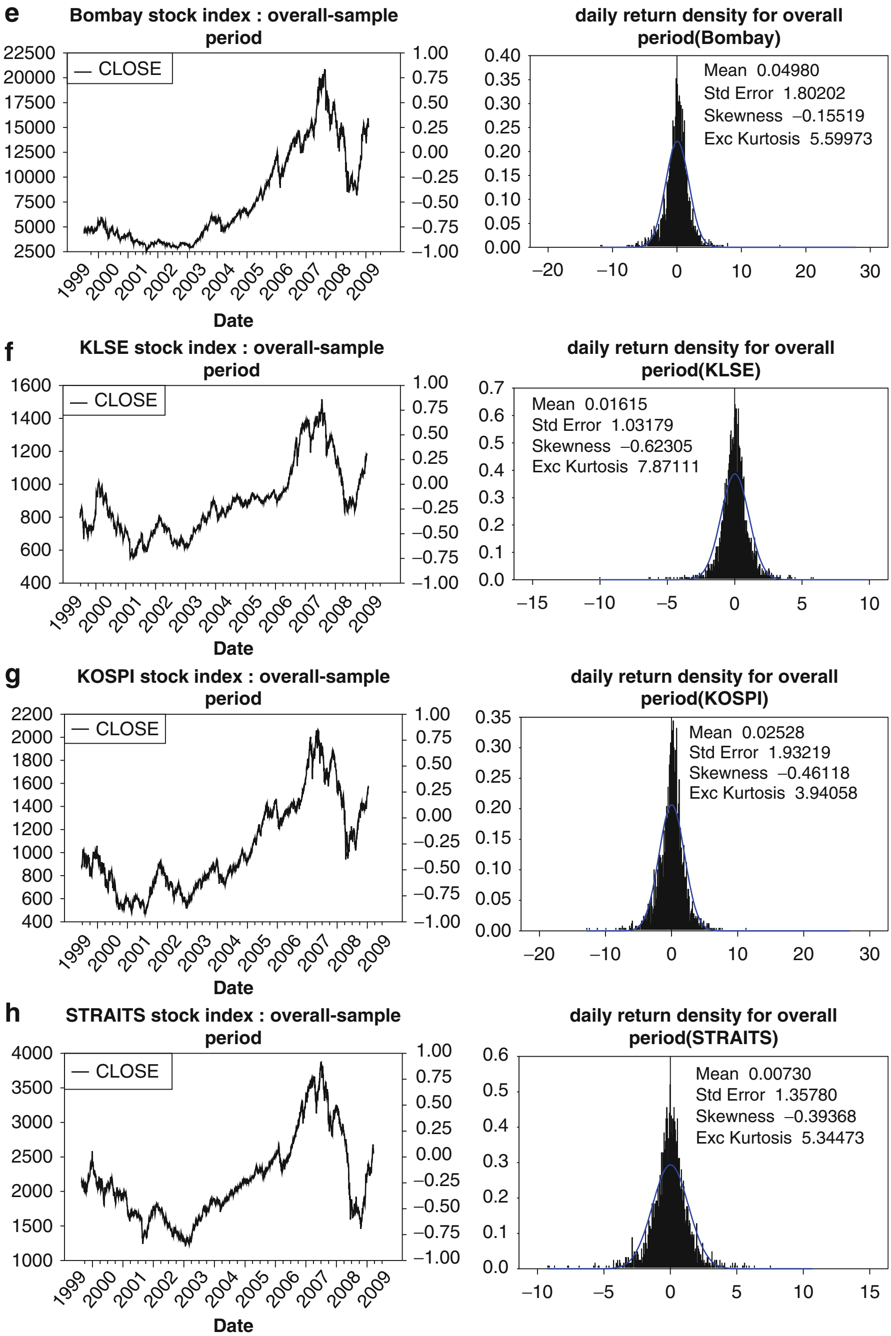


Fig. 51.1 The stock index in level and daily return density (versus normal) for whole sample (**a**) ATX, (**b**) Brussels, (**c**) CAC40, (**d**) Swiss, (**e**) Bombay, (**f**) KLSE, (**g**) KOSPI, (**h**) STRAITS stock indices

most distant day. In this procedure, according to the theory of Sect. 51.2, the out-of-sample VaR is computed for the next 500 days.

Table 51.2 lists the estimation results[4] of the GARCH-N, GARCH-T, and GARCH-SGT models for the ATX, Brussels, CAC40, and Swiss stock indices in Europe and the Bombay, KLSE, STRAITS, and KOSPI stock indices in Asia during the first in-sample period. The variance coefficients ω, α, and β are all positive and significant almost at the 1 % level. Furthermore, the sums of parameters α and β for these three models are less than one thus ensuring that the conditions for stationary covariance hold. As to the fat-tail parameters in student's t distribution, the fat-tail parameter (n) ranges from 4.9906 (KLSE) to 14.9758 (CAC40) for GARCH-T model. All these shape parameters are all significant at 1 % level and obey the constraint $n > 2$ and thereby implying that the distribution of returns has larger, thicker tails than the normal distribution. Turning to the shape parameters in SGT distribution, the fat-tail parameter (n) ranges from 4.9846 (KLSE) to 21.4744 (KOSPI), and the fat-tail parameter (κ) is between 1.5399 (KOSPI) and 2.3917 (Bombay). The skewness parameter (λ) ranges from -0.1560 (Bombay) to -0.0044(KLSE). Moreover, these three coefficients are almost significant at the 1 % level and thereby these negative skewness parameters imply that the distribution of returns has a leftward tail. Therefore, both fat-tails and skewness cannot be ignored in modeling these stock indices returns. The Ljung-Box $Q^2(20)$ statistics for the squared returns are all not significant at the 10 % level and thus indicate that serial correlation does not exist in standard residuals, confirming that the GARCH (1,1) specification in these models is sufficient to correct the serial correlation of these eight return series in the conditional variance equation.

Moreover, as shown in Table 51.2, the LR_N statistics for both GARCH-T and GARCH-SGT models are all significant at the 1 % level, indicating that the null hypothesis of normality for either stock index is rejected. These results thus imply that both the student's t and SGT distributions closely approximate the empirical return series as compared with the normal distribution. Furthermore, except for ATX and KLSE stock indices, the LR_T statistics of GARCH-SGT model are all significant, implying that the SGT distribution more closely approximates the empirical return series than the student's t does. To sum up, the SGT distribution closely approximates the empirical return series followed by student's t and normal distributions.

51.4.3 The Results of VaR Performance Assessment

In this paper, we utilize the parametric approach (GARCH-N, GARCH-T, and GARCH-SGT models) and the semi-parametric approach (HW-N, HW-T, and HW-SGT models), totaling six models, to estimate the VaR.; thereafter, it was compared with the observed return, and both results were recorded. This section

[4]The parameters are estimated by QMLE (quasi-maximum likelihood estimation; QMLE) and the BFGS optimization algorithm, using the econometric package of WinRATS 6.1.

Table 51.2 Estimation results for alternative models (estimation period)

	ATX	Brussels	CAC40	Swiss
Panel A. GARCH(1,1) with normal distribution				
μ	0.1007^{c}(0.0202)	0.0760^{c}(0.0177)	0.0537^{b}(0.0227)	0.0488^{c}(0.0187)
ω	0.0612^{c}(0.0124)	0.0239^{c}(0.0032)	0.0143^{c}(0.0029)	0.0243^{c}(0.0057)
α	0.1146^{c}(0.0158)	0.1470^{c}(0.0090)	0.0799^{c}(0.0100)	0.1184^{c}(0.0136)
β	0.8209^{c}(0.0227)	0.8354^{c}(0.0034)	0.9132^{c}(0.0105)	0.8632^{c}(0.0150)
$Q^2(20)$	16.147	14.731	22.333	19.883
LL	−2648.73	−2649.75	−3156.40	−2748.86
Panel B. GARCH(1,1) with student's t distribution				
μ	0.0998^{c}(0.0178)	0.0775^{c}(0.0162)	0.0622^{c}(0.0216)	0.0589^{c}(0.0167)
ω	0.0623^{c}(0.0160)	0.0179^{c}(0.0049)	0.0118^{c}(0.0043)	0.0177^{c}(0.0052)
α	0.0986^{c}(0.0188)	0.1319^{c}(0.0177)	0.0785^{c}(0.0110)	0.1078^{c}(0.0151)
β	0.8324^{c}(0.0292)	0.8560^{c}(0.0180)	0.9166^{c}(0.0109)	0.8799^{c}(0.0153)
n	7.3393^{c}(1.0609)	9.4946^{c}(1.7035)	14.9758^{c}(4.0264)	9.6205^{c}(1.6621)
$Q^2(20)$	17.334	19.676	21.719	18.712
LL	−2610.15	−2626.31	−3146.96	−2724.27
LR_N	77.16^{c}	46.88^{c}	18.88^{c}	49.18^{c}
Panel C. GARCH(1,1) with SGT distribution				
μ	0.0875^{c}(0.0177)	0.0691^{c}(0.0158)	0.0525^{b}(0.0217)	0.0479^{c}(0.0175)
ω	0.0626^{c}(0.0173)	0.0175^{c}(0.0044)	0.0115^{c}(0.0043)	0.0172^{c}(0.0048)
α	0.0952^{c}(0.0189)	0.1277^{c}(0.0163)	0.0774^{c}(0.0103)	0.1086^{c}(0.0148)
β	0.8343^{c}(0.0323)	0.8590^{c}(0.0161)	0.9170^{c}(0.0106)	0.8787^{c}(0.0150)
n	7.8001^{c}(2.4847)	6.9261^{c}(1.7169)	13.3004^{b}(6.3162)	7.8987^{c}(2.1709)
λ	−0.0660^{b}(0.0290)	−0.1019^{c}(0.0346)	−0.1175^{c}(0.0335)	−0.1175^{c}(0.0323)
κ	1.9710^{c}(0.2290)	2.3745^{c}(0.2782)	2.1219^{c}(0.2220)	2.2601^{c}(0.2519)
$Q^2(20)$	17.779	20.509	21.803	18.791
LL	−2608.01	−2621.51	−3140.58	−2717.94
LR_N (LR_T)	81.44^{c}(4.28)	56.48^{c}(9.6^{c})	31.64^{c}(12.76^{c})	61.84^{c}(12.66^{c})
	Bombay	KLSE	KOSPI	STRAITS
Panel A. GARCH(1,1) with normal distribution				
μ	0.1427^{c}(0.0262)	0.0453^{c}(0.0164)	0.1212^{c}(0.0318)	0.0623^{c}(0.0196)
ω	0.0906^{c}(0.0201)	0.0077^{c}(0.0029)	0.0214^{c}(0.0082)	0.0143^{c}(0.0042)
α	0.1438^{c}(0.0167)	0.0998^{c}(0.0174)	0.0799^{c}(0.0146)	0.1031^{c}(0.0134)
β	0.8189^{c}(0.0206)	0.8989^{c}(0.0165)	0.9177^{c}(0.0141)	0.8938^{c}(0.0123)
$Q^2(20)$	19.954	27.905	11.214	15.574
LL	−3453.18	−2490.05	−3843.01	−2895.36
Panel B. GARCH(1,1) with student's t distribution				
μ	0.1583^{c}(0.0250)	0.0351^{b}(0.0142)	0.1377^{c}(0.0302)	0.0652^{c}(0.0192)
ω	0.0863^{c}(0.0222)	0.0116^{b}(0.0050)	0.0163^{b}(0.0078)	0.0135^{c}(0.0049)
α	0.1417^{c}(0.0198)	0.1116^{c}(0.0254)	0.0639^{c}(0.0128)	0.0806^{c}(0.0136)
β	0.8225^{c}(0.0234)	0.8848^{c}(0.0248)	0.9332^{c}(0.0127)	0.9119^{c}(0.0139)
n	8.3410^{c}(1.3143)	4.9906^{c}(0.5782)	7.2792^{c}(1.1365)	6.7643^{c}(0.9319)

(*continued*)

Table 51.2 (continued)

	Bombay	KLSE	KOSPI	STRAITS
$Q^2(20)$	19.980	24.477	11.304	16.554
LL	−3420.18	−2412.82	−3801.67	−2838.34
LR_N	66.0^c	154.46^c	82.68^c	114.04^c
Panel C. GARCH(1,1) with SGT distribution				
μ	0.1266^c(0.0261)	0.0341^b(0.0147)	0.1021^c(0.0285)	0.0516^c(0.0185)
ω	0.0836^c(0.0201)	0.0116^b(0.0049)	0.0167^b(0.0077)	0.0132^c(0.0045)
α	0.1350^c(0.0196)	0.1117^c(0.0242)	0.0613^c(0.0133)	0.0785^c(0.0135)
β	0.8282^c(0.0228)	0.8847^c(0.0240)	0.9345^c(0.0135)	0.9138^c(0.0136)
n	6.2282^c(1.3602)	4.9846^c(1.0922)	21.4744(15.8310)	6.2641^c(1.4297)
λ	−0.1560^c(0.0303)	−0.0044(0.0281)	−0.1006^c(0.0266)	−0.0745^b(0.0296)
κ	2.3917^c(0.2801)	2.0016^c(0.2450)	1.5399^c(0.1513)	2.1194^c(0.2256)
$Q^2(20)$	21.167	24.455	11.067	16.595
LL	−3408.46	−2412.81	−3791.66	−2835.52
LR_N (LR_T)	89.44^c(23.44^c)	154.48^c(0.02)	102.7^c(20.02^c)	119.68^c(5.64^a)

Notes: 1.$^{a, b}$ and^c denote significantly at the 10 %, 5 %, and 1 % levels, respectively. 2. Numbers in parentheses are standard errors. 3. LL indicates the log-likelihood value. 4. The critical value of the LR_N test statistics at the 1 % significance level is 6.635 for GARCH-T and 11.345 for GARCH-SGT model. 5. The critical value of the LR_T test statistics at the 10 %, 5 %, and 1 % significance level is 4.605, 5.991, and 9.210 for GARCH-SGT model, respectively. 6. $Q^2(20)$ statistics are asymptotically chi-squared distributed with 20 degrees of freedom

then uses three accuracy measures: one likelihood ratio test (the unconditional coverage test (LR_{uc}) of Kupiec (1995)) and two loss functions (the average quadratic loss function (AQLF) of Lopez (1999) and the unexpected loss (UL)) to compare the forecasting ability of the aforementioned models in terms of VaR.

Figure 51.2 graphically illustrates the long VaR forecasts of the GARCH-N, GARCH-T, and GARCH-SGT models at alternate levels (95 %, 99 %, and 99.5 %) for all stock indices. Tables 51.3, 51.4 and 51.5 provide the failure rates and the results of the prior three accuracy evaluation tests (LR_{uc}, AQLF, and UL) for the aforementioned six models at the 95 %, 99 %, and 99.5 % confidence levels, respectively. As observed in Tables 51.3, 51.4 and 51.5, we find that, except for a few cases at the 99 % and 99.5 % confidence levels, all models tend to underestimate real market risk because the empirical failure rate is higher than the theoretical failure rate in most cases. The abovementioned exceptional cases emerge at the GARCH-SGT model of 99 % level (CAC40); both the GARCH-T and GARCH-SGT models of 99.5 % level (KLSE and STRAITS); the HW-N (KLSE), HW-T (KLSE), and HW-SGT (KLSE and STRAITS) models of 99 % level; and the HW-N (STRAITS), HW-T (ATX and STRAITS), and HW-SGT (ATX, KLSE, and STRAITS) models of 99.5 % level, where the stock indices in parentheses behind the models are the exceptional cases. Moreover, the empirical failure rate of the above exceptional cases is lower than the theoretical failure rate, indicating that the non-normal distributions (student's t and SGT) and the semi-parametric approach try to reverse the trend of underestimating real market risk, especially at the 99.5 % level.

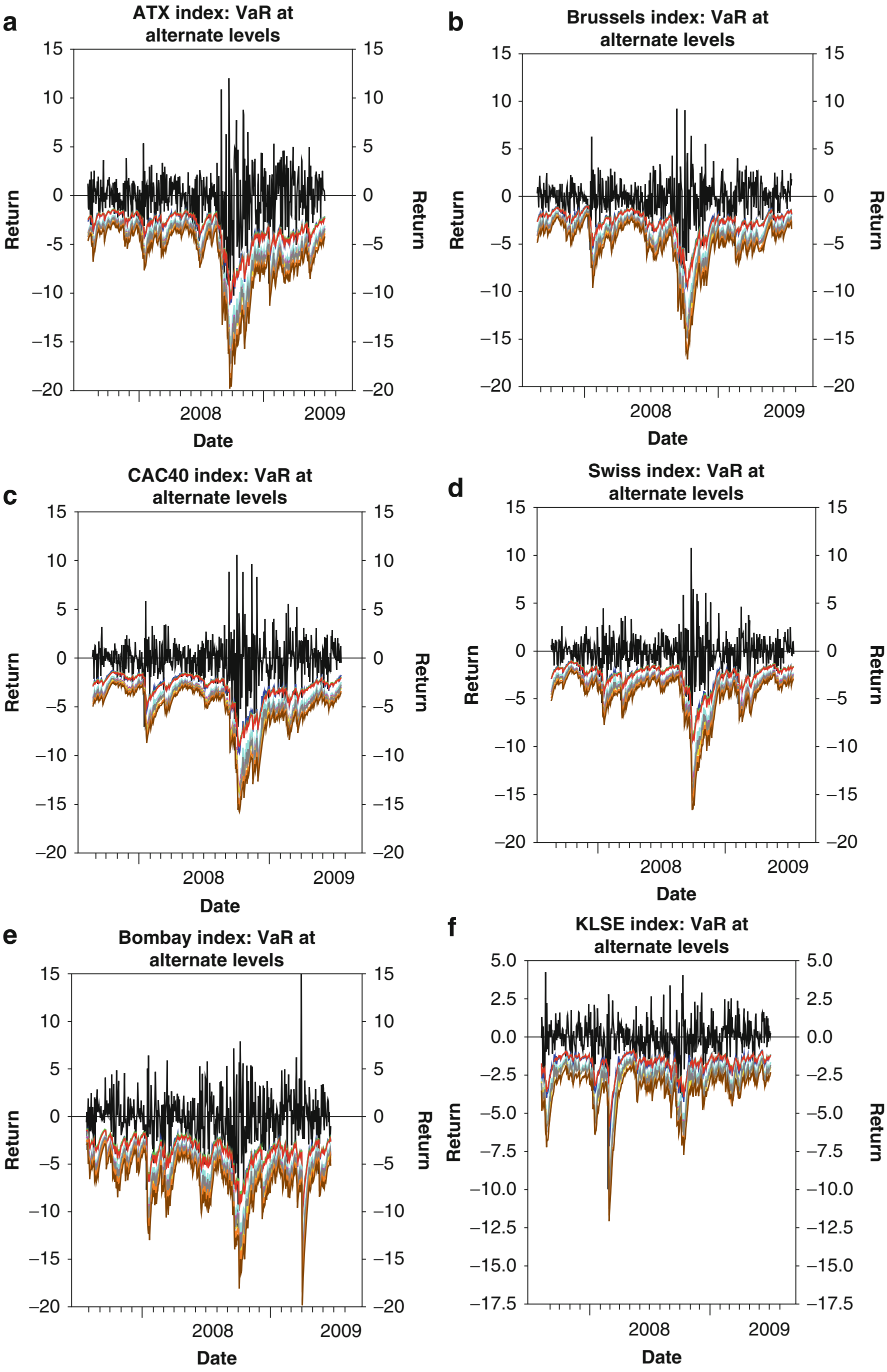


Fig. 51.2 (continued)

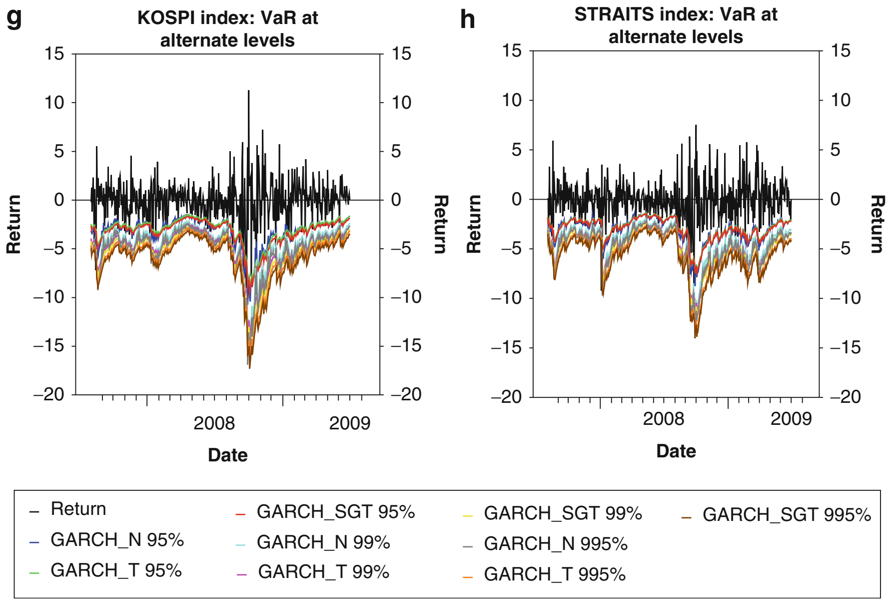


Fig. 51.2 Long VaR forecasts at alternative level for the normal, student's t, and SGT distribution. (**a**) ATX, (**b**) Brussels, (**c**) CAC40, (**d**) Swiss, (**e**) Bombay, (**f**) KLSE, (**g**) KOSPI, (**h**) STRAITS stock indices

As to the back-testing, the back-testing is a specific type of historical testing that determines the performance of the strategy if it had actually been employed during the past periods and market conditions. In this paper, the unconditional coverage tests (LR_{uc}) proposed by Kupiec (1995) is employed to test whether the unconditional coverage rate is statistically consistent with the VaR model's prescribed confidence level and thus is applied as the back-testing to measure the accuracy performance of these six VaR models. To interpret the result of accepting back-testing in Tables 51.3, 51.4 and 51.5, there is an illustration in the following. In Table 51.3, the VaR estimates based on GARCH-N, GARCH-T, and GARCH-SGT models, respectively, have a total of 2 (KLSE and STRAITS), 2 (KLSE and STRAITS), and 5 (Brussels, Bombay, KLSE, STRAITS, and KOSPI) acceptances for the LR_{uc} test when applying to all stock indices returns under 95 % confidence level, where the stock indices in parentheses behind the number are the acceptance cases. For 99 % confidence level, Table 51.4 shows that the GARCH-N, GARCH-T, and GARCH-SGT models pass the LR_{uc} tests with a total of 1, 5, and 8 stock indices, respectively; for 99.5 % confidence level, Table 51.5 gives that the GARCH-N, GARCH-T, and GARCH-SGT models pass the LR_{uc} tests with a total of 3, 7, and 7 stock indices, respectively. Hence, under all confidence levels, there is a total of 6, 14, and 20 acceptances for GARCH-N, GARCH-T, and GARCH-SGT models (the parametric approach), respectively. On the contrary, for 95 % confidence

Table 51.3 Out-of-sample long VaR performance at the 95 % confidence level

	GARCH			HW-GARCH		
	Failure rate (LR_{uc})	AQLF	UL	Failure rate (LR_{uc})	AQLF	UL
Panel A. ATX						
N	0.0960(17.75)	~~0.30360~~	~~–0.10151~~	0.0960(17.75)	0.30642	–0.10308
T	0.0960(17.75)	0.32058	–0.10865	0.0920(15.04)	~~0.30342~~	~~–0.10384~~
SGT	0.0900(13.75)	~~0.29877~~	~~–0.10212~~	0.0980(19.18)	0.32263	–0.10858
Panel B. Brussels						
N	0.0780(7.10)	0.20246	–0.07099	0.0680(3.08*)	~~0.18374~~	~~–0.06673~~
T	0.0720(4.51)	0.19494	–0.07020	0.0620(1.41*)	~~0.16978~~	~~–0.06268~~
SGT	0.0660(2.45*)	0.18178	–0.06602	0.0620(1.41*)	~~0.17067~~	~~–0.06262~~
Panel C. CAC40						
N	0.0740(5.31)	0.21446	–0.06485	0.0700(3.76*)	~~0.20928~~	~~–0.05844~~
T	0.0800(8.07)	0.22598	–0.06887	0.0720(4.51)	~~0.19510~~	~~–0.05617~~
SGT	0.0760(6.18)	0.21146	–0.06281	0.0700(3.76*)	~~0.19390~~	~~–0.05600~~
Panel D. Swiss						
N	0.0760(6.18)	0.18567	–0.06181	0.0620(1.41*)	~~0.15104~~	~~–0.05088~~
T	0.0740(5.31)	0.18729	–0.06345	0.0600(0.99*)	~~0.14222~~	~~–0.04694~~
SGT	0.0740(5.31)	0.17506	–0.05698	0.0560(0.36*)	~~0.13928~~	~~–0.04697~~
Panel E. Bombay						
N	0.0780(7.10)	0.34428	–0.10214	0.0800(8.07)	~~0.33643~~	~~–0.09967~~
T	0.0820(9.11)	0.35999	–0.10617	0.0780(7.10)	~~0.33182~~	~~–0.09639~~
SGT	0.0700(3.76*)	~~0.31488~~	~~–0.09366~~	0.0800(8.07)	0.33629	–0.09824
Panel F. KLSE						
N	0.0560(0.36*)	~~0.20084~~	~~–0.04595~~	0.0740(5.31)	0.23255	–0.05623
T	0.0600(0.99*)	~~0.21827~~	~~–0.05152~~	0.0700(3.76*)	0.23146	–0.05276
SGT	0.0580(0.64*)	~~0.21375~~	~~–0.05007~~	0.0640(1.90*)	0.22378	–0.05306
Panel G. STRAITS						
N	0.0580(0.64*)	~~0.21980~~	~~–0.06233~~	0.0640(1.90*)	0.24194	–0.06675
T	0.0640(1.90*)	0.25978	–0.07222	0.0620(1.41*)	~~0.23921~~	~~–0.06511~~
SGT	0.0620(1.41*)	0.24670	–0.06736	0.0560(0.36*)	~~0.23035~~	~~–0.06428~~
Panel H. KOSPI						
N	0.0740(5.31)	0.27949	–0.08826	0.0600(0.99*)	~~0.24267~~	~~–0.07816~~
T	0.0760(6.18)	0.32360	–0.09624	0.0620(1.41*)	~~0.25234~~	~~–0.07522~~
SGT	0.0672(2.82*)	0.27616	–0.08236	0.0620(1.41*)	~~0.24475~~	~~–0.07531~~

Notes: 1. *Indicates that the model passes the unconditional coverage test at the 5 % significance level and the critical value of the LR_{uc} test statistics at the 5 % significance level is 3.84. 2. The *red* (resp. *blue*) *font* represents the lowest (resp. highest) AQLF and unexpected loss when the predictive accuracies of three different innovations with the same VaR method are compared. 3. The *delete-line font* represents the lowest AQLF and unexpected loss when the predictive accuracies of two different VaR methods with the same innovation are compared. 4. The model acronyms stand for the following methods: *HW-GARCH* non-parametric method proposed by Hull and White (1998), *GARCH* parametric method of GARCH model, *N* the standard normal distribution, *T* the standardized student's t distribution, *SGT* the standardized SGT distribution proposed by Theodossiou (1998)

Table 51.4 Out-of-sample long VaR performance at the 99 % confidence level

	GARCH			HW-GARCH		
	Failure rate (LR_{uc})	AQLF	UL	Failure rate (LR_{uc})	AQLF	UL
Panel A. ATX						
N	0.0300(13.16)	0.20577	–0.02704	0.0160(1.53*)	~~0.06164~~	~~–0.01770~~
T	0.0200(3.91)	0.19791	–0.01986	0.0160(1.53*)	~~0.04977~~	~~–0.01551~~
SGT	0.0160(1.53*)	0.17702	–0.01725	0.0160(1.53*)	~~0.05279~~	~~–0.01626~~
Panel B. Brussels						
N	0.0260(8.97)	0.13491	–0.02538	0.0180(2.61*)	~~0.03685~~	~~–0.01726~~
T	0.0180(2.61*)	0.11962	–0.01947	0.0160(1.53*)	~~0.03465~~	~~–0.01559~~
SGT	0.0160(1.53*)	0.11033	–0.01711	0.0160(1.53*)	~~0.03536~~	~~–0.01596~~
Panel C. CAC40						
N	0.0200(3.91)	0.14794	–0.01913	0.0160(1.53*)	~~0.07472~~	~~–0.01808~~
T	0.0100(0.00*)	0.14325	–0.01494	0.0100(0.00*)	~~0.05628~~	~~–0.01436~~
SGT	0.0060(0.94*)	0.13144	~~–0.01354~~	0.0140(0.71*)	~~0.05992~~	–0.01466
Panel D. Swiss						
N	0.0260(8.97)	0.12387	–0.01967	0.0180(2.61*)	~~0.03938~~	~~–0.01353~~
T	0.0160(1.53*)	0.11943	–0.01516	0.0140(0.71*)	~~0.03528~~	~~–0.01341~~
SGT	0.0140(0.71*)	0.10151	~~–0.01244~~	0.0140(0.71*)	~~0.03512~~	~~–0.01338~~
Panel E. Bombay						
N	0.0300(13.16)	0.23811	–0.03748	0.0120(0.18*)	~~0.04961~~	~~–0.01751~~
T	0.0220(5.41)	0.22706	–0.02822	0.0120(0.18*)	~~0.04539~~	~~–0.01586~~
SGT	0.0180(2.61*)	0.19232	–0.02152	0.0120(0.18*)	~~0.04504~~	~~–0.01617~~
Panel F. KLSE						
N	0.0160(1.53*)	0.16522	–0.01892	0.0080(0.21*)	~~0.07681~~	~~–0.01341~~
T	0.0100(0.00*)	0.16891	–0.01585	0.0060(0.94*)	~~0.08114~~	~~–0.01425~~
SGT	0.0100(0.00*)	0.16278	–0.01551	0.0060(0.94*)	~~0.07965~~	~~–0.01384~~
Panel G. STRAITS						
N	0.0240(7.11)	0.16096	–0.01848	0.0120(0.18*)	~~0.09107~~	~~–0.01763~~
T	0.0100(0.00*)	0.17606	–0.01568	0.0100(0.00*)	~~0.07477~~	~~–0.01403~~
SGT	0.0100(0.00*)	0.16258	–0.01406	0.0080(0.21*)	~~0.07278~~	~~–0.01361~~
Panel H. KOSPI						
N	0.0220(5.41)	0.18050	–0.02675	0.0200(3.91)	~~0.03799~~	~~–0.01639~~
T	0.0200(3.91)	0.20379	–0.02199	0.0180(2.61*)	~~0.04722~~	~~–0.01942~~
SGT	0.0163(1.68*)	0.15465	~~–0.01563~~	0.0180(2.61*)	~~0.04487~~	–0.01941

Note: 1. *indicates that the model passes the unconditional coverage test at the 5 % significance level and the critical value of the LR_{uc} test statistics at the 5 % significance level is 3.84. 2. The *red* (resp. *blue*) *font* represents the lowest (resp. highest) AQLF and unexpected loss when the predictive accuracies of three different innovations with the same VaR method are compared. 3. The *delete-line font* represents the lowest AQLF and unexpected loss when the predictive accuracies of two different VaR methods with the same innovation are compared. 4. The model acronyms stand for the following methods: *HW-GARCH* non-parametric method proposed by Hull and White (1998), *GARCH* parametric method of GARCH model, *N* the standard normal distribution, *T* the standardized student's t distribution, *SGT* the standardized SGT distribution proposed by Theodossiou (1998)

Table 51.5 Out-of-sample long VaR performance at the 99.5 % confidence level

	GARCH			HW-GARCH		
	Failure rate (LR_{uc})	AQLF	UL	Failure rate (LR_{uc})	AQLF	UL
Panel A. ATX						
N	0.0960(17.75)	~~0.30360~~	~~–0.10151~~	0.0960(17.75)	0.30642	–0.10308
T	0.0960(17.75)	0.32058	–0.10865	0.0920(15.04)	~~0.30342~~	~~–0.10384~~
SGT	0.0900(13.75)	~~0.29877~~	~~–0.10212~~	0.0980(19.18)	0.32263	–0.10858
Panel B. Brussels						
N	0.0780(7.10)	0.20246	–0.07099	0.0680(3.08*)	~~0.18374~~	~~–0.06673~~
T	0.0720(4.51)	0.19494	–0.07020	0.0620(1.41*)	~~0.16978~~	~~–0.06268~~
SGT	0.0660(2.45*)	0.18178	–0.06602	0.0620(1.41*)	~~0.17067~~	~~–0.06262~~
Panel C. CAC40						
N	0.0740(5.31)	0.21446	–0.06485	0.0700(3.76*)	~~0.20928~~	~~–0.05844~~
T	0.0800(8.07)	0.22598	–0.06887	0.0720(4.51)	~~0.19510~~	~~–0.05617~~
SGT	0.0760(6.18)	0.21146	–0.06281	0.0700(3.76*)	~~0.19390~~	~~–0.05600~~
Panel D. Swiss						
N	0.0760(6.18)	0.18567	–0.06181	0.0620(1.41*)	~~0.15104~~	~~–0.05088~~
T	0.0740(5.31)	0.18729	–0.06345	0.0600(0.99*)	~~0.14222~~	~~–0.04694~~
SGT	0.0740(5.31)	0.17506	–0.05698	0.0560(0.36*)	~~0.13928~~	~~–0.04697~~
Panel E. Bombay						
N	0.0780(7.10)	0.34428	–0.10214	0.0800(8.07)	~~0.33643~~	~~–0.09967~~
T	0.0820(9.11)	0.35999	–0.10617	0.0780(7.10)	~~0.33182~~	~~–0.09639~~
SGT	0.0700(3.76*)	~~0.31488~~	~~–0.09366~~	0.0800(8.07)	0.33629	–0.09824
Panel F. KLSE						
N	0.0560(0.36*)	~~0.20084~~	~~–0.04595~~	0.0740(5.31)	0.23255	–0.05623
T	0.0600(0.99*)	~~0.21827~~	~~–0.05152~~	0.0700(3.76*)	0.23146	–0.05276
SGT	0.0580(0.64*)	~~0.21375~~	~~–0.05007~~	0.0640(1.90*)	0.22378	–0.05306
Panel G. STRAITS						
N	0.0580(0.64*)	~~0.21980~~	~~–0.06233~~	0.0640(1.90*)	0.24194	–0.06675
T	0.0640(1.90*)	0.25978	–0.07222	0.0620(1.41*)	~~0.23921~~	~~–0.06511~~
SGT	0.0620(1.41*)	0.24670	–0.06736	0.0560(0.36*)	~~0.23035~~	~~–0.06428~~
Panel H. KOSPI						
N	0.0740(5.31)	0.27949	–0.08826	0.0600(0.99*)	~~0.24267~~	~~–0.07816~~
T	0.0760(6.18)	0.32360	–0.09624	0.0620(1.41*)	~~0.25234~~	~~–0.07522~~
SGT	0.0672(2.82*)	0.27616	–0.08236	0.0620(1.41*)	~~0.24475~~	~~–0.07531~~

Note: 1. *Indicates that the model passes the unconditional coverage test at the 5 % significance level and the critical value of the LR_{uc} test statistics at the 5 % significance level is 3.84. 2. The *red* (resp. *blue*) *font* represents the lowest (resp. highest) AQLF and unexpected loss when the predictive accuracies of three different innovations with the same VaR method are compared. 3. The *delete-line font* represents the lowest AQLF and unexpected loss when the predictive accuracies of two different VaR methods with the same innovation are compared. 4. The model acronyms stand for the following methods: *HW-GARCH* non-parametric method proposed by Hull and White (1998), *GARCH* parametric method of GARCH model, *N* the standard normal distribution, *T* the standardized student's t distribution, *SGT* the standardized SGT distribution proposed by Theodossiou (1998) level, Table 51.3 describes that the HW-N, HW-T, and HW-SGT models pass the LR_{uc} tests with a total of 5, 5, and 6 stock indices, respectively. Moreover, for 99 % confidence level, Table 51.4 depicts that the HW-N, HW-T, and HW-SGT models pass the LR_{uc} tests with a total of 7, 8, and 8 stock indices, respectively; for 99.5 % confidence level, Table 51.5 illustrates that the HW-N, HW-T, and HW-SGT models pass the LR_{uc} tests with a total of 7, 8, and 8 stock indices, respectively. Hence, under all confidence levels, there is a total of 19, 21, and 22 acceptances for HW-N, HW-T, and HW-SGT models (the semi-parametric approach), respectively.

From the abovementioned results, we can find the following two important phenomena: First, under the same return distributional setting, the number of acceptance of the HW-based models is greater or equal than those of the GARCH-based models, irrespective of whether the case of individual level (95 %, 99 %, or 99.5 %) or all levels (95 %, 99 %, and 99.5 %) is considered. For example, with regard to all levels, the number of acceptance of the HW-N model (19) is greater than those of the GARCH-N models (6). These results reveal that the HW-based models (semi-parametric approach) have the better VaR forecasting performance as compared with GARCH-based models (parametric approach). Second, the number of acceptance of the SGT distribution is the greatest followed by the student's t and normal distributions, irrespective of whether the GARCH-based model (parametric) or HW-based model (semi-parametric approach) is employed. For instance, with regard to all levels, the number of acceptance of the GARCH-SGT model (20) is the greatest followed by the GARCH-T model (14) and GARCH-N model (6). These results indicate that the SGT has the best VaR forecasting performance followed by student's t while the normal owns the worst VaR forecasting performance.

Turning to the other two accuracy measures (i.e., AQLF and UL), the two loss functions (the average quadratic loss function (AQLF) and the unexpected loss (UL)) reflect the magnitude of the violation which occur as the observed return exceeds the VaR estimation. The smaller the AQLF and UL are generated, the better the forecasting performance of the models is. As observed in Tables 51.3, 51.4 and 51.5, we can also find the following two important phenomena which are similar as those of the back-testing as was mentioned above: First, under the same return distributional setting, the AQLF and UL generated by the HW-based models are smaller than those generated by the GARCH-based models, irrespective of whether the 95 %, 99 %, or 99.5 % level is considered. These results reveal that the HW-based models (semi-parametric approach) significantly have the better VaR forecasting performance as compared with GARCH-based models (parametric approach), which is in line with the results of the back-testing. Second, for all confidence levels, the GARCH-SGT model yields the lowest AQLF and UL for most of the stock indices. Moreover, for most of the stock indices, the GARCH-N model produces the highest AQLF and UL for both 99 % and 99.5 % levels, while the GARCH-T model gives the highest AQLF and UL for 95 % level. These results indicate that the GARCH-SGT model significantly owns the best out-of-sample VaR performance, while the GARCH-N model appears to have the worst out-of-sample VaR performance. On the contrary, for all confidence levels, the HW-N model bears the highest AQLF and UL for most of the stock indices, while the HW-SGT model gives the lowest AQLF and UL for half of the stock indices, indicating that the HW-N model significantly owns the worst out-of-sample VaR performance, while the HW-SGT model appears to bear the highest out-of-sample VaR performance. Consequently, it seems reasonable to conclude that the SGT has the best VaR forecasting performance followed by student's t, while the normal owns the worst VaR forecasting performance, which appears to be consistent with the results of back-testing.

To sum up, according to the three accuracy measures, the HW-based models (semi-parametric approach) have the better VaR forecasting performance as compared with GARCH-based models (parametric approach), and the SGT has the best VaR forecasting performance followed by student's t, while the normal owns the worst VaR forecasting performance. In addition, the kind of VaR approach is more influential than that of return distribution setting on VaR estimate.

51.5 Conclusion

This study utilizes the parametric approach (GARCH-N, GARCH-T, and GARCH-SGT models) and the semi-parametric approach of Hull and White (1998) (HW-N, HW-T, and HW-SGT models), totaling six models, to estimate the VaR for the eight stock indices in Europe and Asia stock markets, then uses three accuracy measures: one likelihood ratio test (the unconditional coverage test (LR_{uc}) of Kupiec (1995)) and two loss functions (the average quadratic loss function (AQLF) of Lopez (1999) and the unexpected loss (UL)) to compare the forecasting ability of the aforementioned models in terms of VaR.

The empirical findings can be summarized as follows. First, according to the results of the log-likelihood ratio test, the SGT distribution closely approximates the empirical return series followed by student's t and normal distributions. Second, in terms of the failure rate, all models tend to underestimate the real market risk in most cases, but the non-normal distributions (student's t and SGT) and the semi-parametric approach try to reverse the trend of underestimating real market risk, especially at the 99.5 % level. Third, the kind of VaR approaches is more influential than that of return distribution settings on VaR estimate. Moreover, under the same return distributional setting, the HW-based models (semi-parametric approach) have the better VaR forecasting performance as compared with the GARCH-based models (parametric approach). Finally, irrespective of whether the GARCH-based model (parametric) or HW-based model (semi-parametric approach) is employed, the SGT has the best VaR forecasting performance followed by student's t, while the normal owns the worst VaR forecasting performance.

Appendix 1: The Left-Tailed Quantiles of the Standardized SGT

The standardized SGT distribution was derived by Lee and Su (2011) and expressed as follows:

$$f(\varepsilon_t) = C\left\{1 + \frac{|\varepsilon_t + \delta|^\kappa}{[1 + \mathrm{sign}(\varepsilon_t + \delta)\lambda]^\kappa \theta^\kappa}\right\}^{-\frac{n+1}{\kappa}} \tag{51.17}$$

where $\theta = \frac{1}{S(\lambda)} B\left(\frac{1}{\kappa}, \frac{n}{\kappa}\right)^{\frac{1}{2}} B\left(\frac{3}{\kappa}, \frac{n-2}{\kappa}\right)^{-\frac{1}{2}}, S(\lambda) = \sqrt{1 + 3\lambda^2 - 4A^2\lambda^2},$

$$A = B\left(\frac{2}{\kappa}, \frac{n-1}{\kappa}\right) B\left(\frac{1}{\kappa}, \frac{n}{\kappa}\right)^{-0.5} B\left(\frac{3}{\kappa}, \frac{n-2}{\kappa}\right)^{-0.5}, \ \delta = \frac{2\lambda A}{S(\lambda)}, \ C = \frac{\kappa}{2\theta} B\left(\frac{1}{\kappa}, \frac{n}{\kappa}\right)^{-1}$$

where κ, n, and λ are scaling parameters and C and θ are normalizing constants ensuring that f(•) is a proper p.d.f. The parameters κ and n control the height and tails of density with constraints $\kappa > 0$ and $n > 2$, respectively. The skewness parameter λ controls the rate of descent of the density around the mode of ε_t with $-1 < \lambda < 1$. In the case of positive (resp. negative) skewness, the density function skews toward the right (resp. left). Sign is the sign function, and B(•) is the beta function. The parameter n has the degrees of freedom interpretation in case $\lambda = 0$ and $\kappa = 2$. Particularly, the SGT distribution generates the student's t distribution for $\lambda = 0$ and $\kappa = 2$. Moreover, the SGT distribution generates the normal distribution for $\lambda = 0$, $\kappa = 2$, and $n = \infty$.

As observed from Table 51.2, the shape parameters in SGT distribution, the fat-tail parameter (n) ranges from 4.9846 (KLSE) to 21.4744 (KOSPI), and the fat-tail parameter (κ) is between 1.5399 (KOSPI) and 2.3917 (Bombay). The skewness parameter (λ) ranges from −0.1560 (Bombay) to −0.0044 (KLSE). Therefore, the left-tailed quantiles of the SGT distribution with various combinations of shape parameters ($-0.15 \le \lambda \le 0.05$; $1.0 \le \kappa \le 2.0$; $n = 10$) at alternate levels are obtained by the composite trapezoid rule and are listed in Table 51.6. Moreover, Fig. 51.3 depicts the left-tailed quantiles surface of SGT (versus normal) distribution with various combinations of shape parameters ($-0.25 \le \lambda \le 0.25$; $0.8 \le \kappa \le 2.0$; $n = 10$ and 20) at 10 %, 5 %, 1 %, and 0.5 % levels. Notably, $F_c(\varepsilon_t; \kappa = 2, \lambda = 0, n = \infty)$ where c = 0.1, 0.05, 0.01, and 0.005 in Fig. 51.3 represents the left-tailed quantiles of normal distribution at 10 %, 5 %, 1 %, and 0.5 % levels, which is −1.28155, −1.64486, −2.32638, and −2.57613, respectively.

Appendix 2: The Procedure of Parametric VaR Approach

The parametric method is very popular because the only variables you need to do the calculation are the mean and standard deviation of the portfolio, indicating the simplicity of the calculations. Moreover, from the literatures' review mentioned above, numerous studies focused on the parametric approach of the GARCH family

Table 51.6 The left-tailed quantiles of SGT distribution with n = 10 and various combinations (κ, λ) at alternate levels

κ\λ	−0.15	−0.10	−0.05	0.00	0.05
Panel A. 10 % level					
1.0	−1.6848(−1.1536)	−1.6443(−1.1359)	−1.6003(−1.1165)	−1.5532(−1.0958)	−1.5033(−1.0739)
1.1	−1.7005(−1.1831)	−1.6615(−1.1659)	−1.6193(−1.1472)	−1.5742(−1.1273)	−1.5266(−1.1062)
1.2	−1.7103(−1.2066)	−1.6729(−1.1901)	−1.6324(−1.1721)	−1.5893(−1.1530)	−1.5439(−1.1329)
1.3	−1.7160(−1.2257)	−1.6800(−1.2098)	−1.6413(−1.1927)	−1.6001(−1.1744)	−1.5568(−1.1552)
1.4	−1.7189(−1.2414)	−1.6843(−1.2261)	−1.6471(−1.2097)	−1.6078(−1.1923)	−1.5665(−1.1740)
1.5	−1.7196(−1.2544)	−1.6863(−1.2398)	−1.6508(−1.2241)	−1.6132(−1.2075)	−1.5738(−1.1901)
1.6	−1.7189(−1.2652)	−1.6869(−1.2512)	−1.6528(−1.2363)	−1.6168(−1.2204)	−1.5792(−1.2039)
1.7	−1.7171(−1.2744)	−1.6864(−1.2610)	−1.6536(−1.2467)	−1.6191(−1.2316)	−1.5831(−1.2159)
1.8	−1.7147(−1.2822)	−1.6850(−1.2694)	−1.6536(−1.2557)	−1.6205(−1.2413)	−1.5860(−1.2264)
1.9	−1.7116(−1.2888)	−1.6831(−1.2766)	−1.6528(−1.2635)	−1.6211(−1.2498)	−1.5881(−1.2356)
2.0	−1.7083(−1.2946)	−1.6807(−1.2828)	−1.6516(−1.2704)	−1.6211(−1.2573)	−1.5894(−1.2438)
Panel B. 5 % level					
1.0	−1.6848(−1.7252)	−1.6443(−1.6851)	−1.6003(−1.6418)	−1.5532(−1.5955)	−1.5033(−1.5468)
1.1	−1.7005(−1.7349)	−1.6615(−1.6966)	−1.6193(−1.6553)	−1.5742(−1.6114)	−1.5266(−1.5652)
1.2	−1.7103(−1.7397)	−1.6729(−1.7031)	−1.6324(−1.6639)	−1.5893(−1.6222)	−1.5439(−1.5785)
1.3	−1.7160(−1.7413)	−1.6800(−1.7064)	−1.6413(−1.6690)	−1.6001(−1.6294)	−1.5568(−1.5880)
1.4	−1.7189(−1.7406)	−1.6843(−1.7073)	−1.6471(−1.6716)	−1.6078(−1.6340)	−1.5665(−1.5947)
1.5	−1.7196(−1.7385)	−1.6863(−1.7065)	−1.6508(−1.6726)	−1.6132(−1.6368)	−1.5738(−1.5995)
1.6	−1.7189(−1.7353)	−1.6869(−1.7047)	−1.6528(−1.6723)	−1.6168(−1.6382)	−1.5792(−1.6027)
1.7	−1.7171(−1.7318)	−1.6864(−1.7021)	−1.6536(−1.6711)	−1.6191(−1.6387)	−1.5831(−1.6048)
1.8	−1.7147(−1.7270)	−1.6850(−1.6990)	−1.6536(−1.6694)	−1.6205(−1.6383)	−1.5860(−1.6061)
1.9	−1.7116(−1.7224)	−1.6831(−1.6955)	−1.6528(−1.6671)	−1.6211(−1.6375)	−1.5881(−1.6067)
2.0	−1.7083(−1.7177)	−1.6807(−1.6918)	−1.6516(−1.6646)	−1.6211(−1.6362)	−1.5894(−1.6068)

Panel C. 1 % level					
1.0	−3.1972(−3.1305)	−3.0964(−3.0358)	−2.9871(−2.9336)	−2.8702(−2.8247)	−2.7467(−2.7100)
1.1	−3.1281(−3.0522)	−3.0323(−2.9630)	−2.9289(−2.8673)	−2.8188(−2.7657)	−2.7029(−2.6589)
1.2	−3.0623(−2.9797)	−2.9711(−2.8956)	−2.8732(−2.8057)	−2.7693(−2.7106)	−2.6602(−2.6109)
1.3	−3.0005(−2.9132)	−2.9136(−2.8336)	−2.8207(−2.7489)	−2.7224(−2.6596)	−2.6195(−2.5661)
1.4	−2.9430(−2.8522)	−2.8600(−2.7768)	−2.7717(−2.6968)	−2.6785(−2.6126)	−2.5811(−2.5247)
1.5	−2.8896(−2.7963)	−2.8103(−2.7247)	−2.7261(−2.6489)	−2.6376(−2.5693)	−2.5452(−2.4864)
1.6	−2.8402(−2.7451)	−2.7643(−2.6769)	−2.6839(−2.6049)	−2.5995(−2.5295)	−2.5117(−2.4511)
1.7	−2.7945(−2.6980)	−2.7216(−2.6329)	−2.6447(−2.5644)	−2.5641(−2.4927)	−2.4804(−2.4184)
1.8	−2.7522(−2.6547)	−2.6821(−2.5924)	−2.6083(−2.5270)	−2.5312(−2.4588)	−2.4512(−2.3881)
1.9	−2.7129(−2.6147)	−2.6454(−2.5550)	−2.5745(−2.4925)	−2.5006(−2.4274)	−2.4240(−2.3600)
2.0	−2.6764(−2.5777)	−2.6113(−2.5204)	−2.5431(−2.4606)	−2.4720(−2.3983)	−2.3986(−2.3339)
Panel D. 0.5 % level					
1.0	−3.9229(−3.7705)	−3.7941(−3.6512)	−3.6542(−3.5224)	−3.5043(−3.3851)	−3.3459(−3.2405)
1.1	−3.8012(−3.6396)	−3.6789(−3.5277)	−3.5470(−3.4077)	−3.4064(−3.2802)	−3.2583(−3.1464)
1.2	−3.6897(−3.5224)	−3.5736(−3.4173)	−3.4490(−3.3050)	−3.3168(−3.1862)	−3.1780(−3.0618)
1.3	−3.5882(−3.4176)	−3.4778(−3.3185)	−3.3599(−3.2131)	−3.2352(−3.1020)	−3.1047(−2.9859)
1.4	−3.4961(−3.3235)	−3.3910(−3.2299)	−3.2791(−3.1307)	−3.1612(−3.0263)	−3.0380(−2.9175)
1.5	−3.4124(−3.2389)	−3.3121(−3.1502)	−3.2057(−3.0564)	−3.0938(−2.9581)	−2.9773(−2.8557)
1.6	−3.3363(−3.1625)	−3.2403(−3.0782)	−3.1389(−2.9893)	−3.0325(−2.8963)	−2.9219(−2.7997)
1.7	−3.2669(−3.0932)	−3.1749(−3.0129)	−3.0779(−2.9284)	−2.9765(−2.8403)	−2.8712(−2.7488)
1.8	−3.2034(−3.0301)	−3.1150(−2.9534)	−3.0221(−2.8730)	−2.9252(−2.7891)	−2.8248(−2.7023)
1.9	−3.1452(−2.9724)	−3.0601(−2.8991)	−2.9710(−2.8223)	−2.8781(−2.7424)	−2.7820(−2.6597)
2.0	−3.0916(−2.9196)	−3.0097(−2.8492)	−2.9239(−2.7757)	−2.8348(−2.6994)	−2.7426(−2.6206)

Note: 1. κ and λ denote the shape parameter and skewness parameter of SGT distribution, respectively. 2. Numbers in parentheses are quantiles with $n = 20$. 3. Numbers in table are obtained using the composite Simpson's rule with WinRATS 6.1 packages

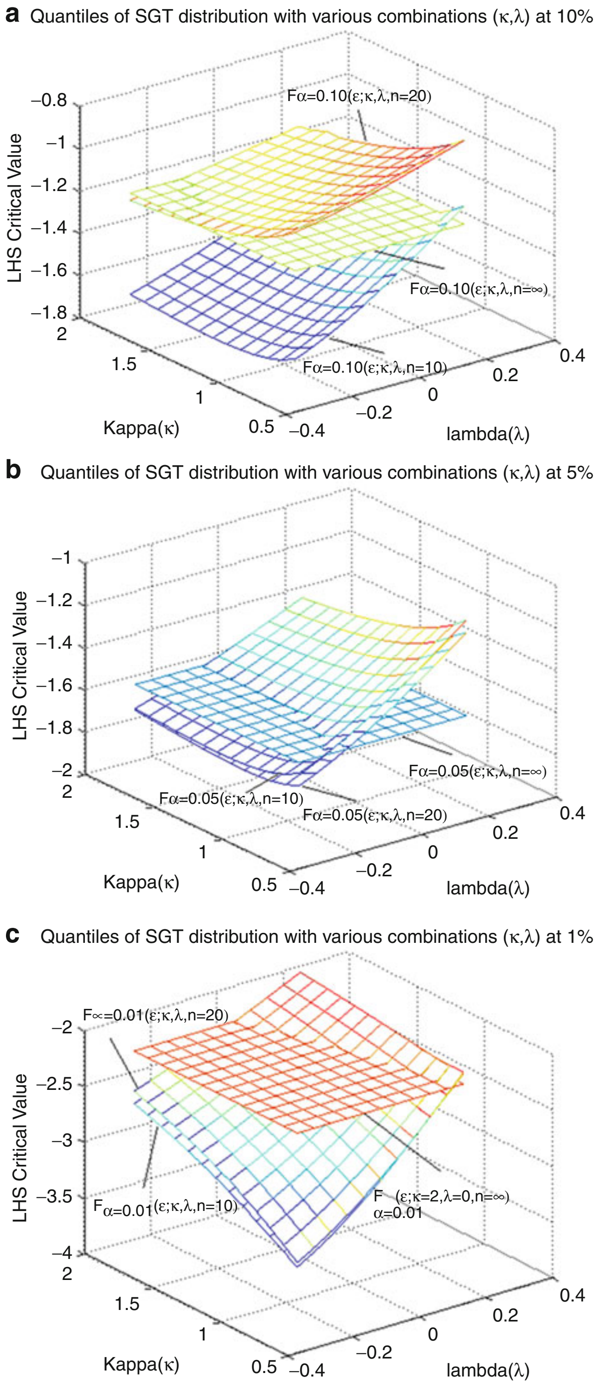


Fig. 51.3 (continued)

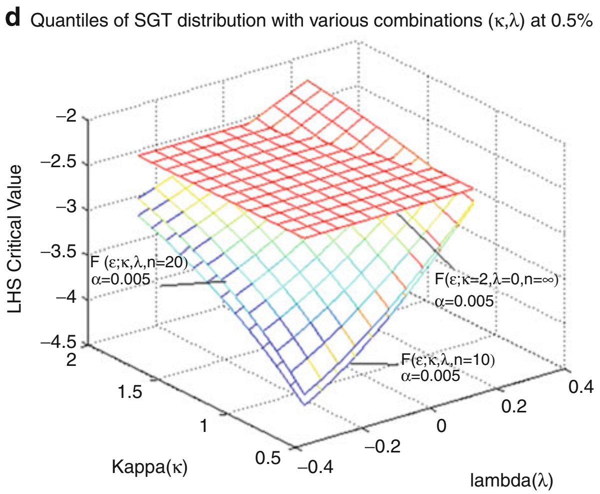


Fig. 51.3 The left-tailed quantiles of SGT distribution with n = 10, 20, and various combinations (κ, λ). (**a**) 10 %, (**b**) 5 %, (**c**) 1 %, (**d**) 0.5 % confidence levels

variance specifications to estimate the VaR. Furthermore, numerous time series data of financial assets appear to exhibit autocorrelated and volatility clustering, and the unconditional distribution of those returns displays leptokurtosis and a moderate amount of skewness. This study thus considers the applicability of the GARCH(1,1) model with three conditional distributions (the normal, student's t, and SGT distributions) to estimate the corresponding volatility in terms of different stock indices, then employs the framework of Jorion (2000) to evaluate the VaR of parametric approach. We take an example of the GARCH-SGT model. The methodology of parametric VaR approach is based on a rolling window procedure. The window size is fixed at 2,000 observations. More specifically, the procedure is conducted in the following manner:

Step 1: For each data series, using the econometric package of WinRATS 6.1, the parameters are estimated with a sample of 2,000 daily returns by quasi-maximum likelihood estimation (QMLE) of log-likelihood function such as Eq. 51.10 and by the BFGS optimization algorithm. Thus, with $\psi = [\mu, \omega, \alpha, \beta, \kappa, \lambda, n]$, the vector of parameters is estimated. The empirical results of GARCH-SGT model are listed in Table 51.2 for all stock indices surveyed in this paper. As to the empirical results of GARCH-N and GARCH-T models, they are also provided by the same approach.

Step 2: Based on the framework of the parametric techniques (Jorion 2000), the 1-day-ahead VaR based on GARCH-SGT model can be calculated by Eq. 51.11. Then the one-step-ahead VaR forecasts are compared with the observed returns, and the comparative results are recorded for subsequent evaluation using statistical tests.

Step 3: The estimation period is then rolled forwards by adding one new day and dropping the most distant day. By replicating step 1 and step 2, the vector of parameters is estimated, and then the 1-day-ahead VaR can be calculated for the next 500 days.

Step 4: For the out-sample period (500 days), via the comparable results between the one-step-ahead VaR forecasts and the observed returns, the 1-day-ahead BLF, QLF, and UL can be calculated by using Eqs. 51.13, 51.14 and 51.16. On the other hand, the unconditional coverage test, LR_{uc}, is evaluated by employing Eq. 51.15. Thereafter, with regard to the GARCH-based models with alternate distributions (GARCH-N, GARCH-T, and GARCH-SGT), the unconditional coverage test (LR_{uc}) and three loss functions (failure rate, AQLF, and UL) are obtained and are reported in the left panel of Tables 51.3, 51.4 and 51.5 for 95 %, 99 %, and 99.5 % levels.

Appendix 3: The Procedure of Semi-parametric VaR Approach

In this paper, we use the approach proposed by Hull and White (1998) as a representative of the semi-parametric approach. This method mainly couples a weighting scheme of volatility with the traditional historical simulation. Hence, it can be regarded as a straightforward extension of traditional historical simulation. The weighting scheme of volatility is expressed as follows. Instead of using the actual historical percentage changes in market variables for the purposes of calculating VaR, we use historical changes that have been adjusted to reflect the ratio of the current daily volatility to the daily volatility at the time of the observation and assume that the variance of each market variable during the period covered by the historical data is monitored using a GARCH-based models. We take an example of the HW-SGT model. This methodology is explained in the following five steps:

Step 1: For each data series, using the econometric package of WinRATS 6.1, the parameters are estimated with a sample of 2,000 daily returns by quasi-maximum likelihood estimation (QMLE) of log-likelihood function such as Eq. 51.10 and by the BFGS optimization algorithm. Thus, with $\psi = [\mu, \omega, \alpha, \beta, \kappa, \lambda, n]$, the vector of parameters is estimated. This step is the same as the first step of parametric approach. Consequently, a series of daily volatility estimates, $\{\sigma_1, \sigma_2, \sigma_3, \ldots\ldots, \sigma_{t=T}\}$, are obtained where T is the number of estimated samples and equals 2,000 in this study.

Step 2: The modified return series are obtained by the raw return series multiplied by the ratio of the current daily volatility to the daily volatility at the time of the observation, σ_T/σ_i. That is, the modified return series are expressed as $\{r_1^*, r_2^*, r_3^*, \ldots\ldots, r_{t=T}^*\}$, where $r_i^* = r_i(\sigma_T/\sigma_i)$.

Step 3: Resort this modified return series ascendingly to achieve the empirical distribution. Thus, VaR is the percentile that corresponds to the specified confidence level. Then the one-step-ahead VaR forecasts are compared with the observed returns, and the comparative results are recorded for subsequent evaluation using statistical tests.

Step 4: The estimation period is then rolled forwards by adding one new day and dropping the most distant day. By replicating steps 1–3, the vector of parameters is estimated, and then the 1-day-ahead VaR can be calculated for the next 500 days. This step is the same as the third step of parametric approach.

Step 5: For the out-sample period (500 days), via the comparable results between the one-step-ahead VaR forecasts and the observed returns, the 1-day-ahead BLF, QLF, and UL can be calculated by using Eqs. 51.13, 51.14, and 51.16. On the other hand, the unconditional coverage test, LR_{uc}, is evaluated by employing Eq. 51.15. Thereafter, with regard to HW-based models with alternate distributions (HW-N, HW-T, and HW-SGT), the unconditional coverage test (LR_{uc}) and three loss functions (failure rate, AQLF, and UL) are obtained and are reported in the right panel of Tables 51.3, 51.4 and 51.5 for 95 %, 99 %, and 99.5 % levels.

References

Aloui, C., & Mabrouk, S. (2010). Value-at-risk estimations of energy commodities via long-memory, asymmetry and fat-tailed GARCH models. *Energy Policy, 38*, 2326–2339.

Angelidis, T., Benos, A., & Degiannakis, S. (2004). The use of GARCH models in VaR estimation. *Statistical Methodology, 1*, 105–128.

Bali, T. G., & Theodossiou, P. (2007). A conditional-SGT-VaR approach with alternative GARCH models. *Annals of Operations Research, 151*, 241–267.

Bhattacharyya, M., Chaudhary, A., & Yadav, G. (2008). Conditional VaR estimation using Pearson's type IV distribution. *European Journal of Operational Research, 191*, 386–397.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics, 31*, 307–327.

Bollerslev, T., Chou, R. Y., & Kroner, K. F. (1992). ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics, 52*, 5–59.

Cabedo, J. D., & Moya, I. (2003). Estimating oil price 'Value at Risk' using the historical simulation approach. *Energy Economics, 25*, 239–253.

Chan, W. H., & Maheu, J. M. (2002). Conditional jump dynamics in stock market returns. *Journal of Business Economics Statistics, 20*, 377–389.

Chang, T. H., Su, H. M., & Chiu, C. L. (2011). Value-at-risk estimation with the optimal dynamic biofuel portfolio. *Energy Economics, 33*, 264–272.

Chen, Q., Gerlach, R., & Lu, Z. (2012). Bayesian value-at-risk and expected shortfall forecasting via the asymmetric Laplace distribution. *Computational Statistics and Data Analysis, 56*, 3498–3516.

Degiannakis, S., Floros, C., & Dent, P. (2012). Forecasting value-at-risk and expected shortfall using fractionally integrated models of conditional volatility: International evidence. *International Review of Financial Analysis, 27*, 21–33.

Faires, J. D., & Burden, R. (2003). *Numerical methods* (3rd ed.). Pacific Grove: Tomson Learning.

Gebizlioglu, O. L., Şenoğlu, B., & Kantar, Y. M. (2011). Comparison of certain value-at-risk estimation methods for the two-parameter Weibull loss distribution. *Journal of Computational and Applied Mathematics, 235*, 3304–3314.

Gençay, R., Selçuk, F., & Ulugülyaˇgci, A. (2003). High volatility, thick tails and extreme value theory in value-at-risk estimation. *Insurance: Mathematics and Economics, 33*, 337–356.

Giot, P., & Laurent, S. (2003a). Value-at-risk for long and short trading positions. *Journal of Applied Econometrics, 18*, 641–664.

Giot, P., & Laurent, S. (2003b). Market risk in commodity markets: A VaR approach. *Energy Economics, 25*, 435–457.

Hartz, C., Mittnik, S., & Paolellad, M. (2006). Accurate value-at-risk forecasting based on the normal-GARCH model. *Computational Statistics & Data Analysis, 51*, 2295–2312.

Huang, Y. C., & Lin, B. J. (2004). Value-at-risk analysis for Taiwan stock index futures: Fat tails and conditional asymmetries in return innovations. *Review of Quantitative Finance and Accounting, 22*, 79–95.

Hull, J., & White, A. (1998). Incorporating volatility updating into the historical simulation method for value-at-risk. *Journal of Risk, 1*, 5–19.

Jarque, C. M., & Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistics Review, 55*, 163–172.

Jorion, P. (2000). *Value at risk: The new benchmark for managing financial risk*. New York: McGraw-Hill.

Kupiec, P. (1995). Techniques for verifying the accuracy of risk management models. *The Journal of Derivatives, 3*, 73–84.

Lee, C. F., & Su, J. B. (2011). Alternative statistical distributions for estimating value-at-risk: Theory and evidence. *Review of Quantitative Finance and Accounting, 39*, 309–331 .

Lee, M. C., Su, J. B., & Liu, H. C. (2008). Value-at-risk in US stock indices with skewed generalized error distribution. *Applied Financial Economics Letters, 4*, 425–431.

Lopez, J. A. (1999). Methods for evaluating value-at-risk estimates. *Federal Reserve Bank of San Francisco Economic Review, 2*, 3–17.

Lu, C., Wu, S. C., & Ho, L. C. (2009). Applying VaR to REITs: A comparison of alternative methods. *Review of Financial Economics, 18*, 97–102.

Sadeghi, M., & Shavvalpour, S. (2006). Energy risk management and value at risk modeling. *Energy Policy, 34*, 3367–3373.

So, M. K. P., & Yu, P. L. H. (2006). Empirical analysis of GARCH models in value at risk estimation. *International Financial Markets, Institutions & Money, 16*, 180–197.

Stavroyiannis, S., Makris, I., Nikolaidis, V., & Zarangas, L. (2012). Econometric modeling and value-at-risk using the Pearson type-IV distribution. *International Review of Financial Analysis, 22*, 10–17.

Su, J. B., & Hung, J. C. (2011). Empirical analysis of jump dynamics, heavy-tails and skewness on value-at-risk estimation. *Economic Modelling, 28*, 1117–1130.

Theodossiou, P. (1998). Financial data and the skewed generalized t distribution. *Management Science, 44*, 1650–1661.

Vlaar, P. J. G. (2000). Value at risk models for Dutch bond portfolios. *Journal of Banking & Finance, 24*, 1131–1154.

Modeling Multiple Asset Returns by a Time-Varying *t* Copula Model

52

Long Kang

Contents

Abstract

We illustrate a framework to model joint distributions of multiple asset returns using a time-varying Student's *t* copula model. We model marginal distributions of individual asset returns by a variant of GARCH models and then use a Student's *t* copula to connect all the margins. To build a time-varying structure for the correlation matrix of *t* copula, we employ a dynamic conditional correlation (DCC) specification. We illustrate the two-stage estimation procedures for the model and apply the model to 45 major US stocks returns selected from nine

L. Kang
Department of Finance, Antai College of Economics and Management, Shanghai Jiao Tong University, Shanghai, China

The Options Clearing Corporation and Center for Applied Economics and Policy Research, Indiana University, Bloomington, IN, USA
e-mail: lkang@indiana.edu; kanglong@gmail.com

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_52,

sectors. As it is quite challenging to find a copula function with very flexible parameter structure to account for difference dependence features among all pairs of random variables, our time-varying *t* copula model tends to be a good working tool to model multiple asset returns for risk management and asset allocation purposes. Our model can capture time-varying conditional correlation and some degree of tail dependence, while it also has limitations of featuring symmetric dependence and inability of generating high tail dependence when being used to model a large number of asset returns.

Keywords

Student's *t* copula • GARCH models • Asset returns • US stocks • Maximum likelihood • Two-stage estimation • Tail dependence • Exceedance correlation • Dynamic conditional correlation • Asymmetric dependence

52.1 Introduction

There have been a large number of applications of copula theory in financial modeling. The popularity of copula mainly results from its capability of decomposing joint distributions of random variables into marginal distributions of individual variables and the copula which links the margins. Then the task of finding a proper joint distribution becomes to find a copula form which features a proper dependence structure given that marginal distributions of individual variables are properly specified. Among many copula functions, Student's *t* copula is a good choice, though not perfect, for modeling multivariate financial data as an alternative to a normal copula, especially for a very large number of assets. The *t* copula models are very useful tools to describe joint distributions of multiple assets for risk management and asset allocation purposes. In this chapter, we illustrate how to model the joint distribution of multiple asset returns under a Copula-GARCH framework. In particular, we show how we can build and estimate a time-varying *t* copula model for a large number of asset returns and how well the time-varying *t* copula accounts for some dependence features of real data.

There are still two challenging issues when applying copula theory to multiple time series. The first is how to choose a copula that best describes the data. Different copulas feature different dependence structure between random variables. Some copulas may fit one particular aspect of the data very well but do not have a very good overall fit, while others may have the opposite performance. What criteria to use when we choose from copula candidates is a major question remaining to be fully addressed. Secondly, how to build a multivariate copula which is sufficiently flexible to simultaneously account for the dependence structure for each pair of random variables in joint distributions is still quite challenging. We hope to shed some light on those issues by working through our time-varying *t* copula model.

Under a Copula-GARCH framework, we first model each asset return with a variant of GARCH specification. Based on different properties of asset returns, we choose a proper GARCH specification to formulate conditional distributions of each return. Then, we choose a proper copula function to link marginal distributions of each return to form the joint distribution. As in marginal distributions of each return, the copula parameters can also be specified as being dependent on previous observations to make the copula structure time varying for a better fit of data. In this chapter, we have an AR(1) process for the conditional mean and a GJR-GARCH (1,1) specification for the conditional volatility for each return. We employ a Student's *t* copula with a time-varying correlation matrix (by a DCC specification) to link marginal distributions. Usually the specified multivariate model contains a huge number of parameters, and the estimation by maximum likelihood estimator (MLE) can be quite challenging. Therefore, we pursue a two-stage procedure, where all the GARCH models for each return are estimated individually first and copula parameters are estimated in the second stage with estimated cumulative distribution functions from the first stage.

We apply our model to modeling log returns of 45 major US stocks selected from nine sectors with a time span ranging from January 3, 2000 to November 29, 2011. Our estimation results show that AR(1) and GJR-GARCH(1,1) can reasonably well capture empirical properties of individual returns. The stock returns possess fat tails and leverage effects. We plot the estimated conditional volatility on selected stocks and volatility spikes which happened during the "Internet Bubbles" in the early 2000s and the financial crisis in 2008. We estimate a DCC specification for the time-varying *t* copula and also a normal copula for comparison purposes. The parameter estimates for time-varying *t* copula are statistically significant, which indicates a significant time-varying property of the dependence structure. The time-varying *t* copula yields significantly higher log-likelihood than normal copula. This improvement of data fitness results from flexibility of *t* copula (relative to normal copula) and its time-varying correlation structure.

We plot the time-varying correlation parameter for selected pairs of stocks under the time-varying *t* copula model. The correlation parameters fluctuate around certain averages, and they spike during the 2008 crisis for some pairs. For 45 asset returns, the estimated degree-of-freedom (DoF) parameter of the *t* copula is around 25. Together with the estimated correlation matrix of the *t* copula, this DoF leads to quite low values of tail dependence coefficients (TDCs). This may indicate the limitation of *t* copulas in capturing possibly large tail dependence behavior for some asset pairs when being used to model a large number of asset returns. Nevertheless, the time-varying Student's *t* copula model has a relatively flexible parameter structure to account for the dependence among multiple asset returns and is a very effective tool to model the dynamics of a large number of asset returns in practice.

This chapter is organized as follows. Section 52.2 gives a short literature review on recent applications of copulas to modeling financial time series. Section 52.3 introduces our copula model where we introduce copula theory, Copula-GARCH framework, and estimation procedures. In particular, we elaborate on how to construct and estimate a time-varying *t* copula model. Section 52.4 documents the data source and descriptive statistics for the data set we use. Section 52.5 reports estimation results and Sect. 52.6 concludes.

52.2 Literature Review

Copula-GARCH models were previously proposed by Jondeau and Rockinger (2002) and Patton (2004, 2006a).[1] To measure time-varying conditional dependence between time series, the former authors use copula functions with time-varying parameters as functions of predetermined variables and model marginal distributions with an autoregressive version of Hansen's (1994) GARCH-type model with time-varying skewness and kurtosis. They show for many market indices, dependency increases after large movements and for some cases it increases after extreme downturns. Patton (2006a) applies the Copula-GARCH model to modeling the conditional dependence between exchange rates. He finds that mark-dollar and yen-dollar exchange rates are more correlated during depreciation against dollar than during appreciation periods. By a similar approach, Patton (2004) models the asymmetric dependence between "large cap" and "small cap" indices and examines the economic and statistical significance of the asymmetries for asset allocations in an out-of-sample setting. As in above literature, copulas are mostly used in capturing asymmetric dependence and tail dependence between times series. Among copula candidates, Gumbel's copula features higher dependence (correlation) at upper side with positive upper tail dependence, and rotated Gumbel's copula features higher dependence (correlation) at lower side with positive lower tail dependence. Hu (2006) studies the dependence structure between a number of pairs of major market indices by a mixed copula approach. Her copula is constructed by a weighted sum of three copulas–normal, Gumbel's, and rotated Gumbel's copulas. Jondeau and Rockinger (2006) model the bivariate dependence between major stock indices by a Student's *t* copula where the parameters are assumed to be modeled by a two-state Markov process.

The task of flexibly modeling dependence structure becomes more challenging for n-dimensional distributions. Tsafack and Garcia (2011) build up a complex multivariate copula to model four international assets (two international equities and two bonds). In his model, he assumes that the copula form has a regime-switching setup where in one regime he uses an n-dimensional normal copula and in the other he uses a mixed copula of which each copula component features the dependence structure of two pairs of variables. Savu and Trede (2010) develop a hierarchical Archimedean copula which renders more flexible parameters to characterize dependency between each pair of variables. In their model, each pair of closely related random variables is modeled by a copula of a particular Archimedean class, and then these pairs are nested by copulas as well. The nice property of Archimedean family easily leads to the validity of the

[1]Alternative approaches are also developed, such as in Ang and Bekaert (2002), Goeij and Marquering (2004), and Lee and Long (2009), to address non-normal joint distributions of asset returns.

joint distribution constructed by this hierarchical structure. (Trivedi and Zimmer 2006) apply trivariate hierarchical Archimedean copulas to model sample selection and treatment effects with applications to the family health-care demand.

Statistical goodness-of-fit tests can provide some guidance for selecting copula models. Chen et al. (2004) propose two simple goodness-of-fit tests for multivariate copula models, both of which are based on multivariate probability integral transform and kernel density estimation. One test is consistent but requires the estimation of the multivariate density function and hence is suitable for a small number of random variables, while the other may not be consistent but requires only kernel estimation of a univariate density function and hence is suitable for a large number of assets. Berg and Bakken (2006) propose a consistent goodness-of-fit test for copulas based on the probability integral transform, and they incorporate in their test a weighting functionality which can increase influence of some specific areas of copulas.

Due to their parameter structure, the estimation of Copula-GARCH models also suffers from "the curse of dimensionality".[2] The exact maximum likelihood estimator (MLE) works in theory.[3] In practice, however, as the number of time series being modeled increases, the numerical optimization problem in MLE will become formidable. Joe and Xu (1996) propose a two-stage procedure, where in the first stage only parameters in marginal distributions are estimated by MLE and then the copula parameters are estimated by MLE in the second stage. This two-stage method is called inference for the margins (IFM) method. Joe (1997) shows that under regular conditions the IFM estimator is consistent and has the property of asymptotic normality and Patton (2006b) also shows similar estimator properties for the two-stage method. Instead of estimating parametric marginal distributions in the IFM method, we can estimate the margins by using empirical distributions, which can avoid the problem of mis-specifying marginal distributions. This method is called canonical maximum likelihood (CML) method by Cherubini et al. (2004). Hu (2006) uses this method and she names it as a semi-parametric method. Based on Genest et al. (1995), she shows that CML estimator is consistent and has asymptotical normality. Moreover, copula models can also be estimated under a nonparametric framework. Deheuvels (1981) introduces the notion of empirical copula and shows that the empirical copula converges uniformly to the underlying true copula. Finally, Xu (2004) shows how the copula models can be estimated with a Bayesian approach. The author shows how a Bayesian approach can be used to account for estimation uncertainty in portfolio optimization based on a Copula-GARCH model, and she proposes to use a Bayesian MCMC algorithm to jointly estimate the copula models.

[2]For a detailed survey on the estimation of Copula-GARCH model, see Chap. 5 of Cherubini et al. (2004).

[3]See Hamilton (1994) and Greene (2003) for more details on maximum likelihood estimation.

52.3 The Model

52.3.1 Copula

We introduce our Copula-GARCH model framework by first introducing the concept of copula. A copula is a multivariate distribution function with uniform marginal distributions as its arguments, and its functional form links all the margins to form a joint distribution of multiple random variables.[4] Copula theory is mainly based on the work of Sklar (1959), and we state the Sklar's theorem for continuous marginal distributions as follows.

Theorem 52.1 *Let $F_1(x_1), \ldots, F_n(x_n)$ be given marginal distribution functions and continuous in $x_1, \ldots, x_n$, respectively. Let H be the joint distribution of $(x_1, \ldots, x_n)$. Then there exists a unique copula C such that*

$$H(x_1, \ldots, x_n) = C(F_1(x_1), \ldots, F_n(x_n)), \quad \forall (x_1, \ldots, x_n) \in \overline{\mathbb{R}}^n. \tag{52.1}$$

Conversely, if we let $F_1(x_1), \ldots, F_n(x_n)$ be continuous marginal distribution functions and C be a copula, then the function H defined by Eq. 52.1 is a joint distribution function with marginal distributions $F_1(x_1), \ldots, F_n(x_n)$.

The above theory allows us to decompose a multivariate distribution function into marginal distributions of each random variable and the copula form linking the margins. Conversely, it also implies that to construct a multivariate distribution, we can first find a proper marginal distribution for each random variable and then obtain a proper copula form to link the margins. Depending on which dependence measure used, the copula function mainly, not exclusively, governs the dependence structure between individual variables. Hence, after specifying marginal distributions of each variable, the task of building a multivariate distribution solely becomes to choose a proper copula form which best describes the dependence structure between variables.

Differentiating Eq. 52.1 with respect to $(x_1, \ldots, x_n)$ leads to the joint density function of random variables in terms of copula density. It is given as

$$h(x_1, \ldots, x_n) = c((F_1(x_1), \ldots, F_n(x_n))\prod_{i=1}^{n} f_i(x_i), \quad \forall (x_1, \ldots, x_n) \in \overline{\mathbb{R}}^n, \tag{52.2}$$

where $c(F_1(x_1), \ldots, F_n(x_n))$ is the copula density and $f_i(x_i)$ is the density function for variable i. Equation 52.2 implies that the log-likelihood of the joint density can be decomposed into components which only involve each marginal density and a component which involves copula parameters. It provides a convenient structure for a two-stage estimation, which will be illustrated in details in the following sections.

[4] See Nelsen (1998) and Joe (1997) for a formal treatment of copula theory, and Bouye et al. (2000), Cherubini et al. (2004), and Embrechts et al. (2002) for applications of copula theory in finance.

To better fit the data, we usually assume the moments of distributions of random variables are time varying and depend on past variables. Therefore, the distribution of random variables at time t becomes a conditional one, and then the above copula theory needs to be extended to a conditional case. It is given as follows.[5]

Theorem 52.2 *Let Ω_{t-1} be the information set up to time t, and let $F_1(x_{1,t}|\Omega_{t-1})$, …, $F_n(x_{n,t}|\Omega_{t-1})$ be continuous marginal distribution functions conditional on Ω_{t-1}. Let H be the joint distribution of $(x_1, \ldots, x_n)$ conditional on Ω_{t-1}. Then there exists a unique copula C such that*

$$H\left(x_1, \ldots, x_n|\Omega_{t-1}\right) = C\left(F_1\left(x_1|\Omega_{t-1}\right), \ldots, F_n\left(x_n|\Omega_{t-1}\right)|\Omega_{t-1}\right), \ \ \forall(x_1, \ldots, x_n) \in \overline{\mathbb{R}}^n. \tag{52.3}$$

Conversely, if we let $F_1(x_{1,t}|\Omega_{t-1})$, …, $F_n(x_{n,t}|\Omega_{t-1})$ be continuous conditional marginal distribution functions and C be a copula, then the function H defined by Eq. 52.3 is a conditional joint distribution function with conditional marginal distributions $F_1(x_{1,t}|\Omega_{t-1})$, …, $F_n(x_{n,t}|\Omega_{t-1})$.

It is worth noting that for the above theorem to hold, the information set Ω_{t-1} has to be the same for the copulas and all the marginal distributions. If different information sets are used, the conditional copula form on the right side of Eq. 52.3 may not be a valid distribution. Generally, the same information set used may not be relevant for each marginal distributions and the copula. For example, the marginal distributions or the copula may be only conditional on a subset of the universally used information set. At the very beginning of estimation of the conditional distributions, however, we should use the same information set based on which we can test for insignificant explanatory variables so as to stick to a relevant subset for each marginal distribution or the copula.

52.3.2 Modeling Marginal Distributions

Before building a copula model, we need to find a proper specification for marginal distributions of individual asset returns, as mis-specified marginal distributions automatically lead to a mis-specified joint distribution. Let $x_{i,t}$ be asset i return at time t, and its conditional mean and variance are modeled as follows:

$$x_{i,t} = \alpha_{0,i} + \alpha_{1,i}x_{i,t-1} + \varepsilon_{i,t}, \tag{52.4}$$

$$\varepsilon_{i,t} = \sqrt{h_{i,t}}\eta_{i,t}, \tag{52.5}$$

$$h_{i,t} = \beta_{0,i} + \beta_{1,i}h_{i,t-1} + \beta_{2,i}\varepsilon^2_{i,t-1} + \beta_{3,i}\varepsilon^2_{i,t-1}1\left(\varepsilon_{i,t-1} < 0\right). \tag{52.6}$$

[5]See Patton (2004).

As shown in Eqs. 52.4, 52.5 and 52.6, we model the conditional mean as an AR(1) process and the conditional variance as a GJR(1,1) specification.[6] We have parameter restrictions as $\beta_{0,i} > 0$, $\beta_{1,i} \geq 0$, $\beta_{2,i} \geq 0$, $\beta_{2,i} + \beta_{3,i} \geq 0$, and $\beta_{1,i} + \beta_{2,i} + \frac{1}{2}\beta_{3,i} < 1$. $1(\varepsilon_{i,t-1} < 0)$ is an indicator function, which equals one when $\varepsilon_{i,t-1} < 0$ and zero otherwise. We believe that our model specifications can capture the features of the individual stock returns reasonably well. It is worth noting that Eqs. 52.4, 52.5 and 52.6 can include more exogenous variables to better describe the data. Alternative GARCH specifications can be used to describe the time-varying conditional volatility. We assume $\eta_{i,t}$ is *i.i.d.* across time and follows a Student's *t* distribution with DoF v_i.

Alternatively, to model the conditional higher moments of the series, we can follow Hansen (1994) and Jondeau and Rockinger (2003) who assume a skewed *t* distribution for the innovation terms of GARCH specifications and find that the skewed *t* distribution fits financial time series better than normal distribution. Accordingly, we can assume $\eta_{i,t} \sim Skewed\,T(\eta_{i,t}|v_{i,t},\lambda_{i,t})$ with zero mean and unitary variance where $v_{i,t}$ is DoF parameter and $\lambda_{i,t}$ is skewness parameter. The two parameters are time varying and depend on lagged values of explanatory variables in a nonlinear form. For illustration purposes, however, we will only use Student's *t* distribution for $\eta_{i,t}$ in this chapter.

52.3.3 Modeling Dependence Structure

Normal copula and Student's *t* copula are two copula functions from elliptical families, which are frequently used in modeling joint distributions of random variables. In this chapter, we also estimate a normal copula model for comparison purposes. Let Φ^{-1} denote the inverse of the standard normal distribution Φ and $\Phi_{\sum,n}$ be *n*-dimensional normal distribution with correlation matrix $\sum$. Hence, the *n*-dimensional normal copula is

$$C(\mathbf{u};\Sigma) = \Phi_{\Sigma,N}\left(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n)\right), \tag{52.7}$$

and its density form is

$$c(\mathbf{u};\Sigma) = \frac{\phi_{\Sigma,n}\left(\Phi^{-1}(u_1), \ldots, \Phi^{-1}(u_n)\right)}{\prod_{i=1}^{n}\phi\left(\Phi^{-1}(u_i)\right)}, \tag{52.8}$$

where ϕ and $\phi_{\sum,n}$ are the probability density functions (*pdf*s) of Φ and $\Phi_{\sum,n}$, respectively. It can be shown via Sklar's theorem that normal copula generates standard joint normal distribution if and only if the margins are standard normal.

[6]See Glosten et al. (1993).

On the other hand, let T_v^{-1} be the inverse of standard Student's t distribution T_v with DoF parameter[7] $v > 2$ and $T_{R,v}$ be n-dimensional Student's t distribution with correlation matrix R and DoF parameter v. Then n-dimensional Student's t copula is

$$C(\mathbf{u};R,v) = T_{R,v}\left(T_v^{-1}(u_1),\ \ldots, T_v^{-1}(u_n)\right), \tag{52.9}$$

and its density function is

$$c(u_1,\ldots,u_n) = \frac{t_{R,v}\left(T_v^{-1}(u_1),\ \ldots, T_v^{-1}(u_n)\right)}{\prod_{i=1}^{n} t_v\left(T_v^{-1}(u_i)\right)},$$

where t_v and $t_{R,v}$ are the *pdf*s of T_v and $T_{R,v}$, respectively.

Borrowing from the dynamic conditional correlation (DCC) structure of multivariate GARCH models, we can specify a time-varying parameter structure in the t copula as follows.[8] For a t copula, the time-varying correlation matrix is governed by

$$Q_t = (1-\alpha-\beta)S + \alpha\left(\varsigma_{t-1}\varsigma_{t-1}'\right) + \beta Q_{t-1}, \tag{52.10}$$

where S is the unconditional covariance matrix of $\varsigma_t = (T_v^{-1}(u_{1,t}), \ldots, T_v^{-1}(u_{n,t}))'$ and α and β are nonnegative and satisfy the condition $\alpha + \beta < 1$. We assign $Q_0 = S$ and the dynamics of Q_t is given by Eq. 52.10. Let $q_{i,j,t}$ be the i,j element of the matrix Q_t, and the i,j element of the conditional correlation matrix R_t can be calculated as

$$\rho_{i,j,t} = \frac{q_{i,j,t}}{\sqrt{q_{i,i,t}q_{j,j,t}}}. \tag{52.11}$$

Moreover, the specification of Eq. 52.10 guarantees that the conditional correlation matrix R_t is positive definite.

Proposition 52.1 *In Eqs.* 52.10 and 52.11, *if*

(a) $\alpha \geq 0$ and $\beta \geq 0$,
(b) $\alpha + \beta < 1$,
(c) *All eigenvalues of S are strictly positive, then the correlation matrix R_t is positive definite.*

[7]In contrast to the previous standardized Student's t distribution, the standard Student's t distribution here has variance as $v/(v-2)$.

[8]Please see Engle and Sheppard (2001) and Engle (2002) for details on the multivariate DCC-GARCH models.

Proof *First, (a) and (b) guarantee the system for $\varsigma_t\varsigma_t'$ is stationary and S exists. With $Q_0 = S$, (c) guarantees Q_0 is positive definite. With (a) to (c), Q_t is the sum of a positive definite matrix, a positive semi-definite matrix, and a positive definite matrix both with nonnegative coefficients and then is positive definite for all t. Based on the proposition Eq.* 52.1 *in Engle and Sheppard* (2001), *we prove that R_t is positive definite.*

52.3.4 Estimation

We illustrate the estimation procedure by writing out the log-likelihoods for observations. Let $\Theta = \{\theta, \gamma_1, \ldots, \gamma_n\}$ be the set of parameters in the joint distribution where θ is the set of parameters in the copula and γ_t is the set of parameters in marginal distributions for asset *i*. Then the conditional cumulative distribution function (*cdf*) of *n* asset returns at time *t* is given as

$$F\left(x_{1,t}, \ldots, x_{n,t} \middle| \underline{X}_{t-1}, \Theta\right) = C\left(u_{1,t}, \ldots, u_{n,t} \middle| \underline{X}_{t-1}, \theta\right) \tag{52.12}$$

where $\underline{X}_{t-1}$ is a vector of previous observations, $C(\cdot|\underline{X}_{t-1}, \theta)$ is the conditional copula, and $u_{i,t} = F_i(x_{i,t}|X_{t-1}, \gamma_i)$ is the conditional *cdf* of the margins. Differentiating both sides with respect to $x_{1t}, \ldots, x_{n,t}$ leads to the density function as

$$f\left(x_{1,t}, \ldots, x_{n,t} \middle| \underline{X}_{t-1}, \Theta\right) = c\left(u_{1,t}, \ldots, u_{n,t} \middle| \underline{X}_{t-1}, \theta\right) \prod_{i=1}^{n} f_i\left(x_{i,t} \middle| \underline{X}_{t-1}, \gamma_i\right), \tag{52.13}$$

where $c(\cdot|\underline{X}_{t-1}, \theta)$ is the density of the conditional copula and $f_i(x_{i,t}|\underline{X}_{t-1}, \gamma_i)$ is the conditional density of the margins. Accordingly, the log-likelihood of the sample is given by

$$L(\Theta) = \sum_{t=1}^{T} \log f\left(x_{1,t}, \ldots, x_{n,t} \middle| \underline{X}_{t-1}, \Theta\right). \tag{52.14}$$

With Eq. 52.13, the log-likelihood can be written as

$$L(\theta, \gamma_1, \ldots, \gamma_n) = \sum_{t=1}^{T} \log c\left(u_{1,t}, \ldots, u_{n,t} \middle| \underline{X}_{t-1}, \theta\right) + \sum_{t=1}^{T} \sum_{i=1}^{n} f_i\left(x_{i,t} \middle| \underline{X}_{t-1}, \gamma_i\right). \tag{52.15}$$

From Eq. 52.15, we observe that the copula and marginal distributions are additively separate. Therefore, we can estimate the model by a two-stage MLE procedure. In the first stage, the marginal distribution parameters for each asset are estimated by MLE, and then with estimated *cdf* of each asset, we estimate the copula parameters by MLE. Based on Joe (1997) and Patton (2006b), this two-stage estimator is consistent and asymptotically normal.

With our model specifications, we first estimate the univariate GJR-GARCH (1,1) with an AR(1) conditional mean and Student's t distribution by MLE. In the second stage, we need to estimate the parameters for the constant normal copula and the time-varying Student's t copula. Let $\mathbf{x}_t = (\Phi^{-1}(u_{1,t}), \ldots, \Phi^{-1}(u_{n,t}))'$, and we can analytically derive the correlation matrix estimator $\hat{\Sigma}$ which maximizes the log-likelihood of the normal copula density as

$$\hat{\Sigma} = \frac{1}{T}\sum_{t=1}^{T} \mathbf{x}_t \mathbf{x}_t'. \tag{52.16}$$

As there is no analytical solution for MLE of Student's t copula, the numerical maximization problem is quite challenging. Following Chen et al. (2004), however, with $\varsigma_t = (T_v^{-1}(u_{1,t}), \ldots, T_v^{-1}(u_{n,t}))'$, we can calculate the sample covariance matrix of ς_t as $\hat{S}$, which is a function of DoF parameter v. By setting $Q_0 = \hat{S}$, we can express Q_t and R_t for all t in terms of α, β, and v using Eq. 52.10. Then we can estimate α, β, and v by maximizing the log-likelihood of t copula density. In the following sections, we apply our estimation procedure to the joint distribution of 45 selected major US stock returns.

52.4 Data

We apply our model to modeling log returns of 45 major US stocks from nine sectors: Consumer Discretionary, Consumer Staples, Energy, Financials, Health Care, Industrials, Technology, Materials, and Utilities. Table 52.1 shows stock symbols and company names of the selected 45 companies. We select five major companies from each sector to form the stock group. The time span ranges from January 3, 2000 to November 29, 2011 with 2990 observations. We download data from yahoo finance (http://finance.yahoo.com/). The log returns are calculated from daily close stock prices adjusted for dividends and splits.

To save space, we only plot and calculate descriptive statistics of nine stocks with each from one sector. Figure 52.1 plots the log returns of those nine selected stocks, and there are two periods of volatility clusterings due to "Internet Bubbles" in the early 2000s and the financial crisis in 2008, respectively. We observe that during the financial crisis in 2008, major banks, such as Citigroup, incurred huge negative and positive daily returns. Table 52.2 shows the calculated mean, standard deviation, skewness, and kurtosis for the nine stocks. The average returns for the nine stocks are close to zero. Major banks, represented by Citigroup, have significantly higher volatility. Most of the stocks are slightly positively skewed, and only two have slight negative skewness. All the stocks have kurtosis greater than three indicating fat tails, and again major banks have significantly fatter tails. All the descriptive statistics indicate that the data property of individual returns needs to be captured by a variant of GARCH specification.

Table 52.1 Symbols and names of 45 selected stocks from nine sectors

Sector	Consumer discretionary	Consumer Staples	Energy
Stock symbol	MCD: McDonald's	WMT: Wal-Mart Stores Inc.	XOM: Exxon Mobil Corp.
	HD: Home Depot	PG: Procter & Gamble Co.	CVX: Chevron Corp.
	DIS: Walt Disney Co.	KO: Coca-Cola Co.	COP: CONOCOPHILLIPS.
	TGT: Target	WAG: Walgreen Co.	DVN: Devon Energy Corp.
	LOW: Lowe's	MO: Altria Group Inc.	SLB: Schlumberger Limited
Sector	**Financials**	**Health Care**	**Industrials**
Stock symbol	C: Citigroup Inc.	JNJ: Johnson & Johnson	GE: General Electric Co.
	BAC: Bank of America Corp.	PFE: Pfizer Inc.	UNP: Union Pacific Corp.
	JPM: JPMorgan Chase & Co.	ABT: Abbott Laboratories	UTX: United Technologies Corp.
	USB: U.S. Bancorp	MRK: Merck & Co. Inc.	MMM: 3 M Co.
	WFC: Wells Fargo & Co.	AMGN: Amgen Inc.	BA: Boeing Co.
Sector	**Technology**	**Materials**	**Utilities**
Stock symbol	T: AT&T Inc.	NEM: Newmont Mining Corp.	EXC: Exelon Corp.
	MSFT: Microsoft Corp.	DD: E.I. DuPont de Nemours & Co.	FE: FirstEnergy Corp.
	IBM: International Business Machines Corp.	DOW: Dow Chemical Co.	PPL: PPL Corporation
	CSCO: Cisco Systems Inc.	FCX: Freeport-McMoRan Copper & Gold Inc.	D: Dominion Resources, Inc.
	HPQ: Hewlett-Packard Co.	PX: Praxair Inc.	DUK: Duke Energy Corp.

52.5 Empirical Results

52.5.1 Marginal Distributions

We briefly report estimation results for marginal distributions of 45 stock returns. For convenience, we only show the estimates and standard errors (in brackets) for nine selected stocks with each from one sector in Table 52.3. The star indicates statistical significance at a 5 % level. Consistent with our observations in Table 52.2, all the nine stocks have low values of DoF indicating fat tails. The parameter $\beta_{3,i}$ is

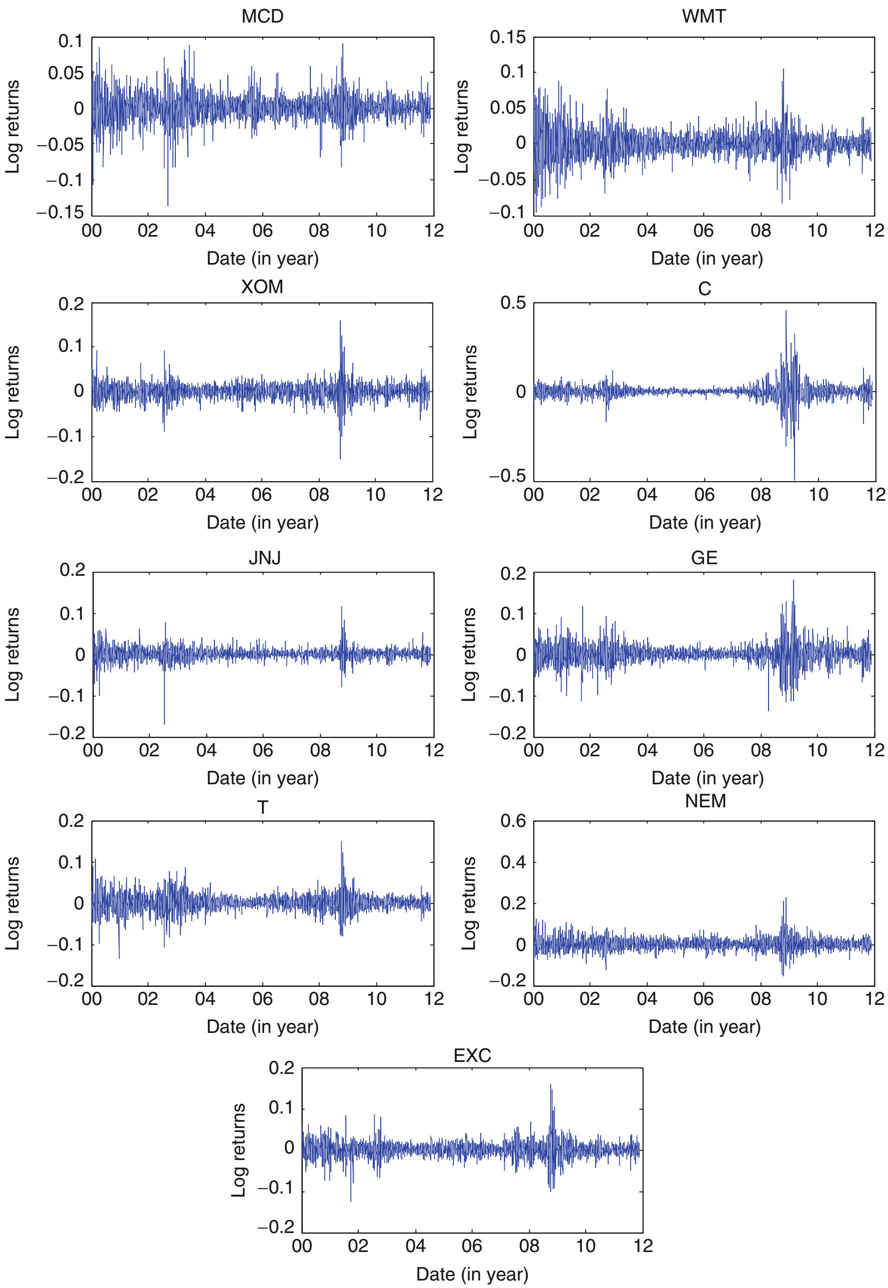


Fig. 52.1 The log returns of the nine of our 45 selected stocks with each from one sector have been plotted

Table 52.2 Descriptive statistics (mean, standard deviation, skewness, and kurtosis) for the nine of our 45 selected stocks with each from each sector

Stock symbol	MCD	WMT	XOM	C	JNJ	GE	T	NEM	EXC
Mean	3.73E-04	7.21E-06	3.15E-04	−8.05E-04	1.96E-04	−2.89E-04	1.23E-05	3.65E-04	4.48E-04
Std. dev.	0.017	0.017	0.017	0.037	0.013	0.022	0.019	0.027	0.018
Skewness	−0.21	0.13	0.02	−0.48	−0.53	0.04	0.12	0.34	0.05
Kurtosis	8.25	7.72	12.52	35.55	17.83	9.99	8.68	8.22	10.58
# obs.	2,990	2,990	2,990	2,990	2,990	2,990	2,990	2,990	2,990

statistically significant at a 5 % level for eight of the nine stocks indicating significant leverage effects for stock returns. The parameters in conditional mean are statistically significant for some stocks and not for others. In Fig. 52.2, we plot estimated conditional volatility for the stocks MCD, WMT, XOM, and C. Consistent with Fig. 52.1, we observe MCD and WMT have significant high volatility in the early 2000s and 2008, while XOM and C have their volatility hikes mainly in 2008 with C, representing Citigroup, having the highest conditional volatility during the 2008 crisis.

52.5.2 Copulas

We report estimation results for the time-varying *t* copula parameters in Table 52.4. All the three parameters α, β, and v are statistically significant. The estimate α is close to zero and the estimate for β is close to one. The estimate for v is about 25. As our estimation is carried out on the joint distribution of 45 stock returns, the estimate for v shed some light on how much Student's *t* copula can capture tail dependence when used to fit a relatively large number of variables. We also report the log-likelihood for time-varying Student's *t* copula and normal copula in Table 52.4. As the correlation matrix in normal copula is estimated by its sample correlation, we did not report it here. We find that time-varying *t* copula has significantly higher log-likelihood than normal copula, which results from the more flexible parameter structure of *t* copula and the time-varying parameter structure.

52.5.3 Time-Varying Dependence

Our time-varying *t* copula features a time-varying dependence structure among all the variables. The DoF parameter, together with the correlation parameters, governs the tail dependence behavior of multiple variables. We plot the estimated

Table 52.3 The GARCH estimation results of individual stock returns for the nine of our selected 45 stocks with each from one sector. Values in brackets are standard errors. The star indicates the statistical significance at a 5 % level

Stock symbol	MCD	WMT	XOM	C	JNJ	GE	T	NEM	EXC
Conditional mean									
$\alpha_{0,i}$	6.69E-04* (2.25E-04)	−1.77E-05 (2.11E-04)	6.50E-04* (2.41E-04)	−1.03E-06 (2.42E-04)	1.17E-04 (1.59E-04)	−1.93E-06 (2.32E-04)	2.73E-04 (2.29E-04)	3.31E-04 (3.87E-04)	7.10E-04* (2.36E-04)
$\alpha_{1,i}$	−0.030 (0.018)	−0.037* (0.018)	−0.076* (0.019)	−0.003 (0.018)	0.004 (0.018)	−0.015 (0.018)	0.009 (0.018)	−0.056* (0.018)	−0.019 (0.019)
Conditional variance									
$\beta_{0,i}$	1.69E-06* (5.21E-07)	8.22E-07* (3.50E-07)	5.97E-06* (1.28E-06)	1.96E-06* (5.53E-07)	1.87E-06* (4.40E-07)	1.26E-06* (3.93E-07)	1.37E-06* (4.47E-07)	3.11E-06* (1.31E-06)	3.91E-06* (9.82E-07)
$\beta_{1,i}$	0.947* (0.007)	0.952* (0.007)	0.901* (0.013)	0.908* (0.009)	0.905* (0.011)	0.948* (0.007)	0.943* (0.007)	0.959* (0.007)	0.897* (0.012)
$\beta_{2,i}$	0.024* (0.008)	0.026* (0.008)	0.025* (0.012)	0.045* (0.011)	0.025* (0.011)	0.017* (0.007)	0.027* (0.008)	0.037* (0.009)	0.066* (0.014)
$\beta_{3,i}$	0.046* (0.012)	0.038* (0.013)	0.092* (0.018)	0.094* (0.017)	0.127* (0.021)	0.067* (0.012)	0.051* (0.012)	−0.001 (0.011)	0.046* (0.018)
Degree of freedom									
v_i	6.187* (0.67)	7.002* (0.92)	9.496* (1.59)	6.505* (0.75)	5.796* (0.62)	6.622* (0.66)	8.403* (1.20)	7.831* (1.09)	9.247* (1.28)

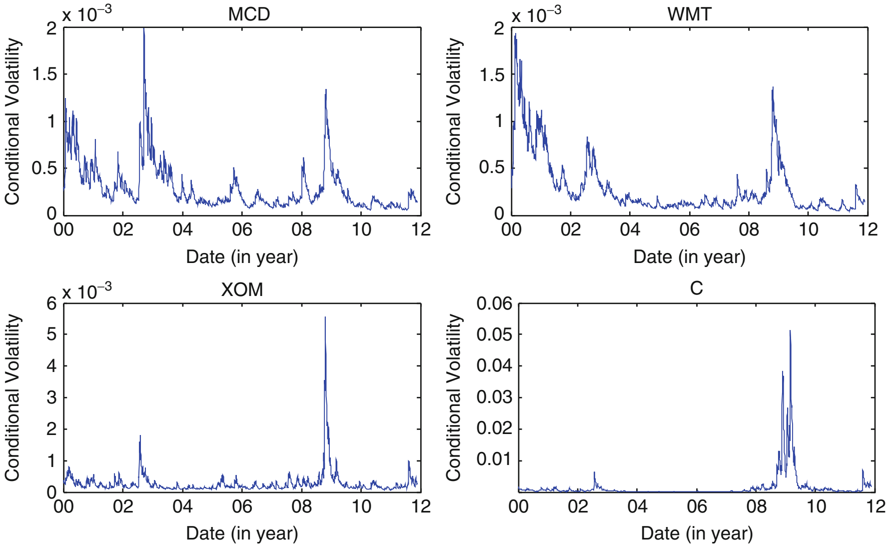


Fig. 52.2 The estimated time-varying conditional volatility for four selected stocks has been plotted

Table 52.4 The estimates and standard errors for time-varying Student's t copula. Values in brackets are standard errors. The star indicates the statistical significance at a 5 % level. We also report the log-likelihood for time-varying t copula and normal copula

	Time-varying t copula	Normal copula
Parameter estimates		
α	0.0031* (0.0002)	
β	0.984* (0.0016)	
v	25.81* (1.03)	
Log-likelihood of copula component		
Log-likelihood	40,445.44	38,096.36

conditional correlation parameters of t copula for four selected pairs of stock returns in Fig. 52.3. For those four pairs, the conditional correlation parameter fluctuates around certain positive averages. The two pairs, MCD-WMT and NEM-EXC, experienced apparent correlation spikes during the 2008 financial crisis. Moreover, Fig. 52.4 shows the estimated TDCs for the four pairs. We find that with the DoF around 25, the TDCs for those pairs of stock returns are very low, though some pairs do exhibit TDC spikes during the 2008 crisis. The low values of TDCs indicate possible limitations of t copula to account for tail dependence when being used to model a large number of variables.

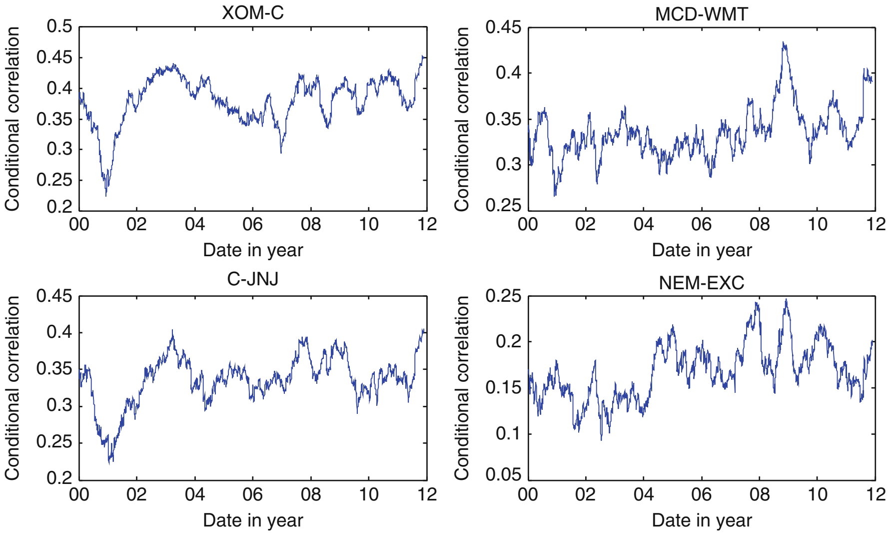


Fig. 52.3 The estimated time-varying correlation parameters in t copula for four selected pairs of stock returns have been plotted

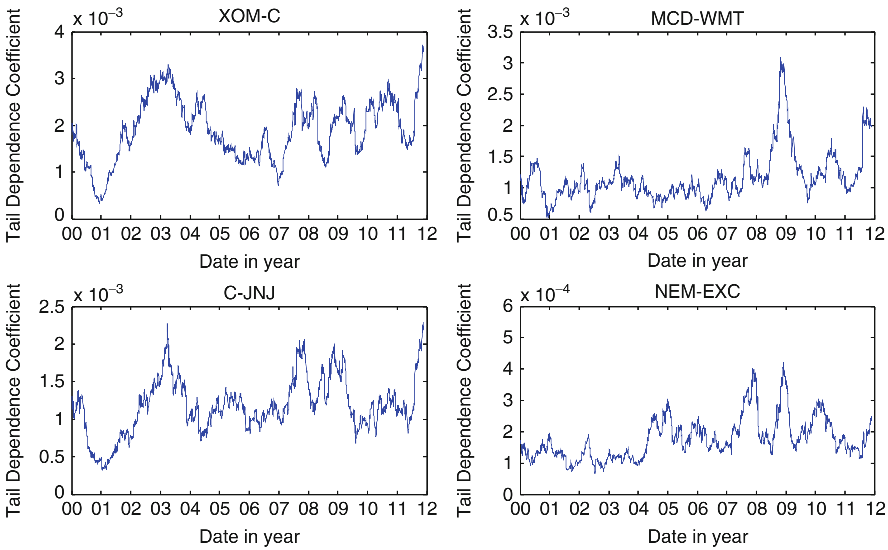


Fig. 52.4 The time-varying tail dependence coefficient (*TDC*) for the four selected pairs of stock returns has been plotted

52.6 Conclusion

We illustrate an effective approach (Copula-GARCH models) to model the dynamics of a large number of multiple asset returns by constructing a time-varying Student's *t* copula model. Under a general Copula-GARCH framework, we specify a proper GARCH model for individual asset returns and use a copula to link the margins to build the joint distribution of returns. We apply our time-varying Student's *t* copula model to 45 major US stock returns, where each stock return is modeled by an AR(1) and GJR-GARCH(1,1) specification and a Student's *t* copula with a DCC dependence structure is used to link all the returns. We illustrate how the model can be effectively estimated by a two-stage MLE procedure, and our estimation results show time-varying *t* copula model has significant better fitness of data than normal copula models.

As it is quite challenging to find a copula function with very flexible parameter structure to account for difference dependence features among all pairs of random variables, our time-varying *t* copula model tends to be a good working tool to model multiple asset returns for risk management and asset allocation purposes. Our model can capture time-varying conditional correlation and some degree of tail dependence, while it also has limitations of featuring symmetric dependence and inability of generating high tail dependence when being used to model a large number of asset returns. Nevertheless, we hope that this chapter provides researchers and financial practitioners with a good introduction on the Copula-GARCH models and a detailed illustration on constructing joint distributions of multiple asset returns using a time-varying Student's *t* copula model.

References

Ang, A., & Bekaert, G. (2002). International asset allocation with regime shifts. *Review of Financial Studies, 15*(4), 1137–1187.

Berg, D., & Bakken, H. (2006). *A copula goodness-of-fit test based on the probability integral transform*. Working Paper at http://www.danielberg.no/publications/Btest.pdf

Bouyé, E., Valdo D., Ashkan N., Gaël R., & Thierry, R. (2000). *Copulas for finance – A reading guide and some applications*. Available at SSRN http://ssrn.com/abstract=1032533. March.

Chen, X., Yanqin F., & Andrew, J. P. (2004). *Simple tests for models of dependence between multiple financial time series, with applications to U.S. Equity returns and exchange rates* (Discussion Paper 483). London School of Economics, Revised July 2004.

Cherubini, U., Luciano, E., & Vecchiato, W. (2004). *Copula methods in finance* (Wiley series in financial engineering). Chichester: Wiley.

De Goeij, P., & Marquering, W. (2004). Modeling the conditional covariance between stock and bond returns: A multivariate GARCH approach. *Journal of Financial Econometrics, 2*(4), 531–564.

Deheuvels, P. (1981). *A Nonparametric test for independence*. Paris: Universite de Paris, Institut de Statistique.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business & Economic Statistics, 20*(3), 339–350.

Engle, R. F., & Sheppard, K. (2001). *Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH* (NBER Working Paper No. 8554). Cambridge, MA: National Bureau of Economic Research.

Genest, C., Ghoudi, K., & Rivest, L. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika, 82*, 543–552.

Glosten, L. R., Ravi, J., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess return on stocks. *Journal of Finance, 48*(5), 1779–1801.

Greene, W. H. (2003). *Econometric analysis*. Upper Saddle River: Prentice Hall.

Hamilton, J. D. (1994). *Time series analysis*. Princeton: Princeton University Press.

Hansen, B. E. (1994). Autoregressive conditional density estimation. *International Economic Review, 35*(3), 705–730.

Hu, L. (2006). Dependence patterns across financial markets: A mixed copula approach. *Applied Financial Economics, 16*, 717–729.

Joe, H. (1997). *Multivariate models and dependence concepts*. London: Chapman and Hall.

Joe, H., & Xu, J. J. (1996). *The estimation method of inference functions for margins for multivariate models* (Technical Report 166). Vancouver: Department of Statistics, University of British Columbia.

Jondeau, E., & Rockinger, M. (2002). *Conditional dependency of financial series: the copula-GARCH model* (FAME Research Paper 69). Geneva: International Center for Financial Asset Management and Engineering.

Jondeau, E., & Rockinger, M. (2003). Conditional volatility, skewness, and kurtosis: Existence, persistence, and comovements. *Journal of Economic Dynamics and Control, 27*(10), 1699–1737.

Jondeau, E., & Rockinger, M. (2006). The copula-GARCH model of conditional dependencies: An international stock market application. *Journal of International Money and Finance, 25*, 827–853.

Lee, T.-H., & Long, X. (2009). Copula-based multivariate GARCH model with uncorrelated dependent errors. *Journal of Econometrics, 150*, 207–218.

Nelson, R. B. (1998). *An introduction to copula*. New York: Springer.

Patton, A. J. (2004). On the out-of-sample importance of skewness and asymmetric dependence for asset allocation. *Journal of Financial Econometrics, 2*(1), 130–168.

Patton, A. J. (2006a). Estimation of multivariate models for time series of possibly different lengths. *Journal of Applied Econometrics, 21*(2), 147–173.

Patton, A. J. (2006b). Modelling asymmetric exchange rate dependence. *International Economic Review, 47*(2), 527–556.

Savu, C., & Trede, M. (2010). Hierarchies of Archimedean copulas. *Quantitative Finance, 10*, 295–304.

Sklar, A. (1959). *Fonctions de repartition an dimensions et leurs marges*. Paris: Publications de l'lnstitut de statistique de l'Universite de Paris.

Trivedi, P. K., & Zimmer, D. M. (2006). Using trivariate copulas to model sample selection and treatment effects: Application to family health care demand. *Journal of Business and Economic Statistics, 24*(1), 63.

Tsafack, G., & Garcia, R. (2011). Dependence structure and extreme comovements in international equity and bond markets. *Journal of Banking and Finance, 35*, 1954–1970.

Xu, Y. (2004). *Incorporating estimation risk in copula-based portfolio optimization*. Miami: Department of Economics, University of Miami.

Internet Bubble Examination with Mean-Variance Ratio

53

Zhidong D. Bai, Yongchang C. Hui, and Wing-Keung Wong

Contents

Abstract

To evaluate the performance of the prospects X and Y, financial professionals are interested in testing the equality of their Sharpe ratios (SRs), the ratios of the excess expected returns to their standard deviations. Bai et al. (Statistics and Probability Letters 81, 1078–1085, 2011d) have developed the mean-variance-ratio (MVR) statistic to test the equality of their MVRs, the ratios of the excess expected returns to its variances. They have also provided theoretical reasoning to use MVR and proved that their proposed statistic is uniformly most powerful unbiased. Rejecting the null hypothesis infers that X will have either smaller

Z.D. Bai (✉)
KLAS MOE & School of Mathematics and Statistics, Northeast Normal University, Changchun, China

Department of Statistics and Applied Probability, National University of Singapore, Singapore, Singapore
e-mail: baizd@nenu.edu.cn

Y.C. Hui
School of Mathematics and Statistics, Xi'an Jiaotong University, Xi'an, China
e-mail: huiyc180@mail.xjtu.edu.cn

W.-K. Wong
Department of Economics, Hong Kong Baptist University, Kowloon, Hong Kong
e-mail: awong@hkbu.edu.hk

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_53,

variance or larger excess mean return or both leading to the conclusion that X is the better investment. In this paper, we illustrate the superiority of the MVR test over the traditional SR test by applying both tests to analyze the performance of the S&P 500 index and the NASDAQ 100 index after the bursting of the Internet bubble in the 2000s. Our findings show that while the traditional SR test concludes the two indices being analyzed to be indistinguishable in their performance, the MVR test statistic shows that the NASDAQ 100 index underperformed the S&P 500 index, which is the real situation after the bursting of the Internet bubble in the 2000s. This shows the superiority of the MVR test statistic in revealing short-term performance and, in turn, enables investors to make better decisions in their investments.

Keywords
Mean-variance ratio • Sharpe ratio • Hypothesis testing • Uniformly most powerful unbiased test • Internet bubble • Fund management

53.1 Introduction

Internet stocks obtained huge gains in the late 1990s, followed by huge losses from early 2000. In just 2 years from 1998 to early March 2000, prices of Internet stocks rose by sixfold and outperformed the S&P 500 by 482 %. Technology stocks generally showed a similar trend based on the fact that NASDAQ 100 index quadrupled in value over the same period and outperformed the S&P 500 index by 268 %. On the other hand, NASDAQ 100 index dropped by 64.28 % in value during the Internet bubble crash and underperformed the S&P 500 index by 173.87 %.

The spectacular rise and fall of Internet stocks in the late 1990s has stimulated research into the causes of the Internet stock bubble. Theories had been developed to explain the Internet bubble. For example, Baker and Stein (2004) develop a model of market sentiment with irrationally overconfident investors and short-sale constraints. Ofek and Richardson (2003) provide circumstantial evidence that Internet stocks attract mostly retail investors who are more prone to be overconfident about their ability to predict future stock prices than institutional investors. Perkins and Perkins (1999) suggest that during the Internet boom, investors were confidently betting on the continued rise of Internet stocks because they knew that high demand and limited equity float implies substantial upside returns. Moreover, Ofek and Richardson (2003) provide indirect evidence that Internet stock prices were supported by a combination of factors such as limited float, short-sale constraints, and aggressive trend chased by retail investors, whereas Statman (2002) shows that this asymmetric payoff must have made Internet stocks appear to be an extremely attractive gamble for risk seekers. On the other hand, Fong et al. (2008) use stochastic dominance methodology (Fong et al. 2005; Broll et al. 2006; Chan et al. 2012; Lean et al. 2012) to identify dominant types of risk preferences in the Internet bull and bear markets. They conclude that investor risk preferences (Wong and Li 1999; Wong and Chan 2008) have changed over this cycle, and the change is related to utility theory (Wong 2007; Sriboonchitta et al. 2009) and behavioral finance (Lam et al. 2010, 2012).

In this paper, we apply both the mean-variance ratio (MVR) test and the Sharpe ratio (SR) test to examine the performance of the NASDAQ 100 index and the S&P 500 index during the bursting of the Internet bubble in the 2000s. The tests are relied on the theory of the mean-variance (MV) portfolio optimization (Markowitz 1952; Bai et al. 2009a, b). The Markowitz efficient frontier also provides the basis for many important financial economics advances, including the Sharpe-Lintner capital asset pricing model (CAPM, Sharpe 1964; Lintner 1965) and the well-known optimal one-fund theorem (Tobin 1958). Originally motivated by the MV analysis, the optimal one-fund theorem, and the CAPM model, the Sharpe ratio, the ratio of the excess expected return to its volatility or standard deviation, is one of the most commonly used statistics in the MV framework. The SR is now widely used in many different areas in Finance and Economics, from the evaluation of portfolio performance to market efficiency tests (see, e.g., Ofek and Richardson 2003).

Jobson and Korkie (1981) develop a SR statistic to test for the equality of two SRs. The test statistic has been modified and improved by Cadsby (1986) and Memmel (2003). Lo (2002) carries out a more thorough study of the statistical property of the SR estimator. Using standard econometric methods with several different sets of assumptions imposed on the statistical behavior of the returns series, Lo derives the asymptotic statistical distribution for the SR estimator and shows that confidence intervals, standard errors, and hypothesis tests can be computed for the estimated SRs in much the same way as regression coefficients such as portfolio alphas and betas are computed.

The SR test statistic developed by Jobson and Korkie (1981) and others provides a formal statistical comparison of performance among portfolios. One deficiency of the SR statistic is that it has only an asymptotic distribution. Hence, the SR test has its statistical properties only for large samples, but not for small samples. Nevertheless, the performance of assets is often compared by using small samples, especially when markets undergo substantial changes resulting from changes in short-term factors and momentum. Under these circumstances, it is more meaningful to use limited data to predict the assets' future performance. In addition, it is not meaningful to measure SRs for extended periods when the means and standard deviations of the underlying assets are found empirically to be nonstationary and/or to possess structural breaks. For small samples, the main difficulty in developing the SR test is that it is impossible to obtain a uniformly most powerful unbiased (UMPU) test to check for the equality of SRs. To circumvent this problem, Bai et al. (2011d) propose to use an alternative statistic, the MVR tests to compare performance of assets. They also discuss the evaluation of the performance of assets for small samples by providing a theoretical framework and then invoking both one-sided and two-sided UMPU MVR tests. Moreover, Bai et al. (2012) further extend the MVR statistics to compare the performance of prospects after the effect of the background risk has been mitigated.

Applying the traditional SR test, we fail to reject the possibility of having any significant difference between the performance of the S&P 500 index and the NASDAQ 100 index during the bursting of the Internet bubble in the 2000s. This finding implies that the two indices being analyzed could be indistinguishable in their performance during the period under the study. However, we conjecture that this conclusion is most likely to be inaccurate as the lack of sensitivity of the SR test in analyzing small samples. Thus, we propose to use the MVR test in the analysis. As expected, the MVR test shows that the MVR of the weekly return on S&P 500 index is different from that on the NASDAQ 100 index. We conclude that the NASDAQ 100 index underperformed the S&P 500 index during the period under the study. The proposed MVR test can discern the performance of the two indices and hence is more informative than tests using the SR statistics for investors to decide on their investments.

The rest of the paper is organized as follows: Section 53.2 discusses the data while Sect. 53.3 provides the theoretical framework and discusses the theory for both one-sided and two-sided MVR tests. In Sect. 53.4, we demonstrate the superiority of the MVR tests over the traditional SR tests by applying both tests to analyze the performance of the S&P 500 index and the NASDAQ 100 index during the bursting of the Internet bubble in the 2000s. This is followed by Sect. 53.4 which summarizes our conclusions and shares our insights.

53.2 Data

The data used in this study consists of weekly returns on two stock indices: the S&P 500 and the NASDAQ 100 index. We use the S&P 500 index to represent non-technology or "old economy" firms. Our proxy for the Internet and technology sectors is the NASDAQ 100 index. Firms represented in the NASDAQ 100 include those in the computer hardware and software, telecommunications, and biotechnology sectors. The NASDAQ 100 index is value weighted.

Our sample period is from January 1, 2000 to December 31, 2002, to study the effect of the crash in the Internet bubble. Before 2000, there is a clear upward trend in technology stock prices emerging from around that period and this period spans a period of intense IPO and secondary market activities for Internet stocks. Schultz and Zaman (2001) report that 321 Internet firms went public between January 1999 and March 2000, accounting for 76 % of all new Internet issues since the first wave of Internet IPOs began in 1996. Ofek and Richardson (2003) find that the extraordinary high valuations of Internet stocks between the early 1998 and February 2000 were accompanied by very high trading volume and liquidity. The unusually high volatility of technology stocks is only partially explained by the rise in the overall market volatility. Our interest centers on the bear market from January 1, 2000 to December 31, 2002. All data for this study are from datastream.

53.3 Methodology

Let X_i and Y_i ($i = 1, 2, \cdots, n$) be independent excess returns drawn from the corresponding normal distributions $N(\mu, \sigma^2)$ and $N(\eta, \tau^2)$ with joint density $p(x, y)$ such that

$$p(x,y) = k \times \exp\left(\frac{\mu}{\sigma^2}\sum x_i - \frac{1}{2\sigma^2}\sum x_i^2 + \frac{\eta}{\tau^2}\sum y_i - \frac{1}{2\tau^2}\sum y_i^2\right) \qquad (53.1)$$

Where $k = (2\pi\sigma^2)^{-n/2}(2\pi\tau^2)^{-n/2}\exp\left(-\frac{n\mu^2}{2\sigma^2}\right)\exp\left(-\frac{n\eta^2}{2\tau^2}\right)$

To evaluate the performance of the prospects X and Y, financial professionals are interested in testing the hypotheses

$$H_0^* : \frac{\mu}{\sigma} \le \frac{\eta}{\tau} \text{ versus } H_1^* : \frac{\mu}{\sigma} > \frac{\eta}{\tau} \qquad (53.2)$$

to compare the performance of their corresponding SRs, $\frac{\mu}{\sigma}$ and $\frac{\eta}{\tau}$, the ratios of the excess expected returns to their standard deviations.

If the hypothesis H_0^* is rejected, it infers that X is the better investment prospect with larger SR because X has either larger excess mean return or smaller standard deviation or both. Jobson and Korkie (1981) and Memmel (2003) develop test statistics to test the hypotheses in Eq. 53.2 for large samples but their tests would not be appropriate for testing small samples as the distribution of their test statistics is only valid asymptotically but not valid for small samples. However, it is especially relevant in investment decisions to test the hypotheses in Eq. 53.2 for small samples to provide useful investment information to investors. Furthermore, as it is impossible to obtain any UMPU test statistic to test the inequality of the SRs in Eq. 53.2 for small samples, Bai et al. (2011d) propose to use the following hypothesis to test for the inequality of the MVRs:

$$H_{01} : \frac{\mu}{\sigma^2} \le \frac{\eta}{\tau^2} \quad \text{versus} \quad H_{11} : \frac{\mu}{\sigma^2} > \frac{\eta}{\tau^2}. \qquad (53.3)$$

In addition, they develop the UMPU test statistic to test the above hypotheses. Rejecting the hypothesis H_0 infers that X will have either smaller variance or larger excess mean return or both leading to the conclusion that X is the better investment. As sometimes investors conduct the two-sided test to compare the MVRs, the following hypotheses are included in our study:

$$H_{02} : \frac{\mu}{\sigma^2} = \frac{\eta}{\tau^2} \quad \text{versus} \quad H_{12} : \frac{\mu}{\sigma^2} \neq \frac{\eta}{\tau^2}. \tag{53.4}$$

One may argue that the MVR test is that SR test is scale invariant, whereas the MV ratio test is not. To support the MVR test to be an acceptable alternative test statistic, Bai et al. (2011d) show the theoretical justification for the use of the MVR test statistic in the following remark:

Remark 53.1 *One may think that the MVR can be less favorable than the SR as the former is not scale invariant while the latter is. However, in some financial processes, the mean change in a short period of time is proportional to its variance change. For example, many financial processes can be characterized by the following diffusion process for stock prices formulated as*

$$dY_t = \mu^P(Y_t)dt + \sigma(Y_t)dW_t^P,$$

where μ^P is an N-dimensional function, σ is an $N \times N$ matrix and W_t^P is an N-dimensional standard Brownian motion under the objective probability measure P. Under this model, the conditional mean of the increment dY_t given Y_t is $\mu^P(Y_t)dt$ and the covariance matrix is $\sigma(Y_t)\sigma^T(Y_t)dt$. When $N = 1$, the SR will be close to 0 while the MVR will be independent of dt. Thus, when the time period dt is small, the MVR will be advantageous over the SR.

To further support for the use of MVR, Bai et al. (2011d) document the MVR in the context of Markowitz MV optimization theory as follows: suppose that there is p-branch of assets $\mathbf{S} = (s_1,\cdots, s_p)^T$ whose returns are denoted by $\mathbf{r} = (r_1,\cdots, r_p)^T$ with mean $\boldsymbol{\mu} = (\mu_1,\cdots, \mu_p)^T$ and covariance matrix $\Sigma = (\sigma_{ij})$. In addition, we suppose that investors will invest capital C on the p-branch of securities $\mathbf{S}$ such that they solve for their optimal investment plans $\mathbf{c} = (c_1,\cdots, c_p)^T$ to allocate their investable wealth on the p-branch of securities to obtain maximize return subject at a given level of risk.

The above maximization problem can be formulated as the following optimization problem:

$$\max R = \mathrm{c}^T\boldsymbol{\mu}, \text{ subject to } \mathrm{c}^T\Sigma\mathrm{c} \leq \sigma_0^2 \tag{53.5}$$

where σ_0^2 is a given risk level. We call R satisfying Eq. 53.5 the optimal return and $\mathbf{c}$ be its corresponding allocation plan. One could easily extend the separation theorem and the mutual fund theorem to obtain the solution of Eq. 53.5[1] from the following lemma:

[1]We note that Bai et al. (2009a, b, 2011c) have also used the same framework as in 53.5.

Lemma 53.1 *For the optimization setting displayed in* Eq. 53.5, *the optimal return, R, and its corresponding investment plan, c, are obtained as follows*:

$$R = \sigma_0 \sqrt{\boldsymbol{\mu}^T \Sigma^{-1} \boldsymbol{\mu}}$$

and

$$c = \frac{\sigma_0}{\sqrt{\boldsymbol{\mu}^T \Sigma^{-1} \boldsymbol{\mu}}} \Sigma^{-1} \boldsymbol{\mu}. \tag{53.6}$$

From Lemma 53.1, the investment plan, **c**, is proportional to the MVR when Σ is a diagonal matrix. Hence, when the asset is concluded as superior in performance utilizing the MVR test, its corresponding weight could then be computed based on the corresponding MVR test value. Thus, another advantage of using the MVR test over the SR test is that it not only allows investors to compare the performance of different assets, but it also provides investors with information of the assets weight. The MVR test enables investors to compute the corresponding allocation for the assets. On the other hand, as the SR is not proportional to the weight of the corresponding asset, an asset with the highest SR would not infer that one should put highest weight on this asset as compared with our MVR. In this sense, the test proposed by Bai et al. (2011d) is superior to the SR test.

Bai et al. (2011d) have also developed both one-sided UMPU test and two-sided UMPU test of equality of the MVRs in comparing the performances of different prospects with hypotheses stated in Eqs. 53.3 and 53.4, respectively. We first state the one-sided UMPU test for the MVRs as follows:

Theorem 53.1 *Let X_i and Y_i $(i = 1, 2, \cdots, n)$ be independent random variables with joint distribution function defined in* Eq. 53.1. *For the hypotheses setup in* Eq. 53.3, *there exists a UMPU level-α test with the critical function $\phi(u, t)$ such that*

$$\phi(u,t) \begin{cases} 1, & \text{when } u \geq C_0(t) \\ 0, & \text{when } u < C_0(t) \end{cases} \tag{53.7}$$

where C_0 is determined by

$$\int_{C_0}^{\infty} f_{n,t}^{*}(u) \ du = K_1; \tag{53.8}$$

with

$$f_{n,t}^{*}(u) = \left(t_2 - \frac{u^2}{n}\right)^{\frac{n-1}{2} - 1} \left(t_3 - \frac{(t_1 - u)^2}{n}\right)^{\frac{n-1}{2} - 1},$$
$$K_1 = \alpha \int_{\Omega} f_{n,t}^{*}(u) \ du;$$

in which

$$U = \sum_{i=1}^{n} X_i, \quad T_1 = \sum_{i=1}^{n} X_i + \sum_{i=1}^{n} Y_i,$$
$$T_2 = \sum_{i=1}^{n} X_i^2, \quad T_3 = \sum_{i=1}^{n} Y_i^2, \quad T = (T_1, T_2, T_3);$$

with $\Omega = \{u | \max(-\sqrt{nt_2}, t_1 - \sqrt{nt_3}) \leq u \leq \min(\sqrt{nt_2}, t_1 + \sqrt{nt_3})\}$ *to be the support of the joint density function of* (U, T).

We call the statistic U in Theorem 53.1 the one-sided MVR test statistic or simply the MVR test statistic for the hypotheses setup in Eq. 53.3 if no confusion arises. In addition, Bai et al. (2011d) have introduced the two-sided UMPU test statistic as stated in the following theorem to test for the equality of the MVRs listed in Eq. 53.4:

Theorem 53.2 *Let* X_i *and* Y_i $(i = 1, 2, \cdots, n)$ *be independent random variables with joint distribution function defined in* Eq. 53.1. *Then, for the hypotheses setup in* Eq. 53.4, *there exists a UMPU level-*α *test with critical function*

$$\phi(u,t) = \begin{cases} 1, & \text{when } u \leq C_1(t) \text{ or } \geq C_2(t) \\ 0, & \text{when } C_1(t) < u < C_2(t) \end{cases} \tag{53.9}$$

in which C_1 *and* C_2 *satisfy*

$$\begin{cases} \int_{C_1}^{C_2} f_{n,t}^*(u)\, du = K_2 \\ \int_{C_1}^{C_2} u f_{n,t}^*(u)\, du = K_3 \end{cases}, \tag{53.10}$$

where

$$K_2 = (1-\alpha)\int_{\Omega} f_{n,t}^*(u)\, du,$$
$$K_3 = (1-\alpha)\int_{\Omega} u f_{n,t}^*(u)\, du.$$

The terms $f_{n,t}^*(u)$, T_i $(i = 1, 2, 3)$ *and* T *are defined in* Theorem 53.1.

We call the statistic U in Theorem 53.2 the two-sided MVR test statistic or simply the MVR test statistic for the hypotheses setup in Eq. 53.4 if no confusion arises. To obtain the critical values C_1 and C_2 for the test, readers may refer to Bai et al. (2011d, 2012).

53.4 Illustration

In this section, we demonstrate the superiority of the MVR tests over the traditional SR tests by illustrating the applicability of the MVR tests to examine the Internet bubble during January 2000 and December 2002. For simplicity, we only demonstrate the two-sided UMPU test.[2] The data for this study consists of weekly returns on two stock indices: the S&P 500 and the NASDAQ 100 index. The sample period covers from January 2000 to December 2002 in which the data from the first week of November 2000 to the last week of January 2001 (3 months) are used to compute the MVR in January 2001, while the data from the first week of December 2000 to the last week of February 2001 are used to compute the MVR in February 2001, and so on. However, if the period used to compute the SRs is too short, the result would not be meaningful as discussed in our previous sections. Thus, we utilize a longer period from the first week of February 2000 to the last week of January 2001 (12 months) to compute the SR ratio in January 2001, from the first week of March 2000 to the last week of February 2001 to compute the SR ratio in February 2001, and so on.

Let X with mean μ_X and variance σ_X^2 be the weekly return on S&P 500 while Y with mean μ_Y and variance σ_Y^2 be the weekly return on the NASDAQ 100 index. We test the following hypotheses:

$$H_0 : \frac{\mu_X}{\sigma_X^2} = \frac{\mu_Y}{\sigma_Y^2} \quad \text{versus} \quad H_1 : \frac{\mu_X}{\sigma_X^2} \neq \frac{\mu_Y}{\sigma_Y^2}. \tag{53.11}$$

To test the hypotheses in Eq. 53.11, we first compute the values of the test function U for the MVR statistic shown in Eq. 53.9, then compute the critical values C_1 and C_2 under the test level of 5 % for the pair of indices and display the values in Table 53.1.

For comparison, we also compute the corresponding SR statistic developed by Jobson and Korkie (1981) and Memmel (2003) such that

$$z = \frac{\hat{\sigma}_Y \hat{\mu}_X - \hat{\sigma}_X \hat{\mu}_Y}{\sqrt{\hat{\theta}}}, \tag{53.12}$$

which follows standard normal distribution asymptotically with

$$\theta = \frac{1}{T}\left(2\sigma_X^2\sigma_Y^2 - 2\sigma_X\sigma_Y\sigma_{X,Y} + \frac{1}{2}\mu_X^2\sigma_Y^2 + \frac{1}{2}\mu_Y^2\sigma_X^2 - \frac{\mu_X\mu_Y}{\sigma_X\sigma_Y}\sigma_{X,Y}^2\right)$$

to test for the equality of the SRs for the funds by setting the following hypotheses such that

$$H_0^* : \frac{\mu_X}{\sigma_X} = \frac{\mu_Y}{\sigma_Y} \quad \text{versus} \quad H_1^* : \frac{\mu_X}{\sigma_X} \neq \frac{\mu_Y}{\sigma_Y}. \tag{53.13}$$

[2]The results of the one-sided test which draw a similar conclusion are available on request.

Table 53.1 The results of the mean-variance ratio test and Sharpe ratio test for NASDAQ and S&P 500, from January 2001 to December 2002

	MVR test			SR test
Date month/year	U	C_1	C_2	Z
01/2001	−0.0556	−0.1812	0.1267	1.0906
02/2001	−0.0636	−0.1843	0.1216	1.8765
03/2001	−0.1291	−0.2291	0.0643	1.1787
04/2001	−0.0633	−0.2465	0.1633	0.9590
05/2001	0.0212	−0.1937	0.2049	0.8313
06/2001	0.0537	−0.1478	0.1983	0.8075
07/2001	−0.0421	−0.1399	0.1132	0.6422
08/2001	−0.1062	−0.1815	0.0886	0.6816
09/2001	−0.1623*	0.1665	0.2728	1.0125
10/2001	−0.1106	−0.3507	0.1742	0.5931
11/2001	0.0051	−0.2386	0.2825	0.1898
12/2001	0.1190	0.0165	0.2041	−0.1573
01/2002	0.0316	−0.0744	0.1389	0.0157
02/2002	−0.0067	−0.1389	0.1013	0.0512
03/2002	−0.0216	−0.1349	0.0853	−0.1219
04/2002	−0.0444	−0.1739	0.0848	0.1885
05/2002	−0.0588	−0.1766	0.1094	0.0446
06/2002	−0.1477	−0.2246	0.0267	0.3408
07/2002	−0.2167*	−0.0101	0.0578	0.0984
08/2002	−0.1526*	0.0452	0.1242	0.1024
09/2002	−0.2121*	−0.0218	0.0551	−0.6304
10/2002	0.0416	−0.1249	0.2344	−0.0361
11/2002	0.0218	−0.1056	0.2150	0.0008
12/2002	0.1265	−0.0015	0.2417	0.3908

Note: The MVR test statistic U is defined in Eq. 53.9 and its critical values C_1 and C_2 are defined in Eqs. 53.10, respectively. The SR test statistic Z is defined in Eq. 53.12. The level is $\alpha = 0.05$, and "*" means significant at levels 5 %. Here, the sample size of the MVR test is 3 months, while the sample size of the SR test 12 months. Recall that $\pm z_{0.025} \approx \pm 1.96$

Instead of using a 2-month data to compute the values of our proposed statistic, we use the overlapping 12-month data to compute the SR statistic. The results are also reported in Table 53.1.

The limitation of applying the SR test is that it would usually conclude indistinguishable performances between the indices, which may not be the situation in reality. In this aspect, looking for a statistic to evaluate the difference between indices for short periods is essential. The situation in reality is that the Internet stocks registered large gains in the late 1990s, followed by large losses from 2000. As we mentioned before, the NASDAQ 100 index comprises 100 of the largest domestic and international technology firms including those in the computer hardware and software, telecommunications, and biotechnology sectors, while the S&P

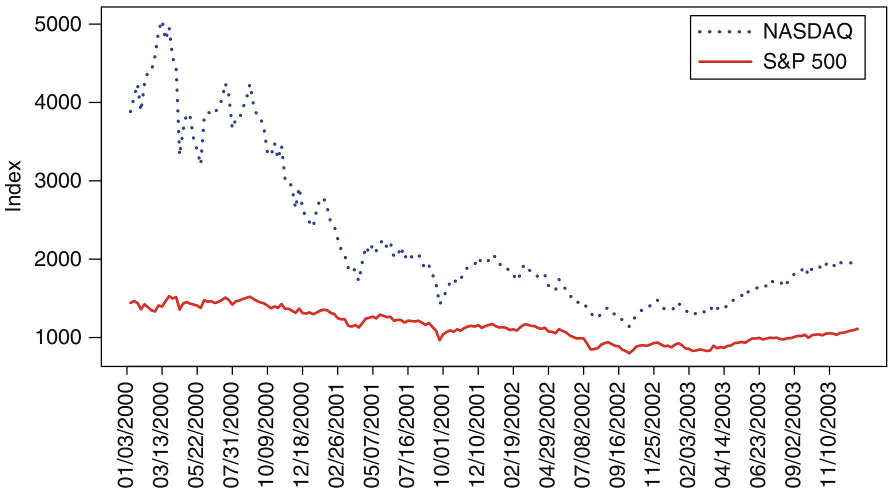


Fig. 53.1 Weekly indices of *NASDAQ* and S&P 500 from January 3, 2000 to December 31, 2003

500 index represents non-technology or "old economy" firms. After the bursting of the Internet bubble in the 2000s, as shown in Fig. 53.1, the NASDAQ 100 declined much more and underperformed the S&P 500. From Table 53.1, we find that the MVR test statistic does not disappoint us in that it does pick up significant differences in performances between the S&P 500 and the NASDAQ 100 index in September 2001, July 2002, August 2002, and September 2002, but SR test does not conclude any distinguishable performances between the indices. Further to say, from Table 53.1, we observe that $\hat{\mu}_X > \hat{\mu}_Y$ in September 2001, July 2002, August 2002, and September 2002. This infers that the MVR test statistics can detect the real situation that the NASDAQ 100 index underperformed the S&P 500 index, but the traditional SR test cannot detect any difference. Thus, we conclude that investors could be able to profiteer from the Internet bubble if they apply the MVR test.

53.5 Concluding Remarks

In this paper, we employ the MVR test statistics developed by Bai et al. (2011d) to examine the performances between the S&P 500 index and the NASDAQ 100 index during Internet bubble from January 2000 to December 2002. We illustrate the superiority of the MVR test over the traditional SR test by applying both tests to analyze the performance of the S&P 500 index and the NASDAQ 100 index after the bursting of the Internet bubble in the 2000s. Our findings show that while the traditional SR test concludes the two indices being analyzed to be indistinguishable in their performance, the MVR test statistic shows that the NASDAQ 100 index underperformed the S&P 500 index, which is the real situation after the bursting of the Internet bubble in the 2000s. This shows the superiority of the MVR test statistic in revealing short-term performance and, in turn, enables the investors to make better decisions about their investments.

There are two basic approaches to the problem of portfolio selection under uncertainty. One approach is based on the concept of utility theory (Gasbarro et al. 2007; Wong et al. 2006, 2008). Several stochastic dominance (SD) test statistics have been developed; see, for example, Bai et al. (2011a) and the references therein for more information. This approach offers a mathematically rigorous treatment for portfolio selection, but it is not popular among investors since investors would have to specify their utility functions and choose a distributional assumption for the returns before making their investment decisions.

The other approach is the mean-risk (MR) analysis that has been discussed in this paper. In this approach, the portfolio choice is made with respect to two measures – the expected portfolio mean return and portfolio risk. A portfolio is preferred if it has higher expected return and smaller risk. These are convenient computational recipes and they provide geometric interpretations for the trade-off between the two measures. A disadvantage of the latter approach is that it is derived by assuming the Von Neumann-Morgenstern quadratic utility function and that returns are normally distributed (Hanoch and Levy 1969). Thus, it cannot capture the richness of the former approach. Among the MR analyses, the most popular measure is the SR introduced by Sharpe (1966). As the SR requires strong assumptions that the returns of assets being analyzed have to be iid, various measures for MR analysis have been developed to improve the SR, including the Sortino ratio (Sortino and van der Meer 1991), the conditional SR (Agarwal and Naik 2004), the modified SR (Gregoriou and Gueyie 2003), value at risk (Ma and Wong 2010), expected shortfall (Chen 2008), and the mixed Sharpe ratio (Wong et al. 2012). However, most of the empirical studies, see, for example, Eling and Schuhmacher (2007), find that the conclusions drawn by using these ratios are basically the same as that drawn by the SR. Nonetheless, Leung and Wong (2008) have developed a multiple SR statistic and find that the results drawn from the multiple Sharpe ratio statistic can be different from its counterpart pair-wise SR statistic comparison, indicating that there are some relationships among the assets that have not being revealed using the pair-wise SR statistics. The MVR test could be the right candidate to reveal these relationships.

One may claim that the limitation of the MVR test statistic is that it can only draw conclusion for investors with quadratic utility functions and for normal-distributed assets. Wong (2006), Wong and Ma (2008), and others have shown that the conclusion drawn from the MR comparison is equivalent to the comparison of expected utility maximization for any risk-averse investor, not necessarily with only quadratic utility function, and for assets with any distribution, not necessarily normal distribution, if the assets being examined belong to the same location-scale family. In addition, one can also apply the results from Li and Wong (1999) and Egozcue and Wong (2010) to generalize the result so that it will be valid for any risk-averse investor and for portfolios with any distribution if the portfolios being examined belong to the same convex combinations of (same or different) location-scale families. The location-scale family can be very large, containing normal distributions as well as t-distributions, gamma distributions, etc. The stock returns could be expressed as convex combinations of normal distributions, t-distributions, and other location-scale families; see, for example, Wong and Bian (2000) and the references therein for more information. Thus, the conclusions drawn from the MVR test statistics are valid for most of the stationary data including most, if not all, of the returns of different portfolios.

Last, we note that to improve the effectiveness of applying the MVR test in evaluating financial assets performance, one may incorporate other techniques/ approaches/models, for example, fundamental analysis (Wong and Chan 2004), technical analysis (Wong et al. 2001, 2003), behavioral finance (Matsumura et al. 1990), prospect theory (Broll et al. 2010; Egozcue et al. 2011), and advanced econometrics (Wong and Miller 1990; Bai et al. 2010, 2011b), to measure the performance of different financial assets and assist investors to make wiser decisions.

Acknowledgment We would like to thank the editor C.-F. Lee for his substantive comments that have significantly improved this manuscript. The third author would also like to thank Professors Robert B. Miller and Howard E. Thompson for their continuous guidance and encouragement. The research is partially supported by grants from North East Normal University, National University of Singapore, Hong Kong Baptist University and the Research Grants Council of Hong Kong. The first author thanks the financial support from NSF China grant 11171057, Program for Changjiang Scholars and Innovative Research Team in University, and the Fundamental Research Funds for the Central Universities and NUS grant R-155-000-141-112.

References

Agarwal, V., & Naik, N. Y. (2004). Risk and portfolios decisions involving hedge funds. *Review of Financial Studies, 17*, 63–98.

Bai, Z. D., Liu, H. X., & Wong, W. K. (2009a). Enhancement of the applicability of Markowitz's portfolio optimization by utilizing random matrix theory. *Mathematical Finance, 19*, 639–667.

Bai, Z. D., Liu, H. X., & Wong, W. K. (2009b). On the markowitz mean-variance analysis of self-financing portfolios. *Risk and Decision Analysis, 1*, 35–42.

Bai, Z. D., Wong, W. K., & Zhang, B. Z. (2010). Multivariate linear and non-linear causality tests. *Mathematics and Computers in Simulation, 81*, 5–17.

Bai, Z. D., Li, H., Liu, H. X., & Wong, W. K. (2011a). Test statistics for prospect and markowitz stochastic dominances with applications. *Econometrics Journal, 14*, 278–303.

Bai, Z. D., Li, H., Wong, W. K., & Zhang, B. Z. (2011b). Multivariate causality tests with simulation and application. *Statistics and Probability Letters, 81*, 1063–1071.

Bai, Z. D., Liu, H. X., & Wong, W. K. (2011c). Asymptotic properties of eigenmatrices of a large sample covariance matrix. *Annals of Applied Probability, 21*, 1994–2015.

Bai, Z. D., Wang, K. Y., & Wong, W. K. (2011d). Mean-variance ratio test, a complement to coefficient of variation test and Sharpe ratio test. *Statistics and Probability Letters, 81*, 1078–1085.

Bai, Z. D., Hui, Y. C., Wong, W. K., & Zitikis, R. (2012). Evaluating prospect performance: Making a case for a non-asymptotic UMPU test. *Journal of Financial Econometrics, 10*(4), 703–732.

Baker, M., & Stein, J. C. (2004). Market liquidity as a sentiment indicator. *Journal of Financial Markets, 7*, 271–300.

Broll, U., Wahl, J. E., & Wong, W. K. (2006). Elasticity of risk aversion and international trade. *Economics Letters, 91*, 126–130.

Broll, U., Egozcue, M., Wong, W. K., & Zitikis, R. (2010). Prospect theory, indifference curves, and hedging risks. *Applied Mathematics Research Express, 2010*, 142–153.

Cadsby, C. B. (1986). Performance hypothesis testing with the Sharpe and Treynor measures: A comment. *Journal of Finance, 41*, 1175–1176.

Chan, C. Y., de Peretti, C., Qiao, Z., & Wong, W. K. (2012). Empirical test of the efficiency of UK covered warrants market: Stochastic dominance and likelihood ratio test approach. *Journal of Empirical Finance, 19*, 162–174.

Chen, S. X. (2008). Nonparametric estimation of expected shortfall. *Journal of Financial Econometrics, 6*, 87–107.

Egozcue, M., & Wong, W. K. (2010). Gains from diversification on convex combinations: A majorization and stochastic dominance approach. *European Journal of Operational Research, 200*, 893–900.

Egozcue, M., Fuentes García, L., Wong, W. K., & Zitikis, R. (2011). Do investors like to diversify? A study of Markowitz preferences. *European Journal of Operational Research, 215*, 188–193.

Eling, M., & Schuhmacher, F. (2007). Does the choice of performance measure influence the evaluation of hedge funds? *Journal of Banking and Finance, 31*, 2632–2647.

Fong, W. M., Wong, W. K., & Lean, H. H. (2005). International momentum strategies: A stochastic dominance approach. *Journal of Financial Markets, 8*, 89–109.

Fong, W. M., Lean, H. H., & Wong, W. K. (2008). Stochastic dominance and behavior towards risk: The market for internet stocks. *Journal of Economic Behavior and Organization, 68*, 194–208.

Gasbarro, D., Wong, W. K., & Zumwalt, J. K. (2007). Stochastic dominance analysis of iShares. *European Journal of Finance, 13*, 89–101.

Gregoriou, G. N., & Gueyie, J. P. (2003). Risk-adjusted performance of funds of hedge funds using a modified Sharpe ratio. *Journal of Wealth Management, 6*, 77–83.

Hanoch, G., & Levy, H. (1969). The efficiency analysis of choices involving risk. *Review of Economic Studies, 36*, 335–346.

Jobson, J. D., & Korkie, B. (1981). Performance hypothesis testing with the Sharpe and Treynor measures. *Journal of Finance, 36*, 889–908.

Lam, K., Liu, T., & Wong, W. K. (2010). A pseudo-Bayesian model in financial decision making with implications to market volatility, under- and overreaction. *European Journal of Operational Research, 203*, 166–175.

Lam, K., Liu, T., & Wong, W. K. (2012). A new pseudo Bayesian model with implications to financial anomalies and investors' behaviors. *Journal of Behavioral Finance, 13*(2), 93–107.

Lean, H. H., Phoon, K. F., & Wong, W. K. (2012). Stochastic dominance analysis of CTA funds. *Review of Quantitative Finance and Accounting*, doi:10.1007/s11156-012-0284-1.

Leung, P. L., & Wong, W. K. (2008). On testing the equality of the multiple Sharpe ratios, with application on the evaluation of iShares. *Journal of Risk, 10*, 1–16.

Li, C. K., & Wong, W. K. (1999). Extension of stochastic dominance theory to random variables. *RAIRO Recherche Opérationnelle, 33*, 509–524.

Lintner, J. (1965). The valuation of risky assets and the selection of risky investment in stock portfolios and capital budgets. *Review of Economics and Statistics, 47*, 13–37.

Lo, A. (2002). The statistics of Sharpe ratios. *Financial Analysis Journal, 58*, 36–52.

Ma, C., & Wong, W. K. (2010). Stochastic dominance and risk measure: A decision-theoretic foundation for VaR and C-VaR. *European Journal of Operational Research, 207*, 927–935.

Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance, 7*, 77–91.

Matsumura, E. M., Tsui, K. W., & Wong, W. K. (1990). An extended multinomial-Dirichlet model for error bounds for dollar-unit sampling. *Contemporary Accounting Research, 6*, 485–500.

Memmel, C. (2003). Performance hypothesis testing with the Sharpe ratio. *Finance Letters, 1*, 21–23.

Ofek, E., & Richardson, M. (2003). A survey of market efficiency in the internet sector. *Journal of Finance, 58*, 1113–1138.
Perkins, A. B., & Perkins, M. C. (1999). *The internet bubble: Inside the overvalued world of high-tech stocks*. New York: Harper Business.
Schultz, P., & Zaman, M. (2001). Do the individuals closest to internet firms believe they are overvalued? *Journal of Financial Economics, 59*, 347–381.
Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance, 19*, 425–442.
Sharpe, W. F. (1966). Mutual funds performance. *Journal of Business, 39*, 119–138.
Sortino, F. A., & van der Meer, R. (1991). Downside risk. *Journal of Portfolio Management, 17*, 27–31.
Sriboonchitta, S., Wong, W. K., Dhompongsa, D., & Nguyen, H. T. (2009). *Stochastic dominance and applications to finance, risk and economics*. Boca Raton: Chapman and Hall/CRC.
Statman, M. (2002). Lottery players/stock traders. *Journal of Financial Planning*, 14–21.
Tobin, J. (1958). Liquidity preference as behavior towards risk. *Review of Economic Studies, 25*, 65–86.
Wong, W. K. (2006). Stochastic dominance theory for location-scale family. *Journal of Applied Mathematics and Decision Sciences, 2006*, 1–10.
Wong, W. K. (2007). Stochastic dominance and mean-variance measures of profit and loss for business planning and investment. *European Journal of Operational Research, 182*, 829–843.
Wong, W. K., & Bian, G. (2000). Robust Bayesian inference in asset pricing estimation. *Journal of Applied Mathematics & Decision Sciences, 4*, 65–82.
Wong, W. K., & Chan, R. (2004). The estimation of the cost of capital and its reliability. *Quantitative Finance, 4*, 365–372.
Wong, W. K., & Chan, R. (2008). Markowitz and prospect stochastic dominances. *Annals of Finance, 4*, 105–129.
Wong, W. K., & Li, C. K. (1999). A note on convex stochastic dominance theory. *Economics Letters, 62*, 293–300.
Wong, W. K., & Ma, C. (2008). Preferences over location-scale family. *Economic Theory, 37*, 119–146.
Wong, W. K., & Miller, R. B. (1990). Analysis of ARIMA-noise models with repeated time series. *Journal of Business and Economic Statistics, 8*, 243–250.
Wong, W. K., Chew, B. K., & Sikorski, D. (2001). Can P/E ratio and bond yield be used to beat stock markets? *Multinational Finance Journal, 5*, 59–86.
Wong, W. K., Manzur, M., & Chew, B. K. (2003). How rewarding is technical analysis? Evidence from Singapore stock market. *Applied Financial Economics, 13*, 543–551.
Wong, W. K., Thompson, H. E., Wei, S., & Chow, Y. F. (2006). Do winners perform better than losers? A stochastic dominance approach. *Advances in Quantitative Analysis of Finance and Accounting, 4*, 219–254.
Wong, W. K., Phoon, K. F., & Lean, H. H. (2008). Stochastic dominance analysis of Asian hedge funds. *Pacific-Basin Finance Journal, 16*, 204–223.
Wong, W. K., Wright, J. A., Yam, S. C. P., & Yung, S. P. (2012). A mixed Sharpe ratio. *Risk and Decision Analysis, 3*(1–2), 37–65.

Quantile Regression in Risk Calibration

54

Shih-Kang Chao, Wolfgang Karl Härdle, and Weining Wang

Contents

Abstract

Financial risk control has always been challenging and becomes now an even harder problem as joint extreme events occur more frequently. For decision makers and government regulators, it is therefore important to obtain accurate

The financial support from the Deutsche Forschungsgemeinschaft via SFB 649 "Ökonomisches Risiko," Humboldt-Universität zu Berlin is gratefully acknowledged

S.-K. Chao (✉) • W. Wang
Ladislaus von Bortkiewicz Chair of Statistics, C.A.S.E. – Center for Applied Statistics and Economics, Humboldt–Universität zu Berlin, Berlin, Berlin, Germany
e-mail: shih-kang.chao@cms.hu-berlin.de; wangwein@cms.hu-berlin.de

W.K. Härdle
Ladislaus von Bortkiewicz Chair of Statistics, C.A.S.E. – Center for Applied Statistics and Economics, Humboldt–Universität zu Berlin, Berlin, Berlin, Germany

Lee Kong Chian School of Business, Singapore Management University, Singapore, Singapore
e-mail: haerdle@wiwi.hu-berlin.de

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_54,

information on the interdependency of risk factors. Given a stressful situation for one market participant, one likes to measure how this stress affects other factors. The CoVaR (Conditional VaR) framework has been developed for this purpose. The basic technical elements of CoVaR estimation are two levels of quantile regression: one on market risk factors; another on individual risk factor.

Tests on the functional form of the two-level quantile regression reject the linearity. A flexible semiparametric modeling framework for CoVaR is proposed. A partial linear model (PLM) is analyzed. In applying the technology to stock data covering the crisis period, the PLM outperforms in the crisis time, with the justification of the backtesting procedures. Moreover, using the data on global stock markets indices, the analysis on marginal contribution of risk (MCR) defined as the local first order derivative of the quantile curve sheds some light on the source of the global market risk.

Keywords

CoVaR • Value-at-Risk • Quantile regression • Locally linear quantile regression • Partial linear model • Semiparametric model

54.1 Introduction

Sufficiently accurate risk measures are needed not only in crisis times. In the last two decades, the world has gone through several financial turmoils, and the financial market is getting riskier and the scale of loss soars. Beside marginal extremes that can shock even a well-diversified portfolio, the focus of intensified research in the recent years has been on understanding the interdependence of risk factors and their conditional structure.

The most popular risk measure is the Value-at-Risk (VaR), which is defined as the τ-quantile of the return distribution at time $t + d$ conditioned on the information set $\mathcal{F}_t$:

$$VaR^{\tau}_{t+d} \overset{\text{def}}{=} \inf\{x \in \mathbb{R} : \mathrm{P}(X_{t+d} \le x|\mathcal{F}_t) \ge \tau\}. \tag{54.1}$$

Here X_t denotes the asset return and τ is taking values such as 0.05, 0.01 or 0.001 to reflect negative extreme risk.

Extracting information in economic variables to predict VaR brings quantile regression into play here, since VaR is the quantile of the conditional asset return distribution. Engle and Manganelli (2004) propose the nonlinear Conditional Autoregressive Value-at-Risk (CaViaR) model, which uses (lag) VaR and lag returns. Chernozhukov and Umantsev (2001) propose linear and quadratic time series models for VaR prediction. Kuan et al. (2009) propose the Conditional AutoRegressive Expectile (CARE) model, and argue that expectiles are more sensitive to the scale of losses. These studies and many others apply quantile regression in a prespecified, often linear functional form. In a more nonparametric context, Cai and Wang (2008) estimate the conditioned cdf by a double kernel local linear estimator and find the quantile by inverting the cdf. Schaumburg (2011) uses the same technique together with extreme value theory for VaR prediction. Taylor (2008) proposes Exponentially Weighted Quantile Regression (EWQR) for estimating VaR time series.

The aforementioned studies focus mainly on the VaR estimation for single assets and do not directly take into account the escalated spillover effect in crisis periods. This risk of joint tail events of asset returns has been identified and studied. Further, Brunnermeier and Pedersen (2008) show that the negative feedback effect of a "loss spiral" and a "margin spiral" leads to the joint depreciation of assets prices. It is therefore important to develop risk measures which can quantify the contagion effects of negative extreme event.

Acharya et al. (2010) propose the concept of marginal expected shortfall (MES), which measures the contribution of individual assets to the portfolio expected shortfall. Via an equilibrium argument, the MES is shown to be a predictor to a financial institution's risk contribution. Brownlees and Engle (2012) demonstrate that the MES can be written as a function of volatility, correlation, and expectation conditional on tail events. Huang et al. (2012) propose the distress insurance premium (DIP), a measure similar to MES but computed under the risk-neutral probability. This measure can therefore be viewed as the market insurance premium against the event that the portfolio loss exceeds a low level. Adams et al. (2012) construct financial indices on return of insurance companies, commercial banks, investment banks, and hedge funds, and use a linear model for the VaRs of the four financial indices to forecast the state-dependent sensitivity VaR (SDSVaR). The risk measures proposed above have some shortcomings though: The computation of DIP is demanding since this involves the simulation of rare events. MES suffers from the scarcity of data because it conditions on a rare event.

In Adrian and Brunnermeier (2011) (henceforth AB), the CoVaR concept of conditional VaR is proposed, which controls the effect of the negative extreme event of some systemically risky financial institutions. Formally, let $C(X_{i,t})$ be some event of a asset i return $X_{i,t}$ at time t and take $X_{j,t}$ as another asset return (e.g., the market index). The $\mathrm{CoVaR}^{\tau}_{j|i,t}$ is defined as the τ-quantile of the conditional probability distribution:

$$\mathrm{P}\left\{X_{j,t} \leq \mathrm{CoVa}R^{\tau}_{j|i,t} \middle| C\left(X_{i,t}\right), M_t\right\} = \tau, \tag{54.2}$$

where M_t is a vector of market variables defined in Sect. 54.2.1. The standard CoVaR approach is to set $C\left(X_{i,t}\right) = \left\{X_{i,t} = VaR^{\tau}_{X_{i,t}}\right\}$. In AB, $X_{j,t}$ is the weekly return which is constructed from a vast data set comprised of all publicly traded commercial banks, broker dealers, insurance companies, and real estate companies in the USA. Further, AB propose ΔCoVaR (measure of marinal risk contribution) as the difference between $CoVaR^{\tau_1}_{j|i,t}$ and $CoVaR^{\tau_2}_{j|i,t}$, where $\tau_1 = 0.5$ is associated with the normal state and $\tau_2 = 0.05$ is associated with the financial distress state.

The formulation of this conditional risk measure has several advantages. First, the cloning property: After dividing a systemically risky firm into several clones, the value of CoVaR conditioned on the entire firm does not differ from the one conditioned on one of the clones. Second, the conservativeness. The CoVaR value is more conservative than VaR, because it conditions on an extreme event. Third, CoVaR is endogenously generated and adapted to the varying environment of the market.

The recipe of AB for CoVaR construction is as follows: In the first step, one predicts the VaR of an individual asset $X_{i,t}$ through a linear model on market variables:

$$X_{i,t} = \alpha_i + \gamma_i^{\mathrm{T}} M_{t-1} + \varepsilon_{i,t}, \tag{54.3}$$

where γ_i^{T} means the transpose of γ_i and M_t is a vector of the state variables (see Sect. 54.2.1). This model is estimated with quantile regression of Koenker and Bassett (1978) to get the coefficients $(\hat{\alpha}_i, \hat{\gamma}_i)$ with $F_{\varepsilon_{i,t}}^{-1}(\tau|M_{t-1}) = 0$. The VaR of asset i is predicted by

$$\widehat{VaR}_{i,t} = \hat{\alpha}_i + \hat{\gamma}_i^{\mathrm{T}} M_{t-1}. \tag{54.4}$$

In the second step, one models the asset j return as a linear function of asset return i and market variables M_t:

$$X_{j,t} = \alpha_{j|i} + \beta_{j|i} X_{i,t} + \gamma_{j|i}^{\mathrm{T}} M_{t-1} + \varepsilon_{j,t}, \tag{54.5}$$

Again one employs quantile regression and obtains coefficients $\left(\hat{\alpha}_{j|i}, \hat{\beta}_{j|i}, \hat{\gamma}_{j|i}\right)$. The CoVaR is finally calculated as:

$$\widehat{CoVaR}_{j|i,t}^{AB} = \hat{\alpha}_{j|i} + \hat{\beta}_{j|i} \widehat{VaR}_{i,t} + \hat{\gamma}_{j|i}^{\mathrm{T}} M_{t-1}. \tag{54.6}$$

In Eq. 54.5, the variable $X_{i,t}$ influences the return $X_{j,t}$ in a linear fashion. However, the linear parametric model may not be flexible enough to capture the tail dependence between i and j. The linearity of the conditioned quantile curves of X_j on X_i is challenged by the confidence bands of the nonparametric quantile curves, as shown in Fig. 54.1. The left tail quantile from linear parametric quantile regression (red) lies well outside the confidence band (gray dashed curve) of Hardle and Song (2010). This motivates empirically that a linear model is not flexible enough for the CoVaR question at hand.

Nonparametric models can be used to account for the nonlinear structure of the conditional quantile, but the challenge for using such models is the curse of dimensionality, as the quantile regression in CoVaR modeling often involves many variables. Thus, we resort to semiparametric partial linear model (PLM) which preserves some flexibility of the nonparametric model while suffers little from the curse of dimensionality.

As an illustration, the VaR/CoVaR of Goldman Sachs (GS) returns are shown, given the returns of Citigroup (C) and S&P500 (SP). S&P500 index return is used as a proxy for the market portfolio return.

Choosing market variables is crucial for the VaR/CoVaR estimation. For the variables representing market states, we follow the most popular choices such as VIX, short-term liquidity spread, etc. In particular, the variable we use for real estate companies is the Dow Jones U.S. real estate index. The daily data date from August 4, 2006 to August 4, 2011.

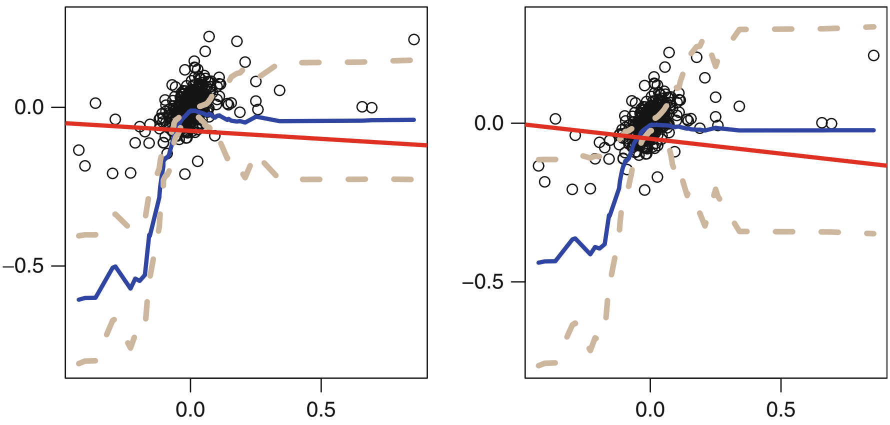


Fig. 54.1 Goldman Sachs (*GS*) and Citigroup (*C*) weekly returns 0.05(*left*) and 0.1(*right*) quantile functions. The *y*-axis is GS daily returns and the *x*-axis is the C daily returns. The *blue curve* are the locally linear quantile regression curves (see Appendix 1). The locally linear quantile regression bandwidth are 0.1026 and 0.0942. The *red lines* are the linear parametric quantile regression line. The antique *white dashed curves* are the asymptotic confidence band (see Appendix 2) with significance level 0.05. The sample size $N = 546$

To see if the estimated VaRs/CoVaRs are accurate, we utilize the backtesting procedures described in Berkowitz et al. (2011). We compare three (Co)VaR estimating methods in this study: VaR computed by linear quantile regression on market variables; CoVaR; PLM CoVaR proposed here. The VaR is one-sided interval prediction, the violations (the asset return exceeds estimated VaR/CoVaR) should happen unpredictably if the VaR algorithm is accurate. In other words, the null hypothesis is that the series of violations of VaR is a martingale difference, given all the past information. Furthermore, if the time series is autocorrelated, we can reject the null hypothesis of martingale difference right away; therefore, autocorrelation tests can be utilized in this context. The Ljung-Box test is not the most appropriate approach here since it has a too strong null hypothesis (i.i.d. sequence). Thus, we additionally apply the Lobato test. The CaViaR test, which is inspired by the CaViaR model, is proposed and shown to have the best overall performance by Berkowitz et al. (2011) among other alternative tests with an exclusive desk-level data set. To illustrate the VaR/CoVaR performances in the crisis time, we separately apply the CaViaR test to the violations of the whole sample period and to the financial crisis period.

The results show that during the financial crisis period from mid-2008 to mid-2009, the PLM CoVaR of GS given C performs better than that constructed from the technique of AB and the PLM CoVaR given SP. In particular, these results suggest that with appropriate modeling techniques (accounting for nonlinearity), the CoVaR of GS calculated from conditioning on C reflects some structurally risk which is not reflected from conditioning on market returns such as SP during financial crisis.

In contrast to ΔCoVaR, we use a mathematically more intuitive way to analyze the marginal effect by taking the first order derivative of the quantile function.

We call it "marginal contribution of risk" (MCR). Bae et al. (2003) and many others have pointed out the phenomenon of financial contagion across national borders. This motivates us to consider the stock indices of a few developed markets and explore their risk contribution to the global stock market. MCR results show that when the global market condition varies, the source of global market risk can be different. To be more specific, when the global market return is bad, the risk contribution from the USA is the largest. On the other hand, during financially stable periods, Hong Kong and Japan are more significant risk contributors than the USA to the global market.

This study is organized as follows: Sect. 54.2 introduces the construction and the estimation of the PLM model of CoVaR. The backtesting methods and our risk contribution measure are also introduced in this section. Section 54.3 presents the Goldman Sachs CoVaR time series and the backtesting procedure results. Section 54.4 presents the conclusion and possible further studies. Appendices describe the detailed estimation and statistical inference procedures used in this study.

54.2 Methodology

Quantile regression is a well-established technique to estimate the conditional quantile function. Koenker and Bassett (1978) focus on the linear functional form. An extension of linear quantile regression is the PLM quantile regression. A partial linear model for the dynamics of assets return quantile is constructed in this section. The construction is justified by a linearity test based on a conservative uniform confidence band proposed in Hardle and Song (2010). For more details on semiparametric modeling and PLM, we refer to Härdle et al. (2004) and Härdle et al. (2000).

The backtesting procedure is done via the CaViaR test. Finally, the methodology of MCR is introduced, which is an intuitive marginal risk contribution measure. We will apply the method to a data set of global market indices in developed countries.

54.2.1 Constructing Partial Linear Model (PLM) for CoVaR

Recall how the CoVaR is constructed:

$$\widehat{VaR}i,t = \hat{\alpha}_i + \hat{\gamma}_i M_{t-1},$$
$$\widehat{CoVaR}^{AB}_{j|i,t} = \hat{\alpha}_{j|i} + \hat{\beta}_{j|i}\widehat{VaR}_{i,t} + \hat{\gamma}^{\mathrm{T}}_{j|i} M_{t-1}.$$

where $(\hat{\alpha}_i, \hat{\gamma}_i)$ and $\left(\hat{\alpha}_{j|i}, \hat{\beta}_{j|i}, \hat{\gamma}_{j|i}\right)$ are estimated from a linear model using standard linear quantile regression.

We have motivated the need for more general functional forms for the quantile curve. We therefore relax the model to a non- or semiparametric model. The market variable M_t is multidimensional, and the data frequency here is daily. The following key variables are entering our analysis:

1. VIX: Measuring the model-free implied volatility of the market. This index is known as the "fear gauge" of investors. The historical data can be found in the Chicago Board Options Exchange's website.
2. Short-term liquidity spread: Measuring short-term liquidity risk by the difference between the 3-month treasury repo rate and the 3-month treasury bill rate. The repo data is from the Bloomberg database and the treasury bill rate data is from the Federal Reserve Board H.15.
3. The daily change in the 3-month treasury bill rate: AB find that the changes have better explanatory power than the levels for the negative tail behavior of asset returns.
4. The change in the slope of the yield curve: The slope is defined by the difference of the 10-year treasury rate and the 3-month treasury bill rate.
5. The change in the credit spread between 10 years BAA-rated bonds and the 10 years treasury rate.
6. The daily Dow Jones U.S. Real Estate index returns: The index reflects the information of lease rates, vacancies, property development, and transactions of real estates in the USA.
7. The daily S&P500 index returns: The approximate of the theoretical market portfolio returns.

The variables 3, 4, 5 are from the Federal Reserve Board H.15 and the data of 6 and 7 are from Yahoo Finance.

First we conduct a statistical check of the linearity between GS return and the market variables using the confidence band as constructed in Appendix 2. As shown in Fig. 54.2, except for some ignorable outsiders, the linear quantile regression line lies in the LLQR asymptotic confidence band.

On the other hand, there is nonlinearity between two individual assets X_i and X_j. To illustrate this, we regress X_j on M_t, and then take the residuals and regress them on X_i. Again the $X_{j,t}$ is GS daily return and X_i is C daily return. The result is shown in Fig. 54.3. The linear QR line (red) lies well outside the LLQR confidence band (magenta) when the C return is negative. The linear quantile regression line is fairly flat. The risk of using a linear model is obvious in this figure: The linear regression can "average out" the humped relation of the underlying structure (blue), and therefore imply a model risk in estimation.

Based on the results of the linearity tests above, we construct a PLM model:

$$X_{i,t} = \alpha_i + \gamma_i^{\mathrm{T}} M_{t-1} + \varepsilon_{i,t}, \tag{54.7}$$

$$X_{j,t} = \hat{\alpha}_{j|i} + \hat{\beta}_{j|i}^{\mathrm{T}} M_{t-1} + l_{j|i}\left(X_{i,t}\right) + \varepsilon_{j,t}, \tag{54.8}$$

where $X_{i,t}, X_{j,t}$ are asset returns of i, j firms. M_t is a vector of market variables at time t as introduced before. If i = S&P500, M_t is set to consist of the first 6 market variables only. Notice the variable $X_{i,t}$ enter the Eq. 54.8 nonlinearly.

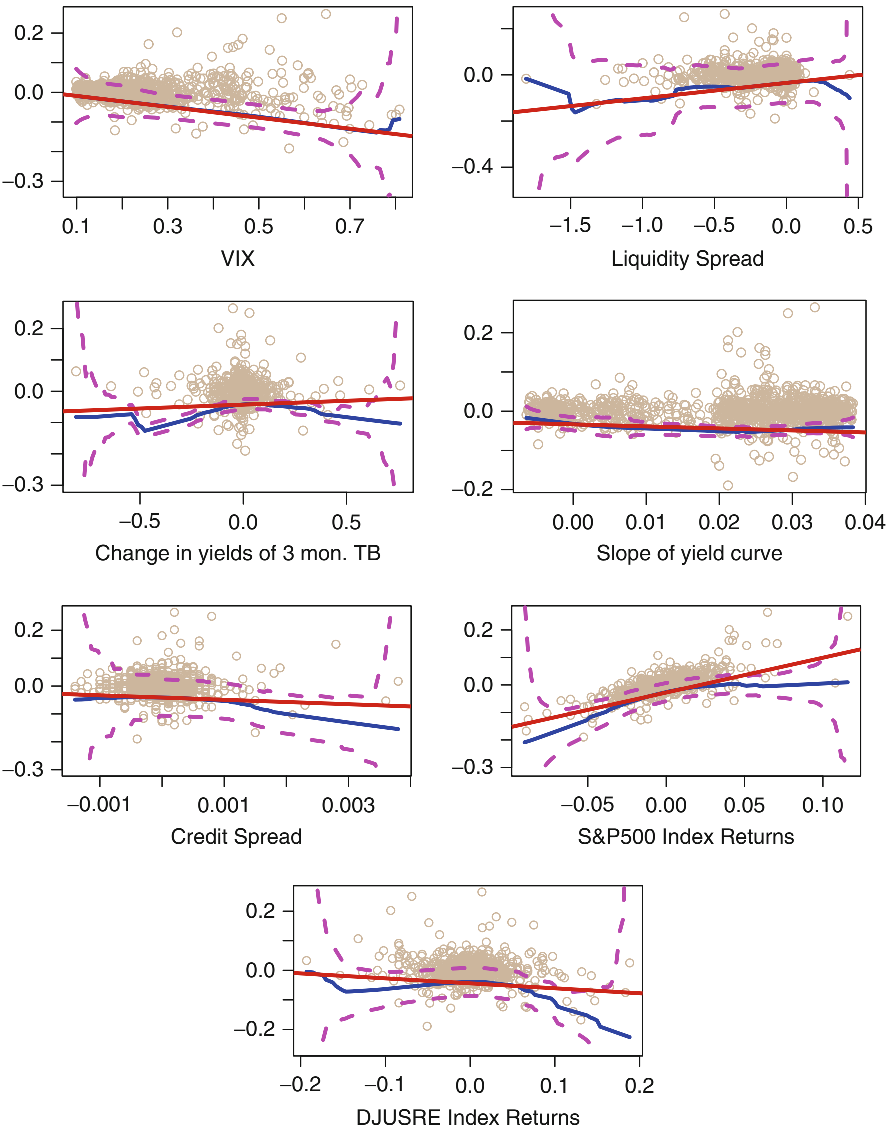


Fig. 54.2 The scatter plots of GS daily returns to the seven market variables with the LLQR curves. The bandwidths are selected by the method described in Appendix 1. The LLQR bandwidths are 0.1101, 0.1668, 0.2449, 0.0053, 0.0088, 0.0295 and 0.0569. The data period is from August 4, 2006, to August 4, 2011. $N = 1260$. $\tau = 0.05$

Applying the algorithm of Koenker and Bassett (1978) to Eq. 54.7 and the process described in Appendix 3 to Eq. 54.8, we get $\{\hat{\alpha}_i, \hat{\gamma}_i\}$ and $\left\{\hat{\alpha}_{j|i}, \hat{\beta}_i, \hat{l}(\cdot)\right\}$ with $F^{-1}_{\varepsilon_{i,t}}(\tau|M_{t-1}) = 0$ for Eq. 54.7 and $F^{-1}_{\varepsilon_{i,t}}\left(\tau|M_{t-1}, X_{i,t}\right) = 0$ for Eq. 54.8. Finally, we estimate the PLM $CoVaR_{j|i,t}$ by

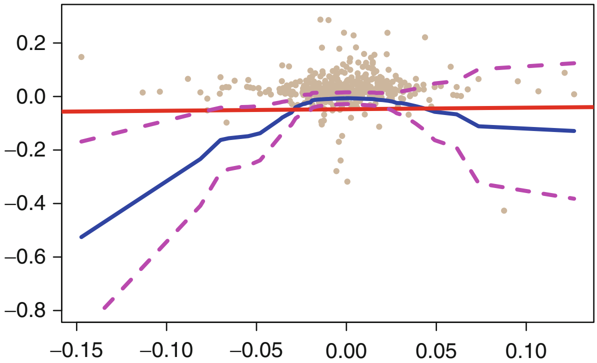


Fig. 54.3 The nonparametric part $\hat{l}_{GS|C}(\cdot)$ of the PLM estimation. The y-axis is the GS daily returns. The x-axis is the C daily returns. The *blue curve* is the LLQR quantile curve. The *red line* is the linear parametric quantile line. The *magenta dashed curves* are the asymptotic confidence band with significance level 0.05. The data is from June 25, 2008, to December 23, 2009. 378 observations. Bandwidth = 0.1255. $\tau = 0.05$

$$\widehat{VaR}_{i,t} = \hat{\alpha}_i + \hat{\gamma}_i^{\mathrm{T}} M_{t-1}, \tag{54.9}$$

$$\widehat{CoVaR}_{j|i,t}^{PLM} = \hat{\tilde{\alpha}}_{j|i} + \hat{\tilde{\beta}}_j^{\mathrm{T}} M_{t-1} + \hat{l}_{j|i}\left(\widehat{VaR}_{i,t}\right). \tag{54.10}$$

54.2.2 Backtesting

The goal of the backtesting procedure is to check if the VaR/CoVaR is accurate enough so that managerial decisions can be made based on them. The VaR forecast is a (one-sided) interval forecast. If the VaR algorithm is correct, then the violations should be unpredictable, after using all the past information. Formally, if we define the violation time series as

$$I_t = \begin{cases} 1, & \text{if } X_t < \widehat{VaR}_t^{\tau}; \\ 0, & \text{otherwise.} \end{cases}$$

where $\widehat{VaR}_t^{\tau}$ can be replaced by $\widehat{CoVaR}_t^{\tau}$ in the case of CoVaR. I_t should form a sequence of martingale difference.

There is a large literature on martingale difference tests. We adopt Ljung-Box test, Lobato test, and the CaViaR test. The Ljung-Box test and Lobato test aim to check whether the time series is autocorrelated. If the time series is autocorrelated, then we reject of course the hypothesis that the time series is a martingale difference.

Particularly, let $\hat{\rho}_k$ be the estimated autocorrelation of lag k of the sequence of violation $\{I_t\}$ and n be the length of the time series. The Ljung-Box test statistics is

$$\mathrm{LB}(m) = n(n+2)\sum_{k=1}^{m} \frac{\hat{\rho}_k^2}{n-k} \xrightarrow{\mathcal{L}} \chi(m), \tag{54.11}$$

as $n \to \infty$.

This test is too strong though in the sense that the asymptotic distribution is derived based on the i.i.d. assumption. A modified Box-Pierce test is proposed by Lobato et al. (2001), who also consider the test of no autocorrelation, but their test is more robust to the correlation of higher (greater than the first) moments. (Autocorrelation in higher moments does not contradict with the martingale difference hypothesis.) The test statistics is given by

$$\mathrm{L}(m) = n\sum_{k=1}^{m} \frac{\hat{\rho}_k^2}{\hat{v}_{kk}} \xrightarrow{\mathcal{L}} \chi(m),$$

as $n \to \infty$, where

$$\hat{v}_{kk} = \frac{\frac{1}{n}\sum_{i=1}^{n-k} (y_i - \overline{y})^2 (y_{i+k} - \overline{y})^2}{\left\{\frac{1}{N}\sum_{i=1}^{n} (y_i - \overline{y})^2\right\}^2}.$$

The CaViaR test, proposed by Berkowitz et al. (2011), is based on the idea that if the sequence of violation is a martingale difference, there ought to be no correlation between any function of the past variables and the current violation. One way to test this uncorrelatedness is through a linear model. The model is

$$I_t = \alpha + \beta_1 I_{t-1} + \beta_2 VaR_t + u_t,$$

where VaR_t can be replaced by $CoVaR_t$ in the case of conditional VaR. The residual u_t follows a Logistic distribution since I_t is binary. We get the estimates of the coefficients $\left(\hat{\beta}_1, \hat{\beta}_2\right)^{\mathrm{T}}$. Therefore, the null hypothesis is $\hat{\beta}_1 = \hat{\beta}_2 = 0$. This hypothesis can be tested by Wald's test.

We set $m = 1$ or 5 for the Ljung-Box and Lobato tests. For the CaViaR test, two data periods are considered separately. The first is the overall data from August 4, 2006, to August 4, 2011. The second is the data from August 4, 2008, to August 4, 2009, the period when the financial market reached its bottom. By separately testing the two periods, we can gain more insights into the PLM model.

54.2.3 Risk Contribution Measure

The risk contribution of one firm to the market is one of the top concerns among central bankers. The regulator can restrict the risky behaviors of the financial institution with high-risk contribution to the market, and reduce the institution's incentive to take more risk. AB propose the idea of ΔCoVaR, which is defined by

$$\Delta CoVaR^{\tau}_{j|i,t} = CoVaR^{\tau}_{j|i,t} - CoVaR^{0.5}_{j|i,t}. \tag{54.12}$$

where $\mathrm{CoVaR}^{\tau}_{j|i,t}$ is defined as in the introduction. j, i represent the financial system and an individual asset. $\tau = 0.5$ corresponds to the normal state of the individual asset i. This is essentially a sensitivity measure quantifying the effect to the financial system from the occurrence of a tail event of asset X_i.

In this study, we adopt a mathematically intuitive way to measure the marginal effect by searching the first order derivative of the quantile function. Because the spillover effect from stock market to stock market has already got much attention, it is important to investigate the risk contribution of a local market to the global stock market. The estimation is conducted as follows:

First, one estimates the following model nonparametrically:

$$X_{j,t} = f_j^{0.05}(X_t) + \varepsilon_j, \tag{54.13}$$

The quantile function $f_j^{0.05}(\cdot)$ is estimated with local linear quantile regression with $\tau = 0.05$, described with more details in Appendix 1. X_j is the weekly return of the stock index of an individual country and X is the weekly return of the global stock market.

Second, with $\hat{f}_j^{0.05}(\cdot)$, we compute the "marginal contribution of risk" (MCR) of institution j by

$$MCR_j^{\tau} = \left.\frac{\partial \hat{f}_j^{0.05}(x)}{\partial x}\right|_{x = \hat{F}_x^{-1}(\tau_k)}, \tag{54.14}$$

where $\hat{F}^{-1}(\tau_k)$ is a consistent estimator of the τ_k quantile of the global market return, and it can be estimated by regressing X_t on the time trend. We put $k = 1, 2$ with $\tau_1 = 0.5$ and $\tau_2 = 0.05$. The quantity Eq. 54.14 is similar to the MES proposed by Acharya et al. (2010) in the sense that the conditioned event belongs to the information set of the market return, but we reformulate it in the VaR framework instead of the expected shortfall framework.

There are some properties of the *MCR* to be described further. First, τ_k determines the condition of the global stock market. This allows us to explore the risk contribution from the index j to the global market, given different global market status. Second, the higher the value of MCR, the more risk factor j imposes on the market in terms of risk. Third, since the function $f_j^{0.05}(\cdot)$ is estimated by LLQR, the quantile curve is locally linear, and therefore, the local first order derivative is straightforward to compute.

We choose indices $j =$ S&P500, NIKKEI225, FTSE100, DAX30, CAC40, Hang Seng as the approximate of the market returns of each developed country or market. The global market is approximated by the MSCI World (developed countries) market index. The data is weekly from April 11, 2004, to April 11, 2011, and $\tau = 0.05$.

54.3 Results

54.3.1 CoVaR Estimation

The estimation results of VaR/CoVaR are shown in this section. We compute three types of VaR/CoVaR of GS, with a moving window size of 126 business days and $\tau = 0.05$.

First, the VaR of GS is estimated using linear quantile regression:

$$\widehat{VaR}_{GS,t} = \hat{\alpha}_{GS} + \hat{\gamma}_{GS}^{\mathrm{T}} M_{t-1}, \tag{54.15}$$

$M_t \in \mathbb{R}^7$ is introduced in Sect. 54.2.1.

Second, the CoVaR of GS given C returns is estimated:

$$\widehat{VaR}_{C,t} = \hat{\alpha}_C + \hat{\gamma}_C^{\mathrm{T}} M_{t-1}; \tag{54.16}$$

$$\widehat{CoVaR}^{AB}_{GS|C,t} = \hat{\alpha}_{GS|C} + \hat{\beta}_{GS|C} \widehat{VaR}_{C,t} + \hat{\gamma}_{GS|C}^{\mathrm{T}} M_{t-1}. \tag{54.17}$$

If the SP replaces C, the estimates are generated from

$$\widehat{VaR}_{SP,t} = \hat{\alpha}_{SP} + \hat{\gamma}_{SP}^{\mathrm{T}} \widetilde{M}_{t-1}; \tag{54.18}$$

$$\widehat{CoVaR}^{AB}_{GS|SP,t} = \hat{\alpha}_{GS|SP} + \hat{\beta}_{GS|SP} \widehat{VaR}_{SP,t} + \hat{\gamma}_{GS|SP}^{\mathrm{T}} \widetilde{M}_{t-1}, \tag{54.19}$$

where $\widetilde{M}_t \in \mathbb{R}^6$ is the vector of market variables without the market portfolio return.

Third, the PLM CoVaR is generated:

$$\widehat{VaR}_{C,t} = \hat{\alpha}_C + \hat{\gamma}_C^{\mathrm{T}} M_{t-1}; \tag{54.20}$$

$$\widehat{CoVaR}^{PLM}_{GS|C,t} = \hat{\hat{\alpha}}_{GS|C} + \hat{\hat{\beta}}_{GS|C}^{\mathrm{T}} M_{t-1} + \hat{l}_{GS|C}\left(\widehat{VaR}_{C,t}\right). \tag{54.21}$$

If SP replaces C:

$$\widehat{VaR}_{SP,t} = \hat{\alpha}_{SP} + \hat{\gamma}_{SP}^{\mathrm{T}} \widetilde{M}_{t-1}; \tag{54.22}$$

$$\widehat{CoVaR}^{PLM}_{GS|SP,t} = \hat{\hat{\alpha}}_{GS|SP} + \hat{\hat{\beta}}_{GS|SP}^{\mathrm{T}} \widetilde{M}_{t-1} + \hat{l}_{GS|SP}\left(\widehat{VaR}_{SP,t}\right). \tag{54.23}$$

The coefficients in Eqs. 54.15–54.20, and 54.22 are estimated from the linear quantile regression and those in Eqs. 54.21 and 54.23 are estimated from the method described in Appendix 3.

Figure 54.4 shows the $\widehat{VaR}_{GS,t}$ sequence. The VaR forecasts (red) seem to form a lower cover of the GS returns (blue). This suggests that the market variables M_t

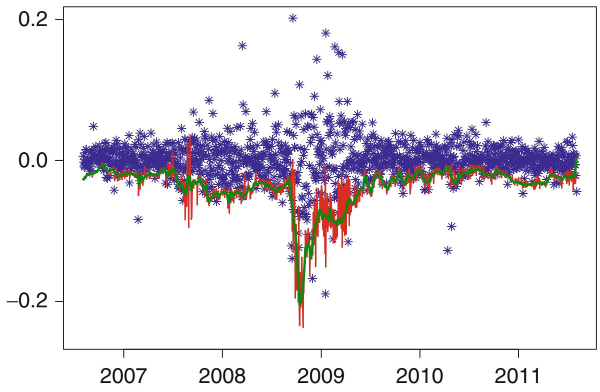


Fig. 54.4 The $\widehat{VaR}_{GS,t}$. The *red line* is the $\widehat{VaR}_{GS,t}$ and *blue stars* are daily returns of GS. The *dark green curve* is the median smoother of the $\widehat{VaR}_{GS,t}$ curve with $h = 2.75$. $\tau = 0.05$. The window size is 252 days

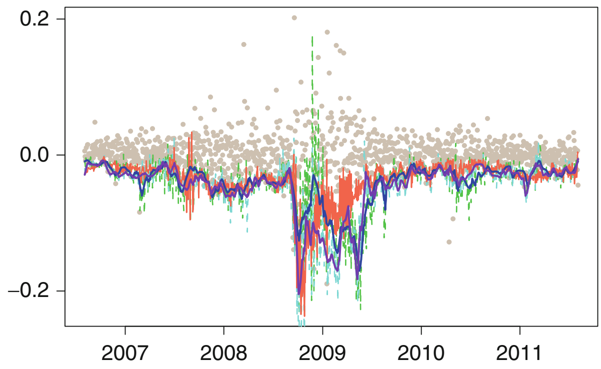


Fig. 54.5 The CoVaR of GS given the VaR of C. The *gray dots* mark the daily returns of GS. The *light green dashed curve* is the $\widehat{CoVaR}^{PLM}_{GS|C,t}$. The *dark blue curve* is the median LLQR smoother of the *light green dashed curve* with $h = 3.19$. The *cyan dashed curve* is the $\widehat{CoVaR}^{AB}_{GS|C,t}$. The *purple curve* is the median LLQR smoother of the *cyan dashed curve* with $h = 3.90$. The *red curve* is the $\widehat{VaR}_{GS,t}$. $\tau = 0.05$. The moving window size is 126 days

have some predictive power for the left tail quantile of the GS return distribution. Figure 54.5 shows the sequences $\widehat{CoVaR}^{AB}_{GS|SP,t}$ (cyan) and $\widehat{CoVaR}^{PLM}_{GS|C,t}$ (light green). As the time series of the estimates is too volatile, we smooth it further by the median LLQR. The two estimates are similar as the market state is stable, but during the period of financial instability (from mid-2008 to mid-2009), the two estimates have different behavior. The performances of these estimates are evaluated by backtesting procedure in Sect. 54.3.2.

Table 54.1 shows the summary statistics of the VaR/CoVaR estimates. The first three rows show the summary statistics of $\widehat{VaR}_{GS,t}$, $\widehat{VaR}_{C,t}$, and $\widehat{VaR}_{SP,t}$. The $\widehat{VaR}_{GS,t}$ has lower mean and higher standard deviation than the other two. Particularly during 2008–2009, the standard deviation of the GS VaR is twice as much as the other two. The mean and standard deviation of the $\widehat{VaR}_{C,t}$ and $\widehat{VaR}_{SP,t}$ are rather similar. The last four rows show the summary statistics of $\widehat{CoVaR}^{PLM}_{GS|C,t}$,

Table 54.1 VaR/CoVaR summary statistics. The overall period is from August 4, 2006, to August 4, 2011. The crisis period is from August 4, 2008, to August 4, 2009. The numbers in the table are scaled up by 10^2

	mean-overall	sd-overall	mean-crisis	sd-crisis
$\widehat{VaR}_{GS,t}$	−3.66	3.08	−7.43	4.76
$\widehat{VaR}_{C,t}$	−2.63	1.67	−4.62	2.25
$\widehat{VaR}_{SP,t}$	−2.09	1.57	−3.88	2.24
$\widehat{CoVaR}^{PLM}_{GS\|C,t}$	−4.26	3.84	−8.79	5.97
$\widehat{CoVaR}^{AB}_{GS\|C,t}$	−4.60	4.30	−10.36	6.32
$\widehat{CoVaR}^{PLM}_{GS\|SP,t}$	−3.86	3.30	−8.20	4.69
$\widehat{CoVaR}^{AB}_{GS\|SP,t}$	−5.81	4.56	−12.65	5.56

$\widehat{CoVaR}^{AB}_{GS|C,t}$, $\widehat{CoVaR}^{PLM}_{GS|SP,t}$, and $\widehat{CoVaR}^{AB}_{GS|SP,t}$. This shows that the CoVaR obtaining from the AB model has smaller mean and greater standard deviation than the CoVaR obtaining from PLM model.

Figure 54.6 shows the bandwidth sequence of the nonparametric part of the PLM estimation. The bandwidth varies with time. Before mid-2007, the bandwidth sequence is stably jumping around 0.2. After that the sequence becomes very volatile. This may have something to do with the rising systemic risk.

54.3.2 Backtesting

For the evaluation of the CoVaR models, we resort to the backtesting procedure described in Sect. 54.2.2. In order to perform the backtesting procedure, the sequences $\{I_t\}$ (defined in Sect. 54.2.2) have to be computed for all VaR/CoVaR estimates. Figure 54.7 shows the timings of the violations $\{t : I_t = 1\}$ of $\widehat{CoVaR}^{PLM}_{GS|C,t}$, $\widehat{CoVaR}^{AB}_{GS|C,t}$ and $\widehat{VaR}_{GS,t}$. This figure shows the total number of violations of PLM CoVaR and CoVaR is similar, while $\widehat{VaR}_{GS,t}$ has more violations than the both. The $\widehat{VaR}_{GS,t}$ has a few clusters of violations in both financial stable and unstable periods. This may result from the failure $\widehat{VaR}_{GS,t}$ to adapt for the negative shocks. The violations of $\widehat{CoVaR}^{PLM}_{GS|C,t}$ are more evenly distributed. The violations of $\widehat{CoVaR}^{AB}_{GS|C,t}$ have large clusters during financially stable period, while the violation during financial crisis period is meager. This contrast suggests that $\widehat{CoVaR}^{AB}_{GS|C,t}$ tend to overreact, as it is slack during the stable period but is too tight during the unstable period.

Figure 54.8 shows the timings of the violations $\{t : I_t = 1\}$ of $\widehat{CoVaR}^{PLM}_{GS|SP,t}$, $\widehat{CoVaR}^{AB}_{GS|SP,t}$, and $\widehat{VaR}_{GS,t}$. The overall number of violations of $\widehat{CoVaR}^{PLM}_{GS|SP,t}$ is more than that of $\widehat{VaR}_{GS,t}$, and it has many clusters. $\widehat{CoVaR}^{PLM}_{GS|SP,t}$ behaves differently from $\widehat{CoVaR}^{PLM}_{GS|C,t}$. The SP may not be more informative than C, though

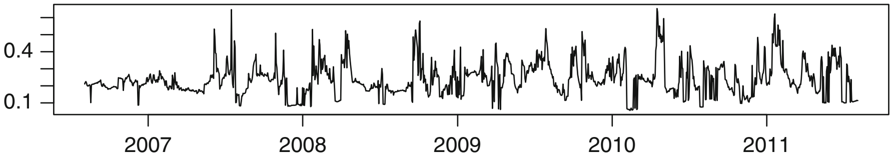


Fig. 54.6 LLQR bandwidth in the daily estimation of $\widehat{CoVaR}^{PLM}_{GS|C,t}$. The average bandwidth is 0.24

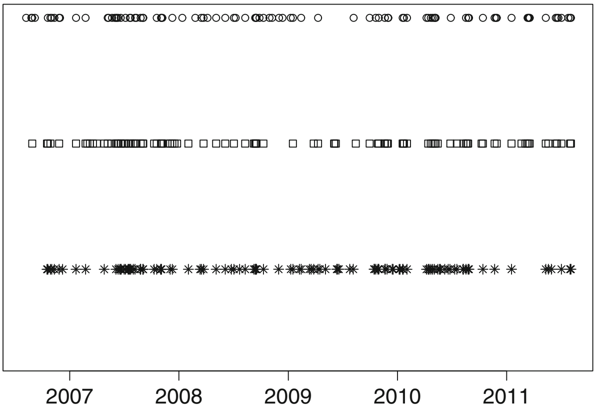


Fig. 54.7 The timings of violations $\{t : I_t = 1\}$. The *top circles* are the violations of the $\widehat{CoVaR}^{PLM}_{GS|C,t}$, totally 95 violations. The *middle squares* are the violations of $\widehat{CoVaR}^{AB}_{GS|C,t}$, totally 98 violations. The *bottom stars* are the violations of $\widehat{VaR}_{GS,t}$, totally 109 violations. Overall data $N = 1260$

the efficient market hypothesis suggests so. The violation of $\widehat{CoVaR}^{AB}_{GS|SP,t}$ is fewer than the other two measures, and the clustering is not significant.

The backtesting procedure is performed separately for each sequence of $\{I_t\}$. The null hypothesis is that each sequence of $\{I_t\}$ forms a series of martingale difference. Six different tests are applied for each $\{I_t\}$: Ljung-Box tests with lags 1 and 5, Lobato test with lags 1 and 5, and finally the CaViaR test with two data periods: overall and crisis period.

The result is shown in Table 54.2. First, in Panel 1 of Table 54.2, the $\widehat{VaR}_{GS,t}$ is rejected by the LB(5) test and the two CaViaR tests. This shows that a linear quantile regression on the seven market variables may not give accurate estimates, in the sense that the violation $\{I_t\}$ of $\widehat{VaR}_{GS,t}$ does not form a martingale sequence. Next we turn to the $\widehat{CoVaR}^{AB}_{GS|SP,t}$ and $\widehat{CoVaR}^{PLM}_{GS|SP,t}$. In Panel 2, the low p-values of the two CaViaR tests show that both the AB model and PLM model conditioned on SP are rejected, though the p-value of the AB model almost reaches the 5 % significant level. In particular, the $\widehat{CoVaR}^{PLM}_{GS|SP,t}$ is rejected by the L(5) and LB (5) tests. Both the parametric and semiparametric models fail with this choice of variable. This suggests that the market return does not provide enough information in risk measurement.

We therefore need more informative variables. Panel 3 of Table 54.2 illustrates this by using C daily returns, which may contain information not revealed in the

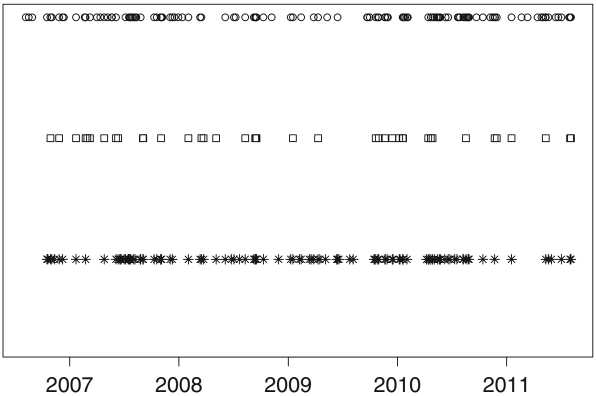


Fig. 54.8 The timings of violations $\{t : I_t = 1\}$. The *top circles* are the violations of $\widehat{CoVaR}^{PLM}_{GS|SP,t}$, totally 123 violations. The *middle squares* are the violations of $\widehat{CoVaR}^{AB}_{GS|SP,t}$, totally 39 violations. The *bottom stars* are the violations of $\widehat{VaR}_{GS,t}$, totally 109 violations. Overall data $N = 1{,}260$

Table 54.2 Goldman Sachs VaR/CoVaR backtesting p-values. The overall period is from August 4, 2006, to August 4, 2011. The crisis period is from August 4, 2008, to August 4, 2009. LB(1) and LB(5) are the Ljung-Box tests of lags 1 and 5. L(1) and L(5) are the Lobato tests of lags 1 and 5. CaViaR-overall and CaViaR-crisis are two CaViaR tests described in Sect. 2.2 applied on the two data periods

Measure	LB(1)	LB(5)	L(1)	L(5)	CaViaR-overall	CaViaR-crisis
Panel 1						
$\widehat{VaR}_{GS,t}$	0.3449	0.0253^{*}	0.3931	0.1310	$1.265 \times 10^{-6***}$	0.0024^{**}
Panel 2						
$\widehat{CoVaR}^{AB}_{GS\|SP,t}$	0.0869	0.2059	0.2684	0.6586	$8.716 \times 10^{-7***}$	0.0424^{*}
$\widehat{CoVaR}^{PLM}_{GS\|SP,t}$	0.0518	0.0006^{***}	0.0999	0.0117^{*}	$2.2 \times 10^{-16***}$	0.0019^{**}
Panel 3						
$\widehat{CoVaR}^{AB}_{GS\|C,t}$	0.0489^{*}	0.2143	0.1201	0.4335	$3.378 \times 10^{-9***}$	0.0001^{***}
$\widehat{CoVaR}^{PLM}_{GS\|C,t}$	0.8109	0.0251^{*}	0.8162	0.2306	$2.946 \times 10^{-9***}$	0.0535

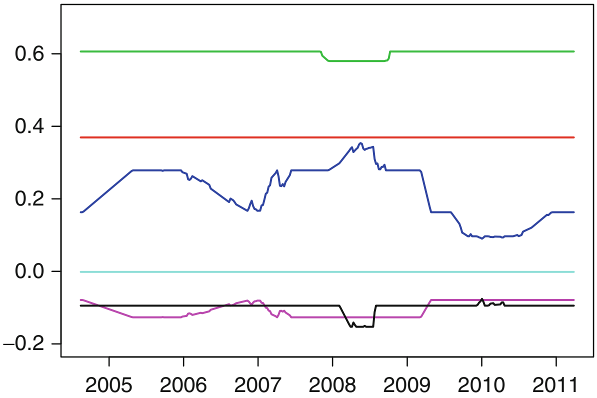
*, ** and *** denote significance at the 5 %, 1 % and 0.1 % levels

market and improve the performance of the estimates. The $\widehat{CoVaR}^{AB}_{GS|C,t}$ is rejected by the two CaViaR tests and the LB(1) test with 0.1 % and 5 % significant level. However, $\widehat{CoVaR}^{PLM}_{GS|C,t}$ is not rejected by the CaViaR-crisis test. This implies that the nonparametric part in the PLM model captures the nonlinear effect of C returns to GS returns, which can lead to better risk-measuring performance.

54.3.3 Global Risk Contribution

In this section, we present the *MCR* (defined in Sect. 54.2.3), which measures the marginal risk contribution of risk factors. We choose $\tau_1 = 0.5$, associated to the

Fig. 54.9 The $MCR_j^{\tau_1}$, $\tau = 0.5$. j:CAC, FTSE, DAX, Hang Seng, S&P500 and NIKKEI225. The global market return is approximated by MSCI World



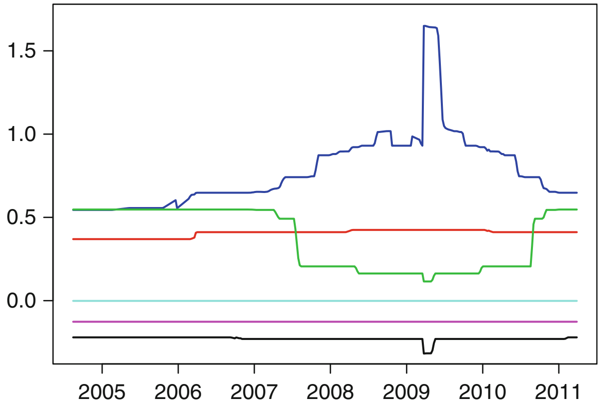


Fig. 54.10 The $MCR_j^{\tau_2}$, $\tau = 0.05$. j:CAC, FTSE, DAX, Hang Seng, S&P500 and NIKKEI225. The global market return is approximated by MSCI World

normal (median) state and $\tau_2 = 0.05$, associated to an negative extreme state. Figure 54.9 shows the $MCR_j^{\tau_1}$ from local markets j to the global market. When the MSCI World is at its normal state, the Hang Seng index in normal times contributes the most risk to the MSCI World at all times. The NIKKEI225 places second; the contribution from S&P500 varies most with the time; the risk contribution from DAX30 is nearly zero. The contributions from CAC40 and FTSE100 are negative.

Assuming that the MSCI World is at its bad state ($\tau_2 = 0.05$), the $MCR_j^{\tau_2}$ differs from $MCR_j^{\tau_1}$, see Fig. 54.10. One sees that the S&P500 imposes more pressure on the world economy than the other countries, especially during the financial crisis of 2008 and 2009. The contribution from Hang Seng is no longer of the same significance. The three European markets are relatively stable.

This analysis suggests that the risk contribution from individual stock market varies a lot with the state of global economy.

54.4 Conclusion

In this study, we construct a PLM model for the CoVaR, and we compare it to the AB model by backtesting. Results show that PLM CoVaR is preferable, especially during a crisis period. The study of the MCR reveals the fact that the risk from each country can vary with the state of global economy.

As an illustration, we only study the Goldman Sachs conditional VaR with Citigroup and S&P500 as conditioned risk sources. In practice, we need to choose variables. In Hautsch et al. (2011), the Least Absolute Shrinkage and Selection Operator (LASSO) techniques are used to determine the most relevant systemic risk sources from a pool of financial institutions. A VAR (Vector Autoregression) model may be also suitable for capturing the asset dynamics, but the estimation may be more involved. We may include other firm-specific variables such as corporate bond yields as these variables can bear other information which is not included in the stock returns or stock indices.

Appendix 1: Local Linear Quantile Regression (LLQR)

Let $\{(X_i, Y_i)\}_{i=1}^n \subset \mathbb{R}^2$ be independently and identically distributed (i.i.d.) bivariate random variables. Denote by $F_{Y|x}(u)$ the conditional cumulative distribution function (cdf) and $l(x) = F_{Y|x}^{-1}(\tau)$ the conditional quantile curve to level τ, given observations $\{(x_i, y_i)\}_{i=1}^n$, one may write this as

$$y_i = l(x_i) + \varepsilon_i,$$

with $F_{\varepsilon|x}^{-1}(\tau) = 0$. A locally linear kernel quantile estimator (LLQR) is estimated as $\hat{l}(x_0) = \hat{a}_0$ from:

$$\left(\hat{a}_0, \hat{b}_0\right) = \underset{\{a_0, b_0\}}{\text{argmin}} \sum_{i=1}^{n} K\left(\frac{x_i - x_0}{h}\right) \tag{54.24}$$

$$\rho_\tau\{y_i - a_0 - b_0(x_i - x_0)\}, \tag{54.25}$$

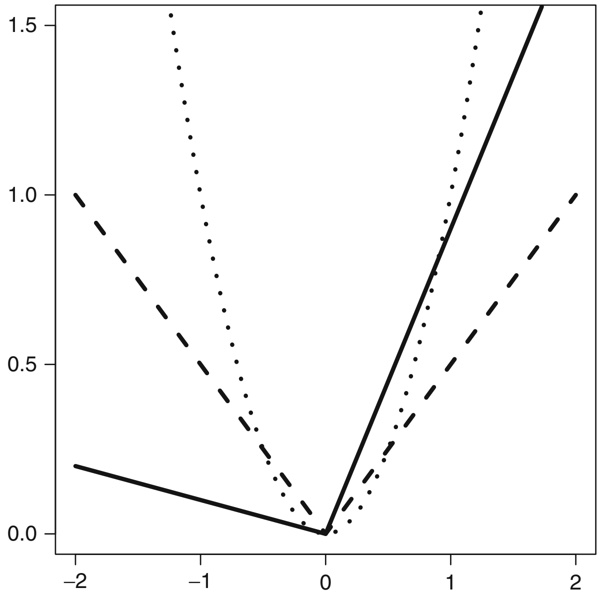
where h is the bandwidth, $K(\cdot)$ is a kernel, and $\rho_\tau(\cdot)$ is the check function given by

$$\rho_\tau(u) = \left(\tau - 1_{\{u<0\}}\right)u \tag{54.26}$$

Figure 54.11 illustrates the check functions. Different loss functions give different estimates. u^2 corresponds to the conditional mean. $\rho_\tau(u)$ corresponds to the conditional τth quantile.

It is shown by Fan et al. (1994) that the locally linear kernel estimator is asymptotically efficient in a minimax sense. It also possesses good finite sampling property which is adaptive to a variety of empirical density $g(x)$ and has good boundary property.

Fig. 54.11 This figure presents the check function. The *dotted line* is u^2. The *dashed and solid lines* are check functions $\rho_\tau(u)$ with $\tau = 0.5$ and 0.9 respectively



Next, we describe the method to compute the bandwidths. The approach used here follows Yu and Jones (1998). The bandwidth is chosen by

$$h_\tau = h_{mean}\left[\tau(1-\tau)\varphi\left\{\Phi^{-1}(\tau)\right\}^{-2}\right]^{1/5}, \tag{54.27}$$

where h_{mean} is the locally linear mean regression bandwidth, which can be computed by the algorithm described in Ruppert and Wand (1995) or Ruppert et al. (1995). $\varphi(\cdot)$ and $\Phi(\cdot)$ are the pdf and cdf of the standard normal distribution. Since we discuss the case for VaR, τ is usually small. h_τ needs to be enlarged to allow for more smoothing (usually taking $1.5h_\tau$ or $2h_\tau$).

The approach is acceptable but not so flexible, because it is based on assuming the quantile functions are parallel. A more flexible approach was developed by Spokoiny et al. (2011). In order to stabilize the bandwidth choice, we first regress y_i on the rank of the corresponding x_i and then rescale the resulted estimated values to the original x space. Carroll and Hardle (1989) show that this local bandwidth estimator and the global bandwidth estimator are asymptotically equivalent.

Appendix 2: Confidence Band for Nonparametric Quantile Estimator

The uniform confidence band of the quantile estimator is based on the Theorem 2.2 and Corollary 2.1 presented in Hardle and Song (2010). The details are as follows.

Let $\{(X_i, Y_i)\}_{i=1}^{n}$ be as in Appendix 1. Define $K_h(u) = h^{-1} K(u/h)$ and similar to Eq. 54.25 let $l_n(x)$ and $l(x)$ are zeros (w.r.t. θ) of the functions:

$$\widetilde{H}_n(\theta, x) \stackrel{\text{def}}{=} n^{-1} \sum_{i=1}^{n} K_h(x - X_i)\rho_\tau(Y_i - \theta);$$

$$\widetilde{H}(\theta, x) \stackrel{\text{def}}{=} \int_{\mathbb{R}} f(x, y)\rho_\tau(y - \theta)dy,$$

where $\rho_\tau(\cdot)$ is the check function defined as Eq. 54.26.

Theorem 1 *Let* $h = n^{-\delta}$, $\frac{1}{5} < \delta < \frac{1}{3}$, $\lambda(K) = \int_{-A}^{A} K^2(u)du$, *where* $K(\cdot)$ *is supported on* $[-A, A]$. $J = [0, 1]$. *Define* $c_1(K) = \{K^2(A) + K^2(-A)\}/2\lambda(K)$, $c_2(K) = \int_{-A}^{A}\{K'(u)\}^2 du/2\lambda(K)$ *and*

$$d_n = \begin{cases} (2\delta \log n)^{1/2} + (2\delta \log n)^{-1/2}\left[\frac{\log\{c_1(K)\}}{\pi^{1/2}}\right\} + \frac{1}{2}\{\log\delta + \log\log n\}], \\ \text{if } \ c_1(K) > 0; \\ (2\delta \log n)^{1/2} + (2\delta \log n)^{-1/2 \frac{\log\{c_2(K)\}}{2\pi}}\}, \quad \text{otherwise.} \end{cases}$$

Then

$$\mathrm{P}\left[(2\delta \log n)^{1/2}\left\{\sup_{x \in J} r(x)\frac{|l_n(x) - l(x)|}{\lambda(K)^{1/2}} - d_n\right\} < z\right] \\ \to \exp\{-2\exp(-z)\},$$

as $n \to \infty$, *with*

$$r(x) = (nh)^{1/2} f\{l(x)|x\}\{f_X(x)/\tau(1 - \tau)\}^{1/2},$$

where $f_X(\cdot)$ *is the marginal pdf for X and* $f(\cdot|x)$ *is the conditional pdf of Y on* $X = x$.

The corollary followed by the theorem explicitly indicates how a uniform confidence interval can be constructed.

Corollary 1 *An approximate* $(1 - \alpha) \times 100\ \%$ *confidence band is*

$$l_n \pm (nh)^{-1/2}\left\{\tau(1 - \tau)\lambda(K)/\hat{f}_X(t)\right\}^{1/2} \hat{f}^{-1}\{l(t)|t\} \times \left\{d_n + c(\alpha)(2\delta \log n)^{-1/2}\right\},$$

where $c(\alpha) = \log 2 - \log|\log(1 - \alpha)|$ *and* $\hat{f}_X(t), \hat{f}\{l(t)|t\}$ *are consistent estimates for* $f_X(t), f\{l(t)|t\}$.

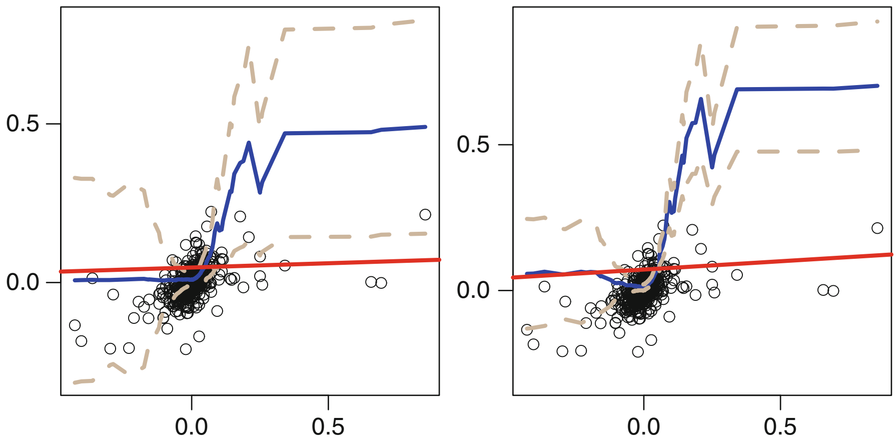


Fig. 54.12 GS and C weekly returns 0.90(*left*) and 0.95(*right*) quantile functions. The y-axis is GS daily returns and the x-axis is the C daily returns. The *blue curves* are the LLQR curves (see Appendix 1). The LLQR bandwidths are 0.0942 and 0.1026. The *red lines* are the linear parametric quantile regression line. The *antique white curves* are the asymptotic confidence band (see Appendix 2) with significance level 0.05. $N = 546$

Figure 54.1 is done by the techniques introduced in Appendices 1 and 2. Another illustration with right tail quantiles is in Fig. 54.12. We plot the LLQR curve for 0.9 and 0.95 quantile. Both the two linear quantile regression lines lie outside the LLQR confidence band as the Citigroup returns are positive.

Appendix 3: PLM Model Estimation

For the PLM estimation, we adopt the algorithm described in Song et al. (2012). Given data $\{(X_t, Y_n)\}_{n=1}^{T}$ bivariate and $\{M_t\}_{n=1}^{T}$ multivariate random variables. The PLM is:

$$Y_t = \alpha + \beta^{\mathrm{T}} M_{t-1} + l(X_t) + \varepsilon_t.$$

Let a_n denote an increasing sequence of positive integers and set $b_n = a_n^{-1}$. For given n, dividing the interval [0, 1] into a_n subintervals I_{nt}, $t = 1, \ldots, a_n$ with equal length b_n. On each I_{nt}, $l(\cdot)$ can approximately be taken as a constant.

The PLM estimation procedure is:

1. Inside each partition I_{nt}, a linear quantile regression is performed to get $\hat{\beta}_i$, then their weighted mean gives $\hat{\beta}$. Formally, let $\rho_\tau(\cdot)$ be the check function defined as Eq. 54.26, $l_1, \ldots, l_{a_n}$ are constants,

$$\hat{\beta} = \underset{\beta}{\operatorname{argmin}} \min_{l_1, \ldots, l_{a_n}} \sum_{t=1}^{n} \rho_\tau \left\{ X_{j,t} - \alpha - \beta^{\mathrm{T}} M_{t-1} - \sum_{m=1}^{a_n} l_m 1\left(X_{i,t} \in I_{nt}\right) \right\}$$

2. Computing the LLQR nonparametric quantile estimates of $l(\cdot)$ as outlined in Appendix 1 from $\left\{ \left(X_{i,t}, X_{j,t} - \hat{\alpha} - \hat{\beta}^{\mathrm{T}} M_{t-1} \right) \right\}_{t=1}^{n}$.

References

Acharya, V. V., Pedersen, L. H., Philippon, T., & Richardson, M. (2010). *Measuring systemic risk* (Working paper 10-02). Federal Reserve Bank of Cleveland.

Adams, Z., Füss, R., & Gropp, R. (2012). Spillover effects among financial institutions: A state-dependent sensitivity value-at-risk approach, SAFE Working Paper Series, No. 20, http://dx.doi.org/10.2139/ssrn.2267853.

Adrian, T., & Brunnermeier, M. K. (2011). *CoVaR* (Staff reports 348). Federal Reserve Bank of New York.

Bae, K.-H., Karolyi, G. A., & Stulz, R. M. (2003). A new approach to measuring financial contagion. *Review of Financial Studies, 16*(3), 717–763.

Berkowitz, J., Christoffersen, P., & Pelletier, D. (2011). *Evaluating value-at-risk models with desk-level data management science, 57*, 2213–2227.

Brownlees, C. T., & Engle, R. F. (2012). Volatility, correlation and tails for systemic risk measurement. Available at SSRN: http://ssrn.com/abstract=1611229 or http://dx.doi.org/10.2139/ssrn.1611229.

Brunnermeier, M., & Pedersen, L. H. (2008). Market liquidity and funding liquidity. *Review of Financial Studies, 22*(2201), 2238.

Cai, Z., & Wang, X. (2008). Nonparametric estimation of conditional VaR and expected shortfall. *Journal of Econometrics, 147*(120), 130.

Carroll, R., & Hardle, W. (1989). Symmetrized nearest neighbor regression estimates. *Statistics & Probability Letters, 7*(315), 318.

Chernozhukov, V., & Umantsev, L. (2001). Conditional value-at-risk: Aspects of modeling and estimation. *Empirical Economics, 26*(271), 292.

Engle, R., & Manganelli, S. (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business & Economic Statistics, 22*(367), 381.

Fan, J., Hu, T.-C., & Truong, Y. K. (1994). Robust nonparametric function estimation. *Scandinavian Journal of Statistics, 21*, 433–446.

Härdle, W., & Song, S. (2010). Confidence bands in quantile regression. *Econometric Theory, 26*, 1180–1200.

Härdle, W., Liang, H., & Gao, J. (2000). *Partially linear models*. Heidelberg: Physica-Verlag.

Härdle, W., Muller, M., Sperlicli, S., & Werwatz, A. (2004). *Nonparametric and semiparametric models*. Berlin: Springer.

Hautsch, N., Schaumburg, J., & Schienle, M. (2011). *Financial network systemic risk contributions* (Discussion paper 2011-072). CRC 649, Humboldt-Universitat zu Berlin.

Huang, X., Zhou, H., & Zhu, H. (2012). *Systemic risk contributions Journal of Financial Services Research, 42*, 55–83. Springer US.

Koenker, R., & Bassett, G. S. (1978). Regression quantiles. *Econometrica, 46*, 33.

Kuan, C.-M., Yeh, J.-H., & Hsu, Y.-C. (2009). Assessing value at risk with CARE, the conditional autoregressive expectile models. *Journal of Econometrics, 150*, 261–270.

Lobato, I., Nankervis, J. C., & Savin, N. (2001). Testing for autocorrelation using a modified Box-Pierce Q test. *International Economic Review, 42*(1), 187–205.

Ruppert, D., & Wand, M. P. (1995). Multivariate locally weighted least, squares regression. *Annals of Statistics, 23*(1346), 1370.

Ruppert, D., Sheatlier, S. J., & Wand, M. P. (1995). An effective bandwidth selector for local least squares regression. *Journal of the American Statistical Association, 90*(1257), 1270.

Schaumburg, J. (2011). *Predicting extreme VaR: Nonparametric quantile regression with refinements from extreme value theory* (Discussion Paper 2010-009). CRC 649, Humboldt-Universitat zu Berlin.

Song, S., Ritov, Y., & Härdle, W. (2012). Partial linear quantile regression and bootstrap confidence bands. *Journal of Multivariate Analysis, 107*, 244–262.

Spokoiny, V., Wang, W., & Härdle, W. (2013). Local quantile regression. *Journal of Statistical Planning and Inference, 143*(7), 1109–1129. doi:10.1016/j.jspi.2013.03.008, with discussions.

Taylor, J. W. (2008). Using exponentially weighted quantile regression to estimate value at risk and expected shortfall. *Journal of Financial Econometrics, 6*, 382–406.

Yu, K., & Jones, M. C. (1998). Local linear quantile regression. *Journal of the American Statistical Association, 98*(228), 237.

Strike Prices of Options for Overconfident Executives

55

Oded Palmon and Itzhak Venezia

Contents

Abstract

We explore via simulations the impacts of managerial overconfidence on the optimal strike prices of executive incentive options. Although it has been shown that, optimally, managerial incentive options should be awarded in-the-money, in practice most firms award them at-the-money. We show that the optimal strike prices of options granted to overconfident executive are directly related to their overconfidence level and that this bias brings the optimal strike prices closer to the institutionally prevalent at-the-money prices. Our results thus support the viability of the common practice of awarding managers with at-the-money incentive options. We also show that overoptimistic CEOs receive lower compensation than their realistic counterparts and that the stockholders benefit from

O. Palmon (✉)
Department of Finance and Economics, Rutgers Business School – Newark and New Brunswick, Piscataway, NJ, USA
e-mail: palmon@business.rutgers.edu; palmon@rbs.rutgers.edu

I. Venezia
School of Business, The Hebrew University, Jerusalem, Israel

Bocconi University, Milan, Italy
e-mail: msvenez@mscc.huji.ac.il

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_55,

their managers bias. The combined welfare of the firm's stakeholders is, however, positively related to managerial overconfidence.

The Monte Carlo simulation procedure described in Sect. 55.3 uses a Mathematica program to find the optimal effort by managers and the optimal (for stockholders) contract parameters. An expanded discussion of the simulations, including the choice of the functional forms and the calibration of the parameters, is provided in Appendix 1.

Keywords

Overconfidence • Managerial effort • Incentive options • Strike price • Simulations • Behavioral finance • Executive compensation schemes • Mathematica optimization • Risk aversion • Effort aversion

55.1 Introduction

The optimal structure of executive compensation has intrigued academic researchers as well as practitioners for a long time. Most principal-agent models dealing with this issue yield rather complex payment schedules, making it quite challenging to test their predictions. In practice there is a widespread use of simple compensation schemes such as linear or piecewise linear (stock option) contracts. An important question that arises in this case is what are the optimal parameters for these simple schemes? In particular what are the optimal strike prices for incentive option schemes? Unfortunately, this important issue has received only little attention. Institutional and tax factors could be to blame for this neglect. Before the 2006 changes in the US tax rules, the "intrinsic value" of executive options was taxed, and this discouraged firms from granting their executives in-the-money options. Granting out-of-the-money options seemed unfair and there is no empirical or theoretical evidence for advantages to such practice (Mahajan 2002). Indeed, only a very small fraction of firms used such strike prices.[1] The virtual monopoly of at-the-money strike prices, their institutional appeal, or some other unknown factors might have discouraged academics from studying the merits and demerits of this practice.

In their landmark paper, Hall and Murphy (2000, 2002) attribute the pervasiveness of granting at-the-money options to their property of being the most sensitive to changes in the stock price. Palmon et al. (2008), however, have shown that issuing the most sensitive options is not necessarily optimal when managers are risk and effort averse.[2] Within a model explicitly considering the choice of the contract parameters by stockholders and the resulting effort chosen by risk-averse and effort-averse managers, they show for a wide range of parameters that well describe

[1]According to Mahajan (2002), less than 1 % of firms used out-of-the-money strike prices. Furthermore, in his study firms did not benefit from awarding such options to their managers.

[2]Hall and Murphy (2000) did not show that at-the-money strike prices are optimal, just that they possess the highest sensitivity to stock prices. They did not assume effort aversion by managers either.

managers' risk and effort aversion, that in-the-money options provide are optimal. Such options provide managers a better risk-return trade-off and ultimately constitute a better form of compensation than either out-of-the-money or at-the-money options.[3] Palmon et al. further argue that the asymmetric tax treatment of options under the old (prior to 2006) tax system, which penalized the issuance of in-the-money options, may have driven firms to use at-the-money options.

Whereas most studies of the issue of optimal incentive contracts assumed that managers as well as stockholders are rational, there exists extensive literature that documents that managers often are overconfident. Of the few studies that explore the effect of cognitive biases on managerial compensation, none however explores the effect of overconfidence on the optimal strike prices for the incentive options. Gervais et al. (2011) investigate the optimal form of managerial compensation under overconfidence but define overconfidence in the sense of too-high-precision-of-estimates (calibration), and the managers in their model exert effort to obtain better information on the investment parameters. There is an abundant literature however that indicates the pervasiveness of overconfidence in the optimism or "better than average" sense rather than in the calibration interpretation (see, e.g., Malmendier and Tate 2005a, b, 2008; Roll 1986; Suntheim 2012).[4] Oyer and Schaefer (2005) and Bergman and Jenter (2007) also consider the effect of optimism, and other sentiments, on managerial compensation, but they do not consider the effect of these sentiments on the optimal strike prices or on the managers' effort.

In this paper we investigate the hitherto unexplored question of the effect of overconfidence on the optimal strike prices for risk-averse and effort-averse managers. We show that overconfidence leads to higher optimal strike prices of managerial incentive schemes, and that awarding overconfident CEOs at-the-money options mitigates the stockholders' vs. managers' agency problem, leading to higher managers' productivity. Our results thus provide support for the viability of the ubiquitous yet seemingly unoptimal practice of awarding CEOs with at-the-money incentive options.

Whereas the main focus of the paper is the interaction between overconfidence and the strike prices of managerial incentive options, it also sheds light on the effect of overconfidence on the firm's stakeholders (stockholders and managers). We predict, as empirically shown by Otto (2011), that overoptimistic CEOs receive lower compensation than their realistic counterparts. However, the stockholders benefit from their managers bias since they pay less and enjoy the productivity of the higher effort the overconfident manager exerts. We construct a measure of the combined welfare of managers and stockholders and demonstrate that it is

[3]Dittman et al. (2010) found that for a range of parameterizations, a principal-agent model with loss-averse agents generates convex compensation contract but did not investigate the parameters of the options to be used in the compensation package. Recently, however, Dittman and Yu (2011) found that in-the-money options are optimal.

[4]Glaser and Weber (2007) note that only overconfidence in the better than average sense affects trading.

positively related to managerial overconfidence, a result helping explain the persistence of this bias.[5]

The paper is constructed as follows. In Sect. 55.2 we present the model. In Sect. 55.3 we explain the simulation method, and in Sect. 55.4 we present the simulations' results. Section 55.5 concludes.

55.2 Overconfidence and the Optimal Exercise Prices of Executive Incentive Options

We consider a one-period Holmstrom (1979)-type model where a risk-neutral firm employs an overconfident, risk-averse, and effort-averse manager.[6] The cash flows, X, of the firm depend on the manager's effort and on exogenous stochastic factors. The manager is assumed to provide some effort which is the minimum necessary to run the firm and hence may be considered observable, but can provide also unobservable extra effort. The more extra effort the manager exerts, the higher will be the expected cash flows. Because stockholders cannot observe managers' extra effort, managerial compensation may depend on the firm's cash flows (which depend on effort), but cannot be determined directly based on extra effort.

We assume that the cash flows of the firm, X, are lognormally distributed with the following distribution function:

$$f(X) = \exp\left\{-0.5\{[\log(X) - \mu(Y)]/\sigma\}^2\right\}/\left(X\sigma\sqrt{2\pi}\right) \tag{55.1}$$

where Y denotes the managerial extra effort (a managerial choice variable) and $\mu(Y)$ and σ denote, respectively, the mean and the standard deviation of the underlying normal distribution of the natural logarithm of X. We assume that managerial effort increases cash flows and that overconfident managers overestimate the impact of their effort on cash flows. Formally, we use the following specification:

$$\mu(Y) = \mathrm{Ln}(\mu_0 + 500\lambda Y) - \sigma^2/2, \tag{55.2}$$

where λ denotes the degree of overconfidence. We assume that stockholders have realistic expectations, which are represented by $\lambda = 1$, and that managers use $\lambda > 1$ to form their expectations. Thus, f(X) can be written as f(X, λ), where f(X, $\lambda = 1$) represents the realistic cash flow distribution, while f(X, $\lambda > 1$)

[5] Palmon and Venezia (2012) explore the effect of managerial overconfidence on the firm's stockholders and show that overconfidence may improve welfare. However, that study does not investigate the optimal strike price of managerial incentive options.

[6] In our model we assume symmetry of information between the manager and the firm regarding the distribution of cash flows of the firm except for the different view of the effect of the manager's effort on cash flows.

represents the cash flow distribution as viewed by overconfident managers. For notation brevity, we suppress the λ in $f(X, \lambda)$. By the known properties of the lognormal distribution, the mean and variance of X equal $e^{\left[\mu(Y)+0.5\sigma^2\right]}$ and $\left[e^{\left[2\mu(Y)+\sigma^2\right]}\left(e^{\sigma^2}-1\right)\right]$, respectively. Thus, it follows from Eq. 55.2 that a person with a λ overconfidence measure believes that the mean of the cash flows X is $e^{\mu(Y)+0.5\sigma^2} = \mu_0 + 500\lambda Y$ and that their coefficient of variation is approximately σ.[7] Since managers and stockholders differ in their perception of the distributions of cash flows, one must be careful in their use. In what follows we refer to the distribution of cash flows as seen by stockholders as the realistic distribution, and will make a special note whenever the manager's overconfident beliefs are used.

Except for her overconfidence, the manager is assumed to be rational and to choose her extra effort so as to maximize the expected value of the following utility function which exhibits constant relative risk aversion (CRRA) with respect to compensation:

$$U(I, Y) = \frac{1}{1-\gamma} NY^{\beta} + \frac{1}{1-\gamma} I^{1-\gamma} \tag{55.3}$$

In Eq. 55.3, I denotes the manager's monetary income, γ denotes the constant relative risk aversion measure, N is a scaling constant representing the importance of effort relative to monetary income in the manager's preferences, and the positive parameter β is related to the convexity of the disutility of effort.

Since stockholders cannot observe the manager's extra effort, they propose compensation schemes that depend on the observed cash flows, but not on Y. Stockholders, which we assume to be risk neutral, strive to make the compensation performance sensitive in order to better align the manager's incentives with their own. Stockholders offer the manager a compensation package that includes two components: a fixed wage (W) that she will receive regardless of her extra effort and of the resulting cash flows and options with a strike price (K) for a fraction (s) of the equity of the firm. We assume that stockholders offer the contract that maximizes the value of their equity.

The following timeline of decisions is assumed. At the beginning of the period, the firm chooses the parameters of the compensation contract (K, W, and s) and offers this contract to the manager. Observing the contract parameters, and taking into account the effects of her endeavors on firm cash flows and hence on her compensation, the manager determines the extra-effort level Y that maximizes her expected utility. At the end of the period, X is revealed, and the firm distributes the cash flows to the manager and to the stockholders and then dissolves. The priority of payments is as follows. The firm first pays the wages or only part of them if the cash flows do not suffice. If the cash flows exceed the wage, W, but not (K + W), then the

[7] More precisely the square of the coefficient of variation is $\left[e^{\sigma^2}-1\right]$ which can be approximated by σ^2 since for any small z, e^z-1 is close to z.

managers just receive their fixed wage. The managers are paid the value of the options s(X − K − W), in addition to W if X exceeds K + W. The manager therefore receives the cash flows I(X) defined by

$$I(x) = \begin{Bmatrix} X \\ W \\ W + s(X - W - K) \end{Bmatrix} \quad \text{when} \quad \begin{Bmatrix} X \le W \\ W \le X \le W + K \\ W + K \le X \end{Bmatrix} \tag{55.4}$$

The shareholders get the residual cash flows.

In the above cash flow formula, the first range covers the case where cash flows do not suffice to pay the entire wage. The second range covers the case where the options expire out-of-the-money, and the manager gets the promised wage. The third range represents cash flows that are large enough so that the options expire in-the-money. In addition to the wage, the manager receives a proportion, s, of the value of the firm above the threshold value of K. The expected utility of the manager E{U[I(X),Y]} which governs her behavior, and her expected compensation E[I(X)], can be obtained by integrating her utility U[I(X),Y] given in Eq. 55.3 and her compensation I(X), given in Eq. 55.4, respectively. We note that the manager chooses the effort level so as to maximize the expected utility using her perception of the distribution of the firm's final cash flows, while stockholders choose the parameters of the compensation contract using the realistic cash flow distribution to calculate the expected cash flows and managerial compensation.

Shareholders receive all cash flows that are not received by the manager. Since stockholders are risk neutral and rational, stockholders' equity value (SEV) is the expected value of these payments, using the realistic distribution function, and hence,[8]

$$SEV = E(Cashflows) - E[I(X)] = \int_0^\infty Xf(X)dX - E[I(X)] \tag{55.5}$$

While the derivation of the optimal contract for any set of exogenous parameters is conceptually straightforward, unfortunately, closed form solutions cannot be obtained in our integrative model. Hence, following Hall and Murphy (2000), we resort to simulations to evaluate the optimal contracts and analyze their properties. In addition, we cannot use the Black-Scholes model to evaluate the executive stock options since this model takes the values of the underlying asset as given, whereas a crucial aspect of the managerial incentive scheme of our model is that managerial extra effort and firm value are endogenously determined. We therefore introduce a model that simultaneously simulates the manager's optimal extra-effort level as well as the expected values of the executive stock options and shareholders' equity

[8]Discounting the cash flows by an appropriate risk-adjusted discount rate would yield a linear transformation of equity values. To simplify the presentation, and as is common in the literature, we abstract from that.

for each compensation package. We check the robustness of our results by using alternative parameters for the manager's utility function and the distribution functions of the cash flows.

55.3 The Simulation Procedures

We assume that managers have external employment opportunities and that stockholders offer managerial compensation packages that provide the managers with a comparable expected utility.[9] Without loss of generality (i.e., by an appropriate definition of the wage units), we assume that these external employment opportunities provide the manager an expected utility that equals the level of utility that is obtained from a fixed compensation of 100 in the absence of any extra effort. Thus, in all the simulations, we set the manager's expected utility to correspond to the level obtained from a fixed compensation of 100 (wage = 100 and no option grants) and no extra effort (which is the optimal extra-effort choice when no options are granted).[10] We then search over a grid of strike prices (using four-digit accuracy) and find for each strike price the percentage of options that should be awarded so that the manager's expected utility equals the expected utility target when the manager chooses the optimal extra-effort level. We identify the strike price that is associated with the highest equity level and refer to this contract as the optimal contract for the given set of parameters.

In calibrating the other parameters for the simulations, we try to approximately conform to Hall and Murphy (2000) and Hall and Liebman (1998); to studies that simulate decisions with effort aversion, such as Bitler et al. (2005); and to studies that explore the effect of overconfidence on corporate decisions, such as Malmendier and Tate (2005a, b, 2008).[11] Accordingly, we set the parameters in our base case as follows. The coefficient of variation, σ, equals 0.3, and thus, the standard deviation is 0.3E(X). Since the expected cash flows serve as numeraire, the volatility is determined solely by the coefficient of variation. In our base case, we set the managerial wage to equal 50.[12] The expected cash flows as viewed by an overconfident manager with an overconfidence measure of λ are $E(X) = 45{,}000 + 500\lambda Y$ (i.e., $\mu_0 = 45{,}000$). The risk aversion and effort aversion parameters are $\gamma = 4$ and $\beta = 3$, respectively.

[9] See Appendix 1 for more details.

[10] When the manager is overconfident, this expected utility is calculated according to the manager's expectations.

[11] See Appendix 1 for the explanation for the calibration of our model. To be on the safe side and in stride with explanations for the risk premium puzzle, we use higher values for the risk aversion parameter.

[12] It should be noted that although the wage level in our base case equals half of the fixed compensation that corresponds to the utility target, it equals only about 11 % of the expected compensation under the optimal contract when managers are realistic. When managers are overconfident, a wage of 50 consists of less than 11 % of total compensation according to the manager's expectations but more than 11 % according to the realistic expectations.

We consider overconfidence levels between $\lambda = 1$ (no overconfidence) and $\lambda = 2.5$ in 0.5 increments.

We examine the robustness of the results to deviations from the base case combination of parameters by simulating with several alternative sets of exogenous parameters. We repeat the analysis for many alternative sets of the exogenous parameters: the manager's risk and extra-effort aversion, γ and β, as well as the volatility measure of cash flows, σ.[13]

55.4 Results and Discussion

In Table 55.1 we present the impact of overconfidence on the strike price that stockholders choose to offer: the moneyness, the percentage of the firm given as options, the effort choice of managers, the stockholders' equity value, and the expected managerial compensation. The expected compensation is calculated both under the realistic distribution and under the subjective distribution of the manager.

One observes from Table 55.1 that the strike price, the options' moneyness, the optimal managerial effort, the value of the stockholders' equity, and the expected compensation according to the managers' expectations are directly related to overconfidence. The optimal strike price (in thousands of dollars; strike prices will be denoted in thousands of dollars in the rest of the study) for a rational manager is 40.71, with a 0.60 moneyness (which can be described as deep-in-the-money), but it rises to 63.74 with a 0.89 moneyness (closer to at-the-money) when $\lambda = 2.5$.[14] Managers also work harder the more overconfident they are (Y increases from around 47 when they are realistic to around 52 when $\lambda = 2.5$). Consequently, in order to hold the managers' expected utility fixed, their subjective expected monetary compensation must increase with overconfidence to compensate for the extra risk resulting from the higher strike price and for the additional effort they exert. The expected compensation the stockholders perceive they pay according to the realistic expectation, however, is inversely related to the overconfidence measure as they take advantage of managers' unrealistic expectations. The SEVs of the optimal contracts increase as managerial overconfidence increases (see column 5, the SEV rises from 68,099 when $\lambda = 1$ to 70,958 when $\lambda = 2.5$, an increase of about 6 %). That is, the stockholders benefit from the managers overestimating their powers.

This analysis suggests that stockholders are able to induce overconfident managers to exert higher effort levels even though the objective contract parameters they offer them are less favorable (the managers work harder but receive lower

[13]Because of scaling there is no need to conduct robustness checks for the expected cash flows.

[14]The moneyness measure depends on the strike price and the value of equity, which in turn depends on effort. Thus, the moneyness measure varies with overconfidence because effort varies with overconfidence, even when the strike price remains constant.

Table 55.1 Stockholders' equity values (SEV), strike prices (K), moneyness (K/S), effort levels (Y), and other parameters of interest. Base case: $\gamma = 4$ $\beta = 3$, $\sigma = 0.3$, wage $= 50$, $N = 4{,}000$

Overconfidence measure, λ	Strike price (K) in 1000s	Moneyness (K/S)	Percentage of firm given as options, (s)	Effort (Y)	Stockholders Equity Value (SEV)	Expected compensation, realistic valuation	Expected compensation, managers valuation	Sum of SEV and realistic expected compensation
	1	2	3	4	5	6	7	8
1	40.71	0.5944	1.40%	47.083	68099.42	441.89	441.89	68541.31
1.5	48.16	0.6891	1.35%	49.867	69580.04	353.41	513.34	69933.45
2	55.88	0.7910	1.35%	51.393	70417.74	278.65	598.49	70696.39
2.5	63.74	0.8962	1.38%	52.351	70958.63	217.12	698.06	71175.75

expected compensation). In particular the optimal strike prices of the options that stockholders award overconfident managers are increasing with their overconfidence. While realistic managers estimate that there is a substantial probability that options with an at-the-money strike price will be worthless regardless of their effort, overconfident managers may believe that their efforts will enhance the values of such options making them valuable.

In practice executive options usually are provided with at-the-money strike prices. When overconfidence or some other behavioral biases are not present, theory has shown (see, e.g., Dittmann et al. 2010; Dittmann and Yu 2011; Palmon et al. 2008), contrary to Hall and Murphy, that at-the-money prices are not optimal. Hall and Murphy argue that at-the-money prices are optimal because they provide maximum sensitivity to stock prices, but their argument does not hold when the managers are risk averse and effort averse. Managers must be adequately compensated for their efforts and for risk taking, and a balance must be reached between their efforts, risk taking, and their pay. As Palmon et al. have shown, the optimal balance is reached by issuing in-the-money options which do not necessarily provide maximum sensitivity to stock prices. If managers are overconfident, however, that makes them more amenable for stock price sensitivity, and hence, they will prefer higher strike prices which are closer to the at-the-money options usually awarded in practice.

We also note in Table 55.1 that managerial overconfidence increases stockholders' equity value. Given that the compensation is determined so as to equate the manager's expected utility to the target expected utility, it follows from Table 55.1 that consistent with the results of Palmon and Venezia (2012), the total welfare of both the managers and the stockholders improves with increased managerial overconfidence. The fixed-level expected utility of the manager is determined according to their subjective, overoptimistic perception. However, when evaluated according to the realistic view, expected managerial compensation falls with overconfidence. We note that, nonetheless, the difference between the monetary expected compensations according to the overoptimistic and realistic expectations is smaller than the monetary gains to stockholders from overconfidence, so that the sum of realistic compensation and SEV rises with overconfidence (see column 8). Thus, also in terms of realistic monetary values, the welfare of the stakeholders (stockholders and managers) increases with overconfidence.

In Table 55.2 we provide sensitivity analysis examining the effect of each of the parameters on the behavior of stockholders and managers. We present the results of only one or two changes in each of the exogenous parameters, but we conduct many other simulations, and all provide the same qualitative results.[15] In all the panels, higher overconfidence measure is associated with higher strike prices, moneyness levels, optimal managerial effort, value of the stockholders' equity, and expected compensation according to the managers' expectations. They also are associated

[15]The results of these simulations can be obtained from the authors upon request.

Table 55.2 Stockholders' equity values (SEV), strike prices (K), moneyness (K/S), effort levels (Y), and other parameters of interest. Departures from the base case

Overconfidence measure, λ	Strike price (K) in 1000s	Moneyness (K/S)	Percentage of firm given as options, (s)	Effort (Y)	Stockholders' Equity Value, SEV	Expected compensation, realistic stockholders' valuation	Expected compensation, managers' valuation	Sum of SEV and realistic expected compensation
	1	2	3	4	5	6	7	8
Panel A: wage = 25								
1	31.56	0.4490	1.25 %	50.63	69,807.58	508.76	508.76	70,316.34
1.5	37.49	0.5252	1.14 %	52.83	71,000.72	413.11	563.28	71,413.83
2	43.57	0.6053	1.08 %	54.01	71,670.56	334.41	622.54	72,004.97
2.5	49.74	0.6875	1.04 %	54.75	72,106.28	269.37	688.74	72,375.65
Panel B: wage = 75								
1	51.70	0.7859	1.63 %	41.71	65,518.91	336.51	336.51	65,855.42
1.5	60.61	0.8965	1.70 %	45.36	67,411.64	269.97	420.45	67,681.61
2	69.91	1.0186	1.82 %	47.42	68,498.86	211.05	522.88	68,709.91
2.5	79.42	1.1462	2.01 %	48.72	69,195.89	166.20	653.44	69,362.09
Panel C: $\sigma = 0.2$								
1	48.97	0.6935	1.80 %	51.33	70,225.15	441.49	441.49	70,666.64
1.5	58.39	0.8138	1.83 %	53.60	71,489.22	312.64	542.64	71,801.86
2	68.10	0.9412	2.02 %	54.82	72,198.16	209.92	692.03	72,408.08
2.5	77.94	1.0716	2.35 %	55.56	72,642.57	137.08	912.80	72,779.65
Panel D: $\sigma = 0.4$								
1	34.02	0.5090	1.19 %	43.77	66,438.75	444.93	444.93	66,883.68
1.5	40.00	0.5850	1.12 %	46.86	68,050.65	377.20	503.87	68,427.85
2	46.22	0.6675	1.09 %	48.59	68,975.32	317.47	566.82	69,292.79
2.5	52.58	0.7534	1.07 %	49.69	69,577.77	266.95	635.35	69,844.72

(continued)

Table 55.2 (continued)

Overconfidence measure, λ	Strike price (K) in 1000s	Moneyness (K/S)	Percentage of firm given as options, (s)	Effort (Y)	Stockholders' Equity Value, SEV	Expected compensation, realistic stockholders' valuation	Expected compensation, managers' valuation	Sum of SEV and realistic expected compensation
	1	2	3	4	5	6	7	8
Panel E: $\gamma = 3$								
1	45.70	0.6801	2.09 %	44.49	66,732.83	513.41	513.41	67,246.24
1.5	53.94	0.7839	2.14 %	47.71	68,445.70	411.53	639.26	68,857.23
2	62.52	0.8969	2.28 %	49.52	69,438.59	319.32	795.50	69,757.91
2.5	71.28	1.0142	2.49 %	50.66	70,086.17	243.48	992.82	70,329.65
Panel F: $\gamma = 5$								
1	36.91	0.5325	1.08 %	48.73	68,964.71	400.77	400.77	69,365.48
1.5	43.74	0.6198	1.01 %	51.24	70,295.63	323.30	449.43	70,618.93
2	50.78	0.7127	0.97 %	52.59	71,038.19	258.80	504.00	71,296.99
2.5	57.95	0.8085	0.96 %	53.44	71,515.08	206.66	567.94	71,721.74
Panel G: N = 2,000								
1	44.54	0.5886	1.36 %	61.44	75,242.70	477.96	477.96	75,720.66
1.5	54.10	0.7013	1.31 %	64.39	76,830.62	364.92	565.18	77,195.55
2	63.94	0.8204	1.32 %	65.98	77,714.30	274.85	672.50	77,989.15
2.5	73.93	0.9426	1.37 %	66.97	78,277.56	205.29	804.84	78,482.85
Panel H: N = 6,000								
1	38.87	0.5977	1.41 %	40.15	64,656.11	421.35	421.35	65,077.47
1.5	45.30	0.6826	1.36 %	42.83	66,067.86	346.87	486.08	66,414.74
2	51.97	0.7745	1.35 %	44.31	66,875.94	279.15	557.17	67,155.09
2.5	58.80	0.8701	1.38 %	45.25	67,401.85	223.48	643.22	67,625.33

with lower expected managerial compensation according to the realistic expectation. This indicates that the qualitative results obtained from the base case prevail also for a host of other parameters.

In panels A and B, we examine the effect of the fixed wages on the results. A higher wage level implies a lower value for the option component of the compensation. Imposing the use of lower-valued options, stockholders choose options that are more responsive to cash flow changes. This is obtained by an increase in the ownership percentage and in the strike price. In panels C and D, we set the coefficient of variation, σ, which equals 0.3 in the base case, to 0.2 and 0.4, respectively. We observe in panel D that facing a higher coefficient of variation, managers prefer a compensation that is less sensitive to firm cash flows, which is achieved by selecting a contract specifying a smaller ownership fraction and a lower strike price. Finally, in panels E and F, we observe the effects of varying risk aversion and in panels G and H those of varying effort aversion. Overall, higher risk aversion levels are associated with less risky compensations as they induce optimal contracts with a smaller option ownership percentage and a lower strike price. Higher effort aversion results in lower SEV and also in lower total monetary welfare.

The above sensitivity analysis shows that the effect of changing the parameters quite conforms to intuition, adding to the robustness of our results. We also note that regardless of the parameters considered, the effect of overconfidence on the qualitative behavior is the same as that observed from the base case. In particular, the higher the overconfidence, the higher the optimal strike prices and the closer they are to the at-the-money levels.

55.5 Conclusion

Our study suggests an explanation for the puzzling questions of why most incentive stock options are issued with at-the-money strike prices. This practice seems arbitrary and beyond its institutional appeal and its expired tax advantages; its main theoretical backing is that it provides the highest sensitivity to stock price. Several studies however have shown that in many cases it is inferior to awarding in-the-money options. Our analysis demonstrates that the optimal strike prices of incentive stock options when managers are overconfident are higher than the corresponding strike prices when managers are realistic, and are closer to the at-the-money strike prices awarded in practice. This makes at-the-money options more attractive to overconfident managers, and hence, given the ubiquity of overconfident managers, it provides support for the popularity of awarding such options. We also show that overoptimistic CEOs receive lower compensation than their realistic counterparts and that the stockholders benefit from their managers' bias. The combined welfare of the firm's stakeholders however is positively related to managerial overconfidence, hence providing support to the survival of managerial overconfidence.

Assef and Santos (2005) interpret the strike price as an intermediate instrument (between wages and stocks) in the incentive schemes for managers. Similarly one can interpret an in-the-money strike price as an intermediate instrument between a stock (zero strike price) and an at-the-money option. Since in practice, because of institutional reasons or inertia, firms are constrained to choose options with at-the-money strike price, they achieve their instrumental in-the-money strike price by choosing an appropriate weight of options relative to stock grants in their compensation contract. According to such an interpretation and from our results showing that higher overconfidence implies higher strike prices, it follows that the observable weight of options in the compensation contract may serve as a proxy for an unobservable degree of confidence.

Appendix 1

In this appendix we expand on the simulations we conduct. These simulations are intended to identify the contracts that yield the highest stockholders' equity value subject to manager's incentive compatibility and participation constraints. That is, the managers choose their effort optimally, and their resulting expected utility equals a predetermined level representing their alternative opportunities. Because we are studying the impact of overconfidence on the strike price, our calculations focus on the trade-off between the strike price and the fraction of the company that is awarded as options. For simplicity, we consider contracts that include only a fixed wage and options.

The first step in our simulation is the selection of the appropriate distribution of the company's cash flow as a function of managerial effort and the manager's utility as a function of managerial effort and compensation. In accordance with conventional assumptions in the options literature, we assume that the firms' cash flows, X, are lognormally distributed with the distribution function (55.1) where Y denotes the managerial extra effort (a managerial choice variable) and $\mu(Y)$ and σ denote, respectively, the mean and the standard deviation of the underlying normal distribution of the natural logarithm of X. We assume that managerial effort increases cash flows, that overconfident managers overestimate the impact of their effort on cash flows, and that the impact of effort on the mean of the natural logarithm of X is presented in Eq. 55.2.

We refer to Hek (1999) and Bitler et al. (2005) for the choice of the parameters and the shape of the manager's utility function that depends also on leisure.[16] We start with a base case of parameters and repeat the analysis for a large set of parameters around the base case. We chose the base case so that these parameters and the deviations around them that we also analyze cover the equivalent parameters used in similar studies. These simulations help verify that our results are robust to the choice of parameter values. They also are used to examine to what

[16]We found additional estimates of effort disutility (leisure utility) in the following papers: Dowell (1985), Kiker and Mendes de Oliveira (1990), and Prasch (2001). These estimates varied in the functional form as well as in the level of effort aversion.

extent the effects of changes in the parameter values on the outcomes coincide with economic intuition.

The parameter μ_0 serves as a numeraire for the other cash flows related parameters, and is chosen, without loss of generality, to equal 45,000. That, in the absence of managerial extra effort, the expected value of the company's cash flows is 45,000. Since the expected cash flows serve as numeraire, the ratio of the standard deviation of the cash flows per share to their expected value is a surrogate for the standard deviation of stock returns. Since Hall and Murphy (2000) used a standard deviation of 0.3, we chose this value also for our base case coefficient of variation.

The appropriate measure of risk aversion is harder to agree upon. Early estimates of risk aversion put this variable at around two (see, e.g., Mehra and Prescott 1985), but they are based on aggregate data and not on CEO compensation data.[17] In our study, in line with more advanced econometric methods (see, e.g., Campbell et al. 1996), we prefer using a base case risk aversion measure of four, slightly higher than the measure of three suggested by Malmendier and Tate (2008) and Hall and Liebman (1998). Our simulations (see, e.g., Glasserman 2003) and sensitivity analysis, of course, cover these parameters as well.

The next step in the simulation process is to identify, for each overconfidence level, the executive options' strike price that is optimal for stockholders. All the simulations were conducted using Mathematica. Because it is not possible to express the equity value as an explicit function of the strike price, we search for the optimal strike price by calculating the equity values that are associated with a set of discrete strike prices. Our search was facilitated by assuming that the stockholders know the manager's reservation expected utility. We assume that reservation utility to equal the utility obtained from a fixed salary of 100 with no extra effort.

For any given wage, the strike price and the fraction of the company awarded to the manager (which is a continuous variable representing the number of options the manager receives; we will henceforth use the latter expression) determine the value of the options to the managers and their cost to the stockholders. For each strike price and number of options, we then find the effort that the manager chooses to apply in order to maximize his/her expected utility Eq. 55.3. For each given strike price, the stockholders, well aware of the managers' reactions, will offer them the number of options that yield their reservation utility. We calculate the value of the stockholders' equity for each strike price (in thousands of dollars, using two digits beyond the decimal point) and identify the strike price that yields a maximum for stockholders' equity.

For each set of parameters for the cash flow distribution function and the managerial utility function, as well as for the several values of fixed salary (50 for the base case, 25 and 75 for the presented robustness simulations), we obtain the optimal effort, stockholders' equity value, and the expected managerial

[17] Similar estimates are provided in other contexts by Carpenter (2000), Constantinides et al. (2002), Epstein and Zin (1991), Friend and Blume (1975), and Levy (1994).

compensation according to the manager's overconfident view and the stockholders' realistic expectations. We repeat these simulations for several values of the overconfidence measure λ. Presented here however are just four such values: 1 for the realistic expectations and 1.5, 2, and 2.5 for increasing levels of overconfidence. We used quite a few simulations but choose to present a subset of the results as all showed the same qualitative results. In addition to verifying the robustness of the results to the choice of the parameter values, they also help examine to what extent the effects of changes in the parameter values on the outcomes coincide with economic intuition.

Acknowledgment We thank Darius Palia, Orly Sade, and seminar participants at Rutgers University and the Universitat Pompeu Fabra for helpful comments and suggestions. The financial support of The Sanger Family Chair for Banking and Risk Management, The Galanter Fund, The Mordecai Zagagi Fund, the Whitcomb Center for Research in Financial Services, and The School of Accounting, the Hebrew University are gratefully acknowledged.

References

Assef, J. G., & Santos, M. (2005). Stock options and managerial optimal contracts. *Economic Theory, 26*, 813–837.

Bergman, N. K., & Jenter, D. (2007). Employee sentiment and stock option compensation. *Journal of Financial Economics, 84*, 667–712.

Bitler, M. P., Moskowitz, T. J., & Vissing-Jorgensen, A. (2005). Testing agency theory with entrepreneur extra-effort and wealth. *Journal of Finance, 60*(2), 539–576.

Campbell, J., Lo, A., & MacKinlay, C. (1996). *The econometrics of financial markets*. New Jersey, Princeton: Princeton University Press.

Carpenter, J. N. (2000). Does option compensation increase managerial risk appetite. *Journal of Finance, 55*, 2311–2331.

Constantinides, G. M., Donaldson, J. M., & Mehra, R. (2002). Junior can't borrow: A new perspective on the equity premium puzzle. *Quarterly Journal of Economics, 117*, 269–296.

Dittmann, I., & Yu, K. C. (2011). *How important are risk-taking incentives in executive compensation*? (Working Paper). Rotterdam: Erasmus University.

Dittmann, I., Maug, E., & Spalt, O. (2010). Sticks or carrots: Optimal CEO compensation when managers are loss-averse. *Journal of Finance, 65*, 2015–2050.

Dowell, R. (1985). Risk preference and the work-leisure trade-off. *Economic-Inquiry, 23*, 691–701.

Epstein, L. G., & Zin, S. E. (1991). Risk aversion, and the temporal behavior of consumption and asset returns: An empirical analysis. *Journal of Political Economy, 99*, 263–86.

Friend, I., & Blume, M. (1975). The demand for risky assets. *American Economic Review, 65*, 900–927.

Gervais, S., Heaton, J. B., & Odean, T. (2011). Overconfidence, compensation contracts, and capital budgeting. *Journal of Finance, 66*, 1735–1777.

Glaser, M., & Weber, M. (2007). Overconfidence and trading volume. *The Geneva Risk and Insurance Review, 32*, 1–36.

Glasserman, P. (2003). *Monte Carlo methods in financial engineering*. Berlin: Springer.

Hall, B., & Liebman, J. B. (1998). Are CEOs really paid like bureaucrats? Quarterly. *Journal of Economics, 113*, 653–691.

Hall, B., & Murphy, K. J. (2000). Optimal exercise prices for executive stock options. *American Economic Review, 90*, 209–14.

Hall, B. J., & Murphy, K. J. (2002). Stock options for undiversified executives. *Journal of Accounting and Economics, 33*, 3–42.

Hek, P. A. (1999). Endogenous technological change with leisure-dependent utility. *Economic Theory, 14*, 669–84.

Holmstrom, B. (1979). Moral hazard and observability. *Bell Journal of Economics, 10*, 74–91.

Kiker, B. F., & Mendes-de-Oliveira, M. (1990). Estimation and valuation of non-leisure time. *Oxford Bulletin of Economics and Statistics, 52*, 115–141.

Levy, H. (1994). Absolute and relative risk aversion: An experimental study. *Journal of Risk and Uncertainty, 8*, 289–307.

Mahajan, S. (2002). *Role of out-of-money options in executive compensation* (Working Paper). New York: NYU, Stern School of Business.

Malmendier, U., & Tate, G. (2005a). Does overconfidence affect corporate investment? CEO overconfidence measures revisited. *European Financial Management, 11*, 649–659.

Malmendier, U., & Tate, G. (2005b). CEO overconfidence and corporate investment. *Journal of Finance, 60*, 2661–2700.

Malmendier, U., & Tate, G. (2008). Who makes acquisitions? CEO overconfidence and the market's reaction. *Journal of Financial Economics, 89*, 20–43.

Mehra, R., & Prescott, E. C. (1985). The equity premium: A puzzle. *Journal of Monetary Economics, 15*, 145–161.

Otto, C. A. (2011). *CEO optimism and incentive compensation* (Working Paper). London: Business School.

Oyer, P., & Schaefer, S. (2005). Why do some firms give stock options to all employees? An empirical examinations of alternative theories. *Journal of Financial Economics, 76*, 99–133.

Palmon, O., & Venezia, I. (2012, forthcoming). A rationale for hiring irrationally overconfident managers. In *Encyclopedia of finance* (2nd ed.). Berlin: Springer.

Palmon, O., Chen, R. R., Bar Yosef, S., & Venezia, I. (2008). Optimal strike prices of stock options for effort averse executives. *Journal of Banking and Finance, 32*, 229–239.

Prasch, R. E. (2001). Work, leisure, and the labor supply curve: A reply to David Spencer. *Journal of Economic Issues, 35*, 1001–1007.

Roll, R. (1986). The hubris hypothesis of corporate takeovers. *Journal of Business, 59*, 197–216.

Suntheim, F. (2012). *CEO overconfidence in banking* (Working Paper). Milano: Bocconi University.

Density and Conditional Distribution-Based Specification Analysis

56

Diep Duong and Norman R. Swanson

Contents

Abstract

The technique of using densities and conditional distributions to carry out consistent specification testing and model selection amongst multiple diffusion processes has received considerable attention from both financial theoreticians and empirical econometricians over the last two decades. In this chapter, we discuss advances to this literature introduced by Corradi and Swanson (J Econom 124:117–148, 2005), who compare the cumulative distribution (marginal or joint) implied by a hypothesized null model with corresponding empirical distributions of observed data. We also outline and expand upon

D. Duong (✉)
Department of Business and Economics, Utica College, Utica, NY, USA
e-mail: dnduong@utica.edu

N.R. Swanson
Department of Economics, Rutgers, The State University of New Jersey, New Brunswick, NJ, USA
e-mail: nswanson@econ.rutgers.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_56,

further testing results from Bhardwaj et al. (J Bus Econ Stat 26:176–193, 2008) and Corradi and Swanson (J Econom 161:304–324, 2011). In particular, parametric specification tests in the spirit of the conditional Kolmogorov test of Andrews (Econometrica 65:1097–1128, 1997) that rely on block bootstrap resampling methods in order to construct test critical values are first discussed. Thereafter, extensions due to Bhardwaj et al. (J Bus Econ Stat 26:176–193, 2008) for cases where the functional form of the conditional density is unknown are introduced, and related continuous time simulation methods are introduced. Finally, we broaden our discussion from single process specification testing to multiple process model selection by discussing how to construct predictive densities and how to compare the accuracy of predictive densities derived from alternative (possibly misspecified) diffusion models. In particular, we generalize simulation steps outlined in Cai and Swanson (J Empir Financ 18:743–764, 2011) to multifactor models where the number of latent variables is larger than three. We finish the chapter with an empirical illustration of model selection amongst alternative short-term interest rate models.

Keywords

Multifactor diffusion process • Specification test • Out-of-sample forecasts • Conditional distribution • Model selection • Block bootstrap • Jump process

56.1 Introduction

The last three decades have provided a unique opportunity to observe numerous interesting developments in finance, financial econometrics, and statistics. For example, although starting as a narrow subfield, financial econometrics has recently transformed itself into an important discipline, equipping financial economic researchers and industry practitioners with immensely helpful tools for estimation, testing, and forecasting. One of these developments has involved the development of "state-of-the-art" consistent specification tests for continuous time models, including not only the geometric Brownian motion process used to describe the dynamics of asset returns (Merton (1973)) but also a myriad of other diffusion models used in finance, such as the Ornstein-Uhlenbeck process introduced by Vasicek (1977); the constant elastic volatility process applied by Beckers (1980); the square root process due to Cox et al. (1985); the so-called CKLS model by Chan et al. (1992); various three-factor models proposed Chen (1996); stochastic volatility processes such as generalized CIR of Andersen and Lund (1997); and the generic class of affine jump diffusion processes discussed in Duffle et al. (2000).[1]

The plethora of available diffusion models allow decision makers to be flexible when choosing a specification to be subsequently used in contexts ranging from equity and option pricing, to term structure modeling and risk management.

[1]For complete details, see Sect. 56.2.2.

Moreover, the use of high-frequency data when estimating such model, in continuous time contexts, allows investors to continuously update their dynamic trading strategies in real time.[2] However, for statisticians and econometricians, the vast number of available models has important implications for formalizing model selection and specification testing methods. This has led to several key papers that have recently been published in the area of parametric and nonparametric specification testing. Most of the papers focus on the ongoing "search" for correct Markov and stationary models that "fit" historical data and associated dynamics. In this literature, it is important to note that correct specification of a joint distribution is not the same as that of a conditional distribution, and hence the recent focus on conditional distributions, given that most models have an interpretation as conditional models. In summary, the key issue in the construction of model selection and specification tests of conditional distributions is the fact that knowledge of the transition density (or conditional distribution) in general cannot be inferred from knowledge of the drift and variance terms of a diffusion model. If the functional form of the density is available parametrically, though, one can test the hypothesis of correct specification of a diffusion via the probability integral transform approach of Diebold et al. (1998); the cross-spectrum approach of Hong (2001), Hong and Li (2005), and Hong et al. (2007); the martingalization-type Kolmogorov test of Bai (2003); or the normality transformation approaches of Bontemps and Meddahi (2005) and Duan (2003). Furthermore, if the transition density is unknown, one can construct a nonparametric test by comparing a kernel density estimator of the actual and simulated data, for example, as in Altissimo and Mele (2009) and Thompson (2008), or by comparing the conditional distribution of the simulated and the historical data, as in Bhardwaj et al. (2008). One can also use the methods of Aït-Sahalia (2002) and Aït-Sahalia et al. (2009), in which they compare closed form approximations of conditional densities under the null, using data-driven kernel density estimates.

For clarity and ease of presentation, we categorize the above literature into two areas. The first area, initiated by the seminal work of Aït-Sahalia (1996) and later followed by Pritsker (1998) and Jiang (1998), breaks new ground in the continuous time specification testing literature by comparing marginal densities implied by hypothesized null models with nonparametric estimates thereof. These sorts of tests examine one-factor specifications. The second area of testing, as initiated in Corradi and Swanson (2005), does not look at densities. Instead, they compare cumulative distributions (marginal, joint, or conditional) implied by a hypothesized null model with corresponding empirical distributions. A natural extension of these sorts of tests involves model selection amongst alternative predictive densities associated with competing models. While Corradi and Swanson (2005) focus on cases where the functional form of the conditional density is known, Bhardwaj et al. (2008) use simulation methods to examine testing in cases where the functional form of the conditional density is unknown. Corradi and Swanson (2011) and

[2]For further discussion, see Duong and Swanson (2010, 2011).

Cai and Swanson (2011) take the analysis of Bhardwaj et al. (2008) on Step further and focus on the comparison of out-of-sample predictive accuracy of possibly misspecified diffusion models, when the conditional distribution is not known in closed form (i.e., they "choose" amongst competing models based on predictive density model performance). The "best" model is selected by constructing tests that compare both predictive densities and predictive conditional confidence intervals associated with alternative models.

In this chapter, we primarily focus our attention on the second area of the model selection and testing literature.[3] One feature of all the tests that we shall discuss is that, given that they are based on the comparison of CDFs, they obtain parametric rates. Moreover, the tests can be used to evaluate single and multiple factor and dimensional models, regardless of whether or not the functional form of the conditional distribution is known.

In addition to discussing simple diffusion process specification tests of Corradi and Swanson (2005), we discuss tests discussed in Bhardwaj et al. (2008) and Corradi and Swanson (2011) and provide some generalizations and additional results. In particular, parametric specification tests in the spirit of the conditional Kolmogorov test of Andrews (1997) that rely on block bootstrap resampling methods in order to construct test critical values are first discussed. Thereafter, extensions due to Bhardwaj et al. (2008) for cases where the functional form of the conditional density is unknown are introduced, and related continuous time simulation methods are introduced. Finally, we broaden our discussion from single dimensional specification testing to multiple dimensional selection by discussing how to construct predictive densities and how to compare the accuracy of predictive densities derived from alternative (possibly misspecified) diffusion models as in Corradi and Swanson (2011). In addition, we generalize simulation and testing procedures introduced in Cai and Swanson (2011) to more complicated multifactor and multidimensional models where the number of latent variables is larger than three. These final tests can be thought of as continuous time generalizations of the discrete time "reality check" test statistics of White (2000), which are widely used in empirical finance (see, e.g., Sullivan et al. (1999, 2001)). We finish the chapter with an empirical illustration of model selection amongst alternative short-term interest rate models, drawing on Bhardwaj et al. (2008), Corradi and Swanson (2011) and Cai and Swanson (2011).

Of the final note is that the test statistics discussed here are implemented via use of simple bootstrap methods for critical value simulation. We use the bootstrap because the covariance kernels of the (Gaussian) asymptotic limiting distributions of the test statistics are shown to contain terms deriving from both the contribution of recursive parameter estimation error (PEE) and the time dependence of data. Asymptotic critical value thus cannot be tabulated in a usual way. Several methods can easily be implemented in this context. First one can use block bootstrapping procedures, as discussed below. Second one can use the conditional p-value

[3]For a recent survey on results in the first area of this literature, see Aït-Sahalia (2007).

approach of Corradi and Swanson (2002) which extends the work of Hansen (1996) and Inoue (2001) to the case of nonvanishing parameter estimation error. Third is the subsampling method of Politis et al. (1999), which has clear efficiency "costs," but is easy to implement. Use of the latter two methods yields simulated (or subsample based) critical values that diverge at rate equivalent to the block size length under the alternative. This is the main drawback to their use in our context. We therefore focus on use of a block bootstrap that mimics the contribution of parameter estimation error in a recursive setting and in the context of time series data. In general, use of the block bootstrap approach is made feasible by establishing consistency and asymptotic normality of both simulated generalized method of moments (SGMM) and nonparametric simulated quasi-maximum likelihood (NPSQML) estimators of (possibly misspecified) diffusion models, in a recursive setting, and by establishing the first-order validity of their bootstrap analogs.

The rest of the paper is organized as follows. In Sect. 56.2, we present our setup and discuss various diffusion models used in finance and financial econometrics. Section 56.3 outlines the specification testing hypotheses, presents the cumulative distribution-based test statistics for one-factor and multiple-factor models, discusses relevant procedures for simulation and estimation, and outlines bootstrap techniques that can be used for critical value tabulation. In Sect. 56.4, we present a small empirical illustration. Section 56.5 summarizes and concludes.

56.2 Setup

56.2.1 Diffusion Models in Finance and Financial Econometrics

For the past two decades, continuous time models have taken center stage in the field of financial econometrics, particularly in the context of structural modeling, option pricing, risk management, and volatility forecasting. One key advantage of continuous time models is that they allow financial econometricians to use the full information set that is available. With the availability of high-frequency data and current computation capability, one can update information, model estimates, and predictions in milliseconds. In this section, we will summarize some of the standard models that have been used in asset pricing as well as term structure modeling. Generally, assume that financial asset returns follow Ito-semimartingale processes with jumps, which are the solution to the following stochastic differential equation system:

$$\begin{aligned} X(t_-) = &\int_0^t b(X(s_-), \theta_0)ds - \lambda_0 t \int_Y y\phi(y)dy \\ &+ \int_0^t \sigma(X(s_-), \theta_0)dW(s) + \sum_{j=1}^{J_t} y_j, \end{aligned} \tag{56.1}$$

where $X(t_{-})$ is a cadlag process (right continuous with left limit) for $t \in \Re^{+}$ and is an N-dimensional vector of variables, $W(t)$ is an N-dimensional Brownian motion, $b(\cdot)$ is N-dimensional function of $X(t_{-})$, and $\sigma(\cdot)$ is an $N \times N$ matrix-valued function of $X(t_{-})$, where θ_0 is an unknown true parameter. J_t is a Poisson process with intensity parameter λ_0, λ_0 finite, and the N-dimensional jump size, y_j, is i.i.d. with marginal distribution given by ϕ. Both J_t and y_j are assumed to be independent of the driving Brownian motion, $W(t)$.[4] Also, note that $\int_Y y\phi(y)dy$ denotes the mean jump size, hereafter denoted by μ_0. Over a unit time interval, there are on average λ_0 jumps, so that over the time span $[0, t]$, there are on average $\lambda_0 t$ jumps. The dynamics of $X(t_{-})$ is then given by

$$\begin{aligned} dX(t) = {} & \big(b(X(t_{-}), \theta_0) - \lambda_0 \mu_{y,0}\big)dt \\ & + \sigma(X(t_{-}), \theta_0)dW(t) + \int_Y yp(dy, dt), \end{aligned} \tag{56.2}$$

where $p(dy, dt)$ is a random Poisson measure giving point mass at y if a jump occurs in the interval dt and $b(\cdot)$, $\sigma(\cdot)$ are the "drift" and "volatility" functions defining the parametric specification of the model. Hereafter, the same (or similar) notation is used throughout when models are specified.

Through not an exhaustive list, we review some popular models. Models are presented with the "true" parameters.

56.2.1.1 Diffusion Models Without Jumps

Geometric Brownian Motion (Log Normal Model)

In this setup, $b(X(t_{-}), \theta_0) = b_0 X(t)$ and $\sigma(X(t_{-}), \theta_0) = \sigma_0 X(t)$.

$$dX(t) = b_0 X(t)dt + \sigma_0 X(t)dW(t),$$

where b_0 and σ_0 are constants and $W(t)$ is a one-dimensional standard Brownian motion. (Below, other constants such as α_0, β_0, λ_0, γ_0, δ_0, η_0, κ_0, and Ω_0 are also used in model specifications.)

This model is popular in the asset pricing literature. For example, one can model equity prices according to this process, especially in the Black-Scholes option setup or in structured corporate finance.[5] The main drawback of this model is that the return process (log(price)) has constant volatility and is not time varying. However, it is widely used as a convenient "first" econometric model.

Vasicek (1977) and *Ornstein-Uhlenbeck Process*: The process is used to model asset prices, specifically in term structure modeling, and the specification is

$$dX(t) = (\alpha_0 + \beta_0 X(t))dt + \sigma_0 dW(t),$$

[4] Hereafter, $X(t_{-})$ denotes the cadlag, while X_t denotes discrete skeleton for $t = 1, 2, \ldots$.

[5] See Black and Scholes (1973) for details.

where $W(t)$ is a standard Brownian motion and α_0, β_0, and σ_0 are constants. β_0 is negative to ensure the mean reversion of $X(t)$.

Cox et al. (1985) use the following square root process to model the term structure of interest rates:

$$dX(t) = \kappa(\alpha_0 - X(t))dt + \sigma_0\sqrt{X(t)}dW(t),$$

where $W(t)$ is a standard Brownian motion, α_0 is the long-run mean of $X(t)$, κ measures the speed of mean reversion, and σ_0 is a standard deviation parameter and is assumed to be fixed. Also, non-negativity of the process is imposed, as $2\kappa\beta_0 > \sigma_0^2$.

Wong (1964) points out that in the CIR model, $X(t)$ with the dynamics evolving according to

$$dX(t) = ((\alpha_0 - \lambda_0) - X(t))dt + \sqrt{\alpha_0 X(t)}dW(t), \tag{56.3}$$

$$\alpha_0 > 0 \text{ and } \alpha_0 - \lambda_0 > 0,$$

belongs to the linear exponential (or Pearson) family with a closed form cumulative distribution. α_0 and λ_0 are fixed parameters of the model.

The constant elasticity of variance or CEV model is specified as follows:

$$dX(t) = \alpha_0 X(t)dt + \sigma_0 X(t)^{\beta_0/2}dW(t),$$

where $W(t)$ is a standard Brownian motion and α_0, σ_0, and β_0 are fixed constants.

Of note is that the interpretation of this model depends on β_0, i.e., in the case of stock prices, if $\beta_0 = 2$, then the price process $X(t)$ follows a lognormal diffusion; if $\beta_0 < 2$, then the model captures exactly the leverage effect as price and volatility are inversely correlated.

Amongst other authors, Beckers (1980) used this CEV model for stocks, Marsh and Rosenfeld (1983) apply a CEV parametrization to interest rates, and Emanuel and Macbeth (1982) utilize this setup for option pricing.

The *generalized constant elasticity of variance model is defined as follows*:

$$dX(t) = \left(\alpha_0 X(t)^{-(1-\beta_0)} + \lambda_0 X(t)\right)dt + \sigma_0 X(t)^{\beta_0/2}dW(t),$$

where the notation follows the CEV case. λ_0 is another parameter of the model. This process nests log diffusion when $\beta_0 = 2$ and nests square root diffusion when $\beta_0 = 1$.

Brennan and Schwartz (1979) and Courtadon (1982) analyze the model:

$$dX(t) = (\alpha_0 + \beta_0 X(t))dt + \sigma_0 X(t)^2 dW(t),$$

where α_0, β_0, σ_0 are fixed constants and $W(t)$ is a standard Brownian motion.

Duffie and Kan (1996) study the specification:

$$dX(t) = (\alpha_0 + X(t))dt + \sqrt{\beta_0 + \gamma_0 X(t)}dW(t),$$

where $W(t)$ is a standard Brownian motion and α_0, β_0, and γ_0 are fixed parameters.

Aït-Sahalia (1996) looks at a general case with general drift and CEV diffusion:

$$dX(t) = \left(\alpha_0 + \beta_0 X(t) + \gamma_0 X(t)^2 + \eta_0/X(t)\right)dt + \sigma_0 X(t)^{\beta_0/2}dW(t).$$

In the above expression, α_0, β_0, γ_0, η_0, σ_0, and β_0 are fixed constants and $W(t)$ is again a standard Brownian motion.

56.2.1.2 Diffusion Models with Jumps

For term structure modeling in empirical finance, the most widely studied class of models is the family of affine processes, including diffusion processes that incorporate jumps.

Affine Jumps Diffusion Model: $X(t_{-})$ is defined to follow an affine jump diffusion if

$$dX(t) = \kappa_0(\alpha_0 \pm X(t))dt + \Omega_0\sqrt{D(t)}dW(t) + dJ(t),$$

where $X(t_{-})$ is an N-dimensional vector of variables of interest and is a cadlag process, $W(t)$ is an N-dimensional independent standard Brownian motion, κ_0 and Ω_0 are square $N \times N$ matrices, α_0 is a fixed long-run mean, and $D(t)$ is a diagonal matrix with ith diagonal element given by

$$d_{ii}(t) = \theta_{0i} + \delta'_{0i}X(t).$$

In the above expressions, θ_{0i} and $\delta\prime_{0i}$ are constants. The jump intensity is assumed to be a positive, affine function of $X(t)$, and the jump size distribution is assumed to be determined by its conditional characteristic function. The attractive feature of this class of affine jump diffusions is that, as shown in Duffie et al. (2000), it has an exponential affine structure that can be derived in closed form, i.e.,

$$\Phi(X(t)) = \exp\left(a(t) + b(t)'X(t)\right),$$

where the functions $a(t)$ and $b(t)$ can be derived from Riccati equations.[6] Given a known characteristic function, one can use either GMM to estimate the

[6]For details, see Singleton (2006), p. 102.

parameters of this jump diffusion, or one can used quasi-maximum likelihood (QML), once the first two moments are obtained. In the univariate case without jumps, as a special case, this corresponds to the above general CIR model with jumps.

56.2.1.3 Multifactor and Stochastic Volatility Model

Multifactor models have been widely used in the literature, particularly in option pricing, term structure, and asset pricing. One general setup has $(X(t), V(t))' = (X(t), V^1(t), \ldots, V^d(t))'$ where only the first element, the diffusion process X_t, is observed while $V(t) = (V^1(t), \ldots, V^d(t))'_{d\times 1}$ is latent. In addition, $X(t)$ can be dependent on $V(t)$. For instance, in empirical finance, the most well-known class of the multifactor models is the stochastic volatility model expressed as

$$\begin{pmatrix} dX(t) \\ dV(t) \end{pmatrix} = \begin{pmatrix} b_1(X(t), \theta_0) \\ b_2(V(t), \theta_0) \end{pmatrix} dt + \begin{pmatrix} \sigma_{11}(V(t), \theta_0) \\ 0 \end{pmatrix} dW_1(t) + \begin{pmatrix} \sigma_{12}(V(t), \theta_0) \\ \sigma_{22}(V(t), \theta_0) \end{pmatrix} dW_2(t), \tag{56.4}$$

where $W_1(t)_{1\times 1}$ and $W_2(t)_{1\times 1}$ are independent standard Brownian motions and $V(t)$ is latent volatility process. $b_1(\cdot)$ is a function of $X(t)$ and $b_2(\cdot)$, $\sigma_{11}(\cdot)$, $\sigma_{22}(\cdot)$, and $\sigma_{22}(\cdot)$ are general functions of $V(t)$, such that system of Eq. 56.4 is well defined. Popular specifications are the square root model of Heston (1993), the GARCH diffusion model of Nelson (1990), lognormal model of Hull and White (1987), and the eigenfunction models of Meddahi (2001). Note that in this stochastic volatility case, the dimension of volatility is $d = 1$. More general setup can involve d driving Brownian motions in $V(t)$ equation.

As an example, Andersen and Lund (1997) study the generalized CIR model with stochastic volatility, specifically

$$dX(t) = \kappa_{x0}(\overline{x}_0 - X(t))dt + \sqrt{V(t)}dW_1(t),$$

$$dX(t) = \kappa_{v0}(\overline{v}_0 - V(t))dt + \sigma_{v0}\sqrt{V(t)}dW_2(t),$$

where $X(t)$ and $V(t)$ are price and volatility processes, respectively, $\kappa_{x0}, \kappa_{v0} > 0$ to ensure stationarity, $\overline{x}_0$ is the long-run mean of (log) price process, and $\overline{v}_0$ and σ_{v0} are constants. $W_1(t)$ and $W_2(t)$ are scalar Brownian motions. However, $W_1(t)$ and $W_2(t)$ are correlated such that $dW_1(t)dW_2(t) = \rho dt$ where the correlation ρ is some constant $\rho \in [-1, 1]$. Finally, note that volatility is a square root diffusion process, which requires that $\kappa_{v0}\overline{v}_0 > \sigma_{v0}^2$.

Stochastic Volatility Model with Jumps (*SVJ*): A standard specification is

$$dX(t) = \kappa_{x0}\left(\overline{x}_0 - X(t)\right)dt + \sqrt{V(t)}dW_1(t) + J_u dq_u - J_d dq_d,$$

$$dV(t) = \kappa_{v0}(\overline{v}_0 - V(t))dt + \sigma_{v0}\sqrt{V(t)}dW_2(t),$$

where q_u and q_d are Poisson processes with jump intensity parameters λ_u and $\lambda_{d,}$ respectively, and are independent of the Brownian motions $W_1(t)$ and $W_2(t)$. In particular, λ_u is the probability of a jump-up, $\Pr(dq_u\,(t) = 1) = \lambda_u$, and λ_d is the probability of a jump-down, $\Pr(dq_d(t) = 1) = \lambda_d$. J_u and J_d are jump-up and jump-down sizes and have exponential distributions: $f(J_u) = \frac{1}{\varsigma_u}\exp\left(-\frac{J_u}{\varsigma_u}\right)$ and $f(J_d) = \frac{1}{\varsigma_d}\exp\left(-\frac{J_d}{\varsigma_d}\right)$, where ς_u, $\varsigma_d > 0$ are the jump magnitudes, which are the means of the jumps, J_u and J_d.

Three-Factor Model (*CHEN*): The three-factor model combines various features of the above models, by considering a version of the oft examined three-factor model due to Chan et al. (1992), which is discussed in detail in Dai and Singleton (2000). In particular,

$$\begin{aligned} dX(t) &= \kappa_{x0}(\theta(t) - X(t))dt + \sqrt{V(t)}dW_1(t), \\ dV(t) &= \kappa_{v0}(\overline{v} - V(t))dt + \sigma_{v0}\sqrt{V(t)}dW_2(t), \\ d\theta(t) &= \kappa_{\theta 0}\left(\overline{\theta}(t) - \theta(t)\right)dt + \sigma_{\theta 0}\sqrt{\theta(t)}dW_3(t), \end{aligned} \tag{56.5}$$

where $W_1(t)$, $W_2(t)$ $W_3(t)$ are independent Brownian motions and V and θ are the stochastic volatility and stochastic mean of $X(t)$, respectively. κ_{x0}, κ_{v0}, $\kappa_{\theta 0}$, $\overline{v_0}$, $\overline{\theta_0}$, σ_{v0}, $\sigma_{\theta 0}$ are constants. As discussed above, non-negativity for $V(t)$ and $\theta(t)$ requires that $2\kappa_{v0}\overline{v}_0 > \sigma_{v0}^2$ and $2\kappa_{\theta 0}\overline{\theta}_0 > \sigma_{\theta 0}^2$.

Three-Factor Jump Diffusion Model (*CHENJ*): Andersen et al. (2004) extend the three-factor Chen (1996) model by incorporating jumps in the short rate process, hence improving the ability of the model to capture the effect of outliers and to address the finding by Piazzesi (2004, 2005) that violent discontinuous movements in underlying measures may arise from monetary policy regime changes. The model is defined as follows:

$$\begin{aligned} dX(t) &= \kappa_{x0}(\theta(t) - X(t))dt + \sqrt{V(t)}dW_1(t) + J_u dq_u - J_d dq_d, \\ dV(t) &= \kappa_{v0}\Big(\overline{v}_0 - V(t))dt + \sigma_{v0}\sqrt{V(t)}dW_2(t), \\ d\theta(t) &= \kappa_{\theta 0}\Big(\overline{\theta}_0 - \theta(t))dt + \sigma_{\theta 0}\sqrt{\theta(t)}dW_3(t) \end{aligned} \tag{56.6}$$

where all parameters are similar as in Eq. 56.5; $W_1(t)$, $W_2(t)$, and $W_3(t)$ are independent Brownian motions; and q_u and q_d are Poisson processes with jump intensities λ_{u0} and λ_{d0}, respectively, and are independent of the Brownian motions $W_r(t)$, $W_v(t)$, and $W_\theta(t)$. In particular, λ_{u0} is the probability of a jump-up, $\Pr(dq_u(t) = 1) = \lambda_{u0}$, and λ_{d0} is the probability of a jump-down, $\Pr(dq_d(t) = 1) = \lambda_{d0}$. J_u and J_d are jump-up and jump-down sizes and have exponential distributions $f(J_u) = \frac{1}{\varsigma_{u0}} \exp\left(-\frac{J_u}{\varsigma_{u0}}\right)$ and $f(J_d) = \frac{1}{\varsigma_{d0}} \exp\left(-\frac{J_d}{\varsigma_{d0}}\right)$, where $\varsigma_{u0}, \varsigma_{d0} > 0$ are the jump magnitudes, which are the means of the jumps J_u and J_d.

56.2.2 Overview on Specification Tests and Model Selection

The focus in this chapter is specification testing and model selection. The "tools" used in this literature have been long established. Several key classical contributions include the Kolmogorov-Smirnov test (see, e.g., Kolmogorov (1933) and Smirnov (1939)), various results on empirical processes (see, e.g., Andrews (1993) and the discussion in Chap. 19 of van der Vaart (1998) on the contributions of Glivenko, Cantelli, Doob, Donsker, and others), the probability integral transform (see, e.g., Rosenblatt (1952)), and the Kullback–Leibler information criterion (see, e.g., White (1982) and Vuong (1989)). For illustration, the empirical distribution mentioned above is crucial in our discussion of predictive densities because it is useful in estimation, testing, and model evaluation. Let Y_t is a variable of interest with distribution F and parameter θ_0. The theory of empirical distributions provides a result that

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} (1\{Y_t \leq u\} - F(u|\theta_0))$$

satisfies a central limit theorem (with a parametric rate) if T is large (i.e., asymptotically). In the above expression, $1\{Y_t \leq u\}$ is the indicator function which takes value 1 if $Y_t \leq u$ and 0 otherwise. In the case where there is parameter estimation error, we can use more general results in Chap. 19 of van der Vaart (1998). Define

$$P_T(f) = \frac{1}{T} \sum_{i=1}^{T} f(Y_i) \text{ and } P(f) = \int f dP,$$

where P is a probability measure associated with F. Here, $P_n(f)$ converges to $P(f)$ almost surely for all the measurable functions f for which $P(f)$ is defined. Suppose one wants to test the null hypothesis that P belongs to a certain family $\{P_{\theta_0} : \theta_0 \in \Theta\}$, where θ_0 is unknown; it is natural to use a measure of the discrepancy between P_n and $P_{\hat{\theta}}$ for a reasonable estimator $\hat{\theta}_t$ of θ_0. In particular, if $\hat{\theta}_t$ converges to θ_0 at a root-T rate, $\frac{1}{\sqrt{T}} \left(P_T - P_{\hat{\theta}t}\right)$ has been shown to satisfy a central limit theorem.[7]

With regard to dynamic misspecification and parameter estimation error, the approach discussed for the class of tests in this chapter allows for the construction

[7]See Theorem 19.23 in van der Vaart (1998) for details.

of statistics that admit for dynamic misspecification under both hypotheses. This differs from other classes of tests such as the framework used by Diebold et al. (1998), Hong (2001), and Bai (2003) in which correction dynamic specification under the null hypothesis is assumed. In particular, DGT use the probability integral transform to show that $F_t(Y_t|\Im_{t-1},\theta_0) = \int_{-\infty}^{Y_t} f_t(y|\Im_{t-1},\theta_0)dy$ is identically and independently distributed as a uniform random variable on [0; 1], where $F_t(\cdot)$ and $f_t(\cdot)$ are a parametric distribution and density with underlying parameter θ_0, Y_t is again our random variable of interest, and $\Im_t$ is the information set containing all "relevant" past information. They thus suggest using the difference between the empirical distribution of $F_t\left(Y_t\middle|\Im_{t-1},\hat{\theta}_t.\right)$ and the 45° line as a measure of "goodness of fit," where $\hat{\theta}_t$ is some estimator of θ_0. This approach has been shown to be very useful for financial risk management (see, e.g., Diebold et al. (1999)), as well as for macroeconomic forecasting (see, e.g., Diebold et al. (1998) and Clements and Smith (2000, 2002)). Similarly, Bai (2003) develops a Kolmogorov-type test of $F_t(Y_t|\Im_{t-1},\theta_0)$ on the basis of the discrepancy between $F_t\left(Y_t\middle|\Im_{t-1},\hat{\theta}_t.\right)$ and the CDF of a uniform on [0; 1]. As the test involves estimator $\hat{\theta}_t$, the limiting distribution reflects the contribution of parameter estimation error and is not nuisance parameter-free. To overcome this problem, Bai (2003) proposes a novel approach based on a martingalization argument to construct a modified Kolmogorov test which has a nuisance parameter-free limiting distribution. This test has power against violations of uniformity but not against violations of independence. Hong (2001) proposes another related interesting test, based on the generalized spectrum, which has power against both uniformity and independence violations, for the case in which the contribution of parameter estimation error vanishes asymptotically. If the null is rejected, Hong (2001) also proposes a test for uniformity robust to nonindependence, which is based on the comparison between a kernel density estimator and the uniform density. Two features differentiate the tests surveyed in this chapter from the tests outlined in the other papers mentioned above. First, the tests discussed here assume strict stationarity. Second, they allow for dynamic misspecification under the null hypothesis. The second feature allows us to obtain asymptotically valid critical values even when the conditioning information set does not contain all of the relevant past history. More precisely, assume that we are interested in testing for correct specification, given a particular information set which may or may not contain all of the relevant past information. This is important when a Kolmogorov test is constructed, as one is generally faced with the problem of defining $\Im_{t-1}$. If enough history is not included, then there may be dynamic misspecification. Additionally, finding out how much information (e.g., how many lags) to include may involve pre-testing, hence leading to a form of sequential test bias. By allowing for dynamic misspecification, such pre-testing is not required. Also note that critical values derived under correct specification given $\Im_{t-1}$ are not in general valid in the case of correct specification given a subset of $\Im_{t-1}$. Consider the following example. Assume that we are interested in testing whether the conditional distribution of $Y_t|Y_{t-1}$ follows normal distribution $N(\alpha_1 Y_{t-1}, \sigma_1)$. Suppose also that in actual fact the "relevant" information set has $\Im_{t-1}$ including both Y_{t-1} and Y_{t-2}, so that the true conditional model is $Y_t|\Im_{t-1} = Y_t|Y_{t-1}, Y_{t-2} = N(\alpha_1 Y_{t-1} + \alpha_2 Y_{t-2}, \sigma_2)$. In this case, correct specification holds with respect to the information contained in X_{t-1}; but there is dynamic misspecification with respect to Y_{t-1} and Y_{t-2}. Even without taking account of parameter estimation error, the critical values obtained assuming correct dynamic specification are invalid, thus leading to invalid inference. Stated differently, tests that are designed to have power against both uniformity and independence violations (i.e., tests that assume correct dynamic specification under the null) will reject an inference which is incorrect, at least in the sense that the "normality" assumption is not false. In summary, if one is interested in the particular problem of testing for correct specification for a given information set, then the approach of tests in this chapter is appropriate.

56.3 Consistent Distribution-Based Specification Tests and Predictive Density-Type Model Selection for Diffusion Processes

56.3.1 One-Factor Models

In this section, we outline the setup for the general class of one-factor jump diffusion specifications. All analyses carry through to the more complicated case of multifactor stochastic volatility models which we will elaborate upon in the next subsection. In the presentation of the tests, we follow a view that all candidate models, either single or multiple dimensional ones, are approximations of reality and can thus be misspecified. The issue of correct specification (or misspecification) of a single model and the model selection test for choosing amongst multiple competing models allow for this feature.

To begin, fix the time interval $[0, T]$ and consider a given single one-factor candidate model the same as Eq. 56.1, with the true parameters θ_0, λ_0, μ_0 to be replaced by its pseudo true analogs $\theta^{\dagger}$, λ, μ, respectively, and $0 \leq t \leq T$:

$$X(t_-) = \int_0^t b\left(X(s_-), \theta^{\dagger}\right) ds - \lambda t \int_Y y\phi(y) dy + \int_0^t \sigma\left(X(s_-), \theta^{\dagger}\right) dW(s) + \sum_{j=1}^{J_t} y_j,$$

or

$$\begin{aligned} dX(t-) = {} & \left(b\left(X(t-), \theta^{\dagger}\right) - \lambda\mu\right) dt. \\ & + \sigma\left(X(t-), \theta^{\dagger}\right) dW(t) + \int_Y y p(dy, dt), \end{aligned} \tag{56.7}$$

where variables are defined the same as in Eqs. 56.1 and 56.2. Note that as the above model is the one-factor version of Eqs. 56.1 and 56.2 where the dimension of $X(t_-)$ is 1×1, $W(t)$ is a one-dimensional standard Brownian motion and jump size, and y_j

is one-dimensional variable for all j. Also note both J_t and y_j are assumed to be independent of the driving Brownian motion.

If the single model is correctly specified, then $b(X(t-),\theta^{\dagger}) = b_0(X(t-),\theta_0)$, $\sigma(X(t-),\theta^{\dagger}) = \sigma_0(X(t-),\theta_0)$, $\lambda = \lambda_0$, $\mu = \mu_0$, and $\phi = \phi_0$ where $b_0(X(t-),\theta_0)$, $\sigma_0(X(t-),\theta_0)$, λ_0, μ_0, ϕ_0 are unknown and belong to the true specification.

Now consider a different case (not a single model) where m candidate models are involved. For model k with $1 \le k \le m$, denote its corresponding specification to be $(b_k(X(t_-),\theta_k^{\dagger}), \sigma_k(X(t_-),\theta_k^{\dagger}), \lambda_k, \mu_k, \phi_k)$. Two scenarios immediate arise. Firstly, if the model k is correctly specified, then $b_k(X(t_-),\theta_k^{\dagger}) = b_0(X(t_-),\theta_0)$, $\sigma_k(X(t_-),\theta_k^{\dagger}) = \sigma_0(X(t_-),\theta_0)$, $\lambda_k = \lambda_0$, $\mu_k = \mu_0$, and $\phi_k = \phi_0$ which are similar to the case of a single model. In the second scenario, all the models are likely to be misspecified and modelers are faced with the choice of selecting the "best" one. This type of problem is well fitted into the class of accuracy assessment tests initiated earlier by Diebold and Mariano (1995) or White (2000).

The tests discussed hereafter are Kolmogorov-type tests based on the construction of cumulative distribution functions (CDFs). In a few cases, the CDF is known in closed form. For instance, for the simplified version of the CIR model as in Eq. 56.3, $X(t)$ belongs to the linear exponential (or Pearson) family with the gamma CDF of the form[8]

$$F(u,\alpha,\lambda) = \frac{\int_0^u \left(\frac{\lambda}{2}\right)^{-2(1-\alpha/\lambda)-1} \exp\left(-x/\left(\frac{\lambda}{2}\right)\right)dx}{\Gamma(2(1-\alpha/\lambda))} \tag{56.8}$$

where $\Gamma(x) = \int_0^{\infty} t^x \exp(-t)dt$, and α, λ are constants.

Furthermore, if we look at the pure diffusion process without jumps

$$dX(t) = b\left(X(t),\theta^{\dagger}\right)dt + \sigma\left(X(t),\theta^{\dagger}\right)dW(t), \tag{56.9}$$

where $b(\cdot)$ and $\sigma = \sigma(\cdot)$ are drift and volatility functions, it is known that the stationary density, say $f(x,\theta^{\dagger})$, associated with the invariant probability measure can be expressed explicitly as[9]

$$f\left(x,\theta^{\dagger}\right) = \frac{c\left(\theta^{\dagger}\right)}{\sigma^2\left(x,\theta^{\dagger}\right)}\exp\left(\int^x \frac{2b\left(u,\theta^{\dagger}\right)}{\sigma^2\left(u,\theta^{\dagger}\right)}du\right),$$

where $c(\theta^{\dagger})$ is a constant ensuring that f integrates to one. The CDF, say $F(u,\theta^{\dagger}) = \int^u f(x,\theta^{\dagger})\, dx$, can then be obtained using available numerical integration procedures.

However, in most cases, it is impossible to derive the CDFs in closed form. To obtain a CDF in such cases, a more general approach is to use simulation. Instead of

[8]See Wong (1964) for details.

[9]See Karlin and Taylor (1981) for details.

estimating the CDF directly, simulation techniques estimate the CDF indirectly utilizing its generated sample paths and the theory of empirical distributions. The specification of a specific diffusion process will dictate the sample paths and thereby corresponding test outcomes.

Note that in the historical context, many early papers in this literature are probability density based. For example, in a seminal paper, Ait-Sahalia (1996) compares the marginal densities implied by hypothesized null models with nonparametric estimates thereof. Following the same framework of correct specification tests, Corradi and Swanson (2005) and Bhardwaj et al. (2008), however, do not look at densities. Instead, they compare the cumulative distribution (marginal or joint) implied by a hypothesized null model with the corresponding empirical distribution. While Corradi and Swanson (2005) focus on the known unconditional distribution, Bhardwaj et al. (2008) look at the conditional simulated distributions. Corradi and Swanson (2011) make extensions to multiple models in the context of out-of-sample accuracy assessment tests. This approach is somewhat novel to this continuous time model testing literature.

Now suppose we observe a discrete sample path $X_1, X_2, \ldots, X_T$ (also referred as skeletons).[10] The corresponding hypotheses can be set up as follows:

Hypothesis 1 Unconditional Distribution Specification Test of a Single Model

$H_0 : F(u,\theta^\dagger) = F_0(u,\theta_0)$, for all u, a.s.

$H_A : \Pr(F(u,\theta^\dagger) - F_0(u,\theta_0) \neq 0) > 0$, for some $u \in U$, with nonzero Lebesgue measure.

where $F_0(u, \theta_0)$ is the true cumulative distribution implied by the above density, i.e., $F_0(u,\theta_0) = \Pr(X_t \leq u)$. $F\left(u, \theta^\dagger\right) = \Pr\left(X_t^{\theta^\dagger} \leq u\right)$ is the cumulative distribution of the proposed model. $X_t^{\theta^\dagger}$ is a skeleton implied by model (56.7).

Hypothesis 2 Conditional Distribution Specification Test of a Single Model

$H_0 : F_\tau(u|X_t, \theta^\dagger) = F_{0,\tau}(u|X_t, \theta_0)$, for all u, a.s.

$H_A : \Pr(F_\tau(u|X_t, \theta^\dagger) - F_{0,\tau}(u|X_t, \theta_0) \neq 0) > 0$, for some $u \in U$, with nonzero Lebesgue measure.

where $F_\tau\left(u|X_t, \theta^\dagger\right) = \Pr\left(X_{t+\tau}^{\theta^\dagger} \leq u \middle| X_t^{\theta^\dagger} = X_t\right)$ is τ-step ahead conditional distributions and $t = 1, \ldots, T - \tau$. $F_{0,\tau}(u|X_t, \theta_0)$ is τ-step ahead true conditional distributions.

Hypothesis 3 Predictive Density Test for Choosing Amongst Multiple Competing Models

The null hypothesis is that no model can outperform model 1 which is the benchmark model.[11]

[10] As mentioned earlier, we follow Corradi and Swanson (2005) by using notation $X(\cdot)$ when defining continuous time processes and X_t for a skeleton.

[11] See White (2000) for a discussion of a discrete time series analog to this case, whereby point rather than density-based loss is considered; Corradi and Swanson (2007b) for an extension of White (2000) that allows for parameter estimation error; and Corradi and Swanson (2006) for an extension of Corradi and Swanson (2007b) that allows for the comparison of conditional distributions and densities in a discrete time series context.

H_0:

$$\max_{k=2,\ldots,m} E_X\left(\left(F_{X_{1,t+\tau}^{\theta_1^\dagger}(X_t)}(u_2) - F_{X_{1,t+\tau}^{\theta_1^\dagger}(X_t)}(u_1)\right)\right.$$
$$\left. -(F_0(u_2|X_t) - F_0(u_1|X_t))\right)^2$$
$$- E_X\left(\left(F_{X_{k,t+\tau}^{\theta_k^\dagger}(X_t)}(u_2) - F_{X_{k,t+\tau}^{\theta_k^\dagger}(X_t)}(u_1)\right)\right.$$
$$\left. -(F_0(u_2|X_t) - F_0(u_1|X_t))\right)^2 .$$

H_A: negation of H_0
where

$$F_{X_{k,t+\tau}^{\theta_k^\dagger}(X_t)}(u) = F_k^\tau\left(u\middle|X_t, \theta_k^\dagger\right) = P_{\theta_k^\dagger}^\tau\left(X_{k,t+\tau}^{\theta_k^\dagger} \le u\middle|X_t^{\theta_k^\dagger} = X_t\right),$$

which is the conditional distribution of $X_{t+\tau}$, given X_t, and evaluated at u under the probability law generated by model k. $X_{k,t+\tau}^{\theta_k^\dagger}(X_t)$ with $1 \le \tau \le T - t$ is the skeleton implied by model k, parameter $\theta_k^\dagger$, and initial value X_t. Analogously, define $F_0^\tau(u|X_t, \theta_0) = P_{\theta_0}^\tau(X_{t+\tau} \le u|X_t)$ to be the "true" conditional distribution.

Note that the three hypotheses expressed above apply exactly the same to the case of multifactor diffusions. Now, before moving to the statistics description section, we briefly explain the intuitions in facilitating construction of the tests.

In the first case (Hypothesis 1), Corradi and Swanson (2005) construct a Kolmogorov-type test based on comparison of the empirical distribution and the unconditional CDF implied by the specification of the drift, variance, and jumps. Specifically, one can look at the scaled difference between

$$F\left(u, \theta^\dagger\right) = \Pr\left(X_t^{\theta^\dagger} \le u\right) = \int^u f\left(x, \theta^\dagger\right)dx$$

and estimator of the true $F_0(u|X_t, \theta_0)$, the empirical distribution of X_t defined as

$$\frac{1}{T}\sum_{t=1}^{T} 1\{X_t \le u\},$$

where $1\{Y_t \le u\}$ is indicator function which takes value 1 if $Y_t \le u$ and 0 otherwise.

Similarly for the second case of conditional distribution (Hypothesis 2), the test statistic V_T can be a measure of the distance between the τ-*step* ahead conditional distribution of $X_{t+\tau}^{\theta^\dagger}$, given $X_t^{\theta^\dagger} = X_t$, as

$$F_\tau\left(u\middle|X_t,\theta^\dagger\right) = \Pr\left(X_{t+\tau}^{\theta^\dagger} \le u\middle|X_t^{\theta^\dagger} = X_t\right),$$

to an estimator of the true $F_{0,\tau}(u|X_t,\theta_0)$, the conditional empirical distribution of $X_{t+\tau}$ conditional on the initial value X_t defined as

$$\frac{1}{T-\tau}\sum_{t=1}^{T-\tau} 1\{X_{t+\tau} \le u\},$$

In the third case (Hypothesis 3), model accuracy is measured in terms of a distributional analog of mean square error. As is commonplace in the out-of-sample evaluation literature, the sample of T observations is divided into two subsamples, such that $T = R + P$, where only the last P observations are used for predictive evaluation. A τ-step ahead prediction error under model k is $1\{u_1 \le X_{t+\tau} \le u_2\} - (F_k^\tau(u_2|X_t,\ \theta_k^\dagger) - F_k^\tau(u_1|X_t,\ \theta_k^\dagger))$ where $2 \le k \le m$ and similarly for model 1 by replacing index k with index 1. Suppose we can simulate $P - \tau$ paths of τ-step ahead skeleton[12] using X_t as starting values where $t = R, \ldots, R + P - \tau$, from which we can construct a sample of $P - \tau$ prediction errors. Then, these prediction errors can be used to construct a test statistic for model comparison. In particular, model 1 is defined to be more accurate than model k if

$$\begin{aligned} &F\left(\left(\begin{array}{c}\left(F_1^\tau\left(u_2\middle|X_t,\theta_1^\dagger\right) - F_1^\tau\left(u_1\middle|X_t,\theta_1^\dagger\right)\right)\\ -\left(F_1^\tau(u_2|X_t,\theta_0) - F_1^\tau(u_1|X_t,\theta_0)\right)\end{array}\right)^2\right)\\ &\quad < E\left(\left(\begin{array}{c}\left(F_k^\tau\left(u_2^\tau\middle|X_t,\theta_1^\dagger\right) - F_k^\tau\left(u_1^\tau\middle|X_t,\theta_k^\dagger\right)\right)\\ -\left(F_0^\tau(u_2|X_t,\theta_0) - F_0^\tau(u_1|X_t,\theta_0)\right)\end{array}\right)^2\right),\end{aligned}$$

where $E(\cdot)$ is an expectation operator and $E(1\{u_1 \le X_{t+\tau} \le u_2\}|X_t) = F_0^\tau(u_2|X_t,\theta_0) - F_0^\tau(u_1|X_t,\theta_0)$. Concretely, model k is worse than model 1 if on average τ-step ahead prediction errors under model k is larger than that of model 1.

Finally, it is important to point out some main features characterized by all the three test statistics. Processes $X(t)$ hereafter are required to satisfy the regular conditions, i.e., assumptions A1–A8 in Corradi and Swanson (2011). Regarding model estimation (in Sect. 56.3.3), $\theta^\dagger$ and $\theta_k^\dagger$ are unobserved and need to be estimated. While Corradi and Swanson (2005) and Bhardwaj et al. (2008) utilize (recursive) simulated generalized method of moments (SGMM), Corradi and Swanson (2011) make extension to (recursive) nonparametric simulated

[12]See Sect. 56.3.3.1 for model simulation details.

quasi-maximum likelihood (NPSQML). For the unknown distribution and conditional distribution, it will be pointed out in Sect. 56.3.3.2 that $F(u,\theta^{\dagger})$, $F_{\tau}(u|X_t, \theta^{\dagger})$, and $F_{X_{k,t+\tau}^{\theta_k^{\dagger}}(X_t)}(u)$ can be replaced by their simulated counterparts using the (recursive) SGMM and NPSQML parameter estimators. In addition, test statistics converge to functional of Gaussian processes with covariance kernels that reflect time dependence of the data and the contribution of parameter estimation error (PEE). Limiting distributions are not nuisance parameter-free, and critical values thereby cannot be tabulated by the standard approach. All the tests discussed in this chapter rely on the bootstrap procedures for obtaining the asymptotically valid critical values, which we will describe in Sect. 56.3.4.

56.3.1.1 Unconditional Distribution Tests

For one-factor diffusions, we outline the construction of unconditional test statistics in the context where CDF is known in closed form. In order to test the **Hypothesis 1**, consider the following statistic:

$$V_{T,N,h}^2 = \int_U V_{T,N,h}^2(u)\pi(u),$$

where

$$V_{T,N,h} = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(1\{X_t \le u\} - F\left(u, \hat{\theta}_{T,N,h}\right)\right).$$

In the above expression, U is a compact interval and $\int_U \pi(u)du = 1, 1\{X_t \le u\}$ is again the indicator function which returns value 1 if $X_t \le u$ and 0 otherwise. Further, as defined in Sect. 56.3.3, $\hat{\theta}_{T,N,h}$ hereafter is a simulated estimator where T is sample size and h is the discretization interval used in simulation. In addition, with the abuse of notation, N is a generic notation throughout this chapter, i.e., $N = L$, the length of each simulation path for (recursive) SGMM, and $N = M$, the number of random draws (simulated paths) for (recursive) NPQML estimator.[13] Also note in our notation that as the above test is in sample specification test, the estimator and the statistics are constructed using the entire sample, i.e., $\hat{\theta}_{T,N,h}$.

It has been shown in Corradi and Swanson (2005) that under regular conditions and if the estimator is estimated by SGMM, the above statistics converges to a functional of Gaussian process.[14] In particular, pick the choice T, $N \to \infty$, $h \to 0$, $T/N \to 0$, and $Th^2 \to 0$.

[13] M is often chosen to coincide with S, the number of simulated paths used when simulating distributions.

[14] For details and the proof, see Theorem 1 in Corradi and Swanson (2005).

Under the null,

$$V_{T,N,h}^2 \to \int_U Z^2(u)\pi(u),$$

where Z is a Gaussian process with covariance kernel. Hence, the limiting distribution of $V_{T,N,h}^2$ is a functional of a Gaussian process with a covariance kernel that reflects both PEE and the time series nature of the data. As $\hat{\theta}_{T,N,h}$ is root-T consistent, PEE does not disappear in the asymptotic covariance kernel.

Under H_A, there exists an $\varepsilon > 0$ such that

$$\lim_{T\to\infty} \Pr\left(\frac{1}{T}V_{T,N,h}^2 > \varepsilon\right) = 1.$$

For the asymptotic critical value tabulation, we use the bootstrap procedure. In order to establish validity of the block bootstrap under SGMM with the presence of PEE, the simulated sample size should be chosen to grow at a faster rate than the historical sample, i.e., $T/N \to 0$.

Thus, we can follow the steps in appropriate bootstrap procedure in Sect. 56.3.4. For instance, if the SGMM estimator is used, the bootstrap statistic is

$$V_{T,N,h}^{2*} = \int\limits_U V_{T,N,h}^{2*}(u)\pi(u)du,$$

where

$$V_{T,N,h}^{2*} = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\Big(\left(1\left\{X_t^* \le u\right\} - 1\{X_t \le u\}\right)$$
$$-\left(F\left(u, \hat{\theta}_{T,N,h}^*\right) - F\left(u, \hat{\theta}_{T,N,h}\right)\right)\Big).$$

In the above expression, $\hat{\theta}_{T,N,h}^*$ is the bootstrap analog of $\hat{\theta}_{T,N,h}^*$ and is estimated by the bootstrap sample $X_1^*, \ldots, X_T^*$ (see Sect. 56.3.4). With appropriate conditions, Corradi and Swanson (2005) show that under the null, $V_{T,N,h}^{2*}$ has a well-defined limiting distribution which coincides with that of $V_{T,N,h.}^2$ We then can straightforwardly derive the bootstrap critical value by following Steps 1–5 in Sect. 56.3.4. In particular, in Step 5, the idea is to perform B bootstrap replications (B large) and compute the percentiles of the empirical distribution of the B bootstrap statistics. Reject H_0 if $V_{T,N,h}^2$ is greater than the $(1-\alpha)$th percentile of this empirical distribution. Otherwise, do not reject H_0.

56.3.1.2 Conditional Distribution Tests

Hypothesis 2 tests correct specification of the conditional distribution, implied by a proposed diffusion model. In practice, the difficulty arises from the fact that the functional form of neither τ-step ahead conditional distributions $F_\tau(u|X_t, \theta^\dagger)$ nor $F_{0,\tau}(u|X_t, \theta_0)$ is unknown in most cases. Therefore, Bhardwaj et al. (2008) develop bootstrap specification test on the basis of simulated distribution using the SGMM estimator.[15] With the important inputs leading to the test such as simulated estimator, distribution simulation, and bootstrap procedures to be presented in the next section,[16] the test statistic is defined as

$$Z_T = \sup_{u\times v\in U\times V} |Z_T(u, v)|,$$

where

$$Z_T(u, v) = \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} \left(\begin{array}{c} \frac{1}{S}\sum_{s=1}^{S} 1\left\{X_{s,t+\tau}^{\hat{\theta}_{T,N,h}} \le u\right\} \\ -(1\{X_{t+\tau} \le u\}) \end{array} \right),$$

with U and V compact sets on the real line. $\hat{\theta}_{T,N,h}$ is the simulated estimator using entire sample $X_1, \ldots, X_T$, and S is the number of simulated replications used in the estimation of conditional distributions as described in Sect. 56.3.3. If SGMM estimator is used (similar to unconditional distribution case and the same as in Bhardwaj et al. (2008)), then $N = L$, where L is the simulation length used in parameter estimation.

The above statistic is a simulation-based version of the conditional Kolmogorov test of Andrews (1997), which compare the joint empirical distribution:

$$\frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} 1\{X_{t+\tau} \le u\} 1\{X_t \le v\},$$

with its semi-empirical/semi-parametric analog given by the product of

$$\frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} F_{0,\tau}(u|X_t, \theta_0) 1\{X_t \le v\}.$$

Intuitively, if the null is not rejected, the metric distance between the two should asymptotically disappear. In the simulation context with parameter estimation error, the asymptotic limit of Z_T however is a nontrivial one. Bhardwaj et al. (2008) show that with the proper choice of T, N, S, h, i.e., T, N, S, $T^2/S \to \infty$ and h, T/N, T/S, Nh, $h^2T \to 0$, then

$$Z_T \xrightarrow{d} \sup_{u\times v\in U\times V} |Z(u, v)|,$$

where $Z(u, v)$ is a Gaussian process with a covariance kernel that characterizes (1) long-run variance we would have if we knew $F_{0,\tau}(u|X_t, \theta_0)$), (2) the contribution

[15] In this chapter, we assume that $X(\cdot)$ satisfies the regularity conditions stated in Corradi and Swanson (2011), i.e., assumptions A1–A8. Those conditions also reflect requirements A1–A2 in Bhardwaj et al. (2008). Note that the SGMM estimator used in Bhardwaj et al. (2008) satisfies the root-N consistency condition that Corradi and Swanson (2011) impose on their parameter estimator (see Assumption 4).

[16] See Sects. 56.3.3 and 56.3.4 for further details.

of parameter estimation error, and (3) The correlation between the first two.

Furthermore, under H_A, there exists some $\varepsilon > 0$ such that

$$\lim_{P\to\infty} \Pr\left(\frac{1}{\sqrt{T}} Z_T > \varepsilon\right) = 1.$$

As $T/S \to 0$, the contribution of simulation error is asymptotically negligible. The limiting distribution is not nuisance parameter-free and hence critical values cannot be tabulated directly from it. The appropriate bootstrap statistic in this context is

$$Z_T^* = \sup_{u\times v\in U\times V} \left|Z_T^*(u,v)\right|,$$

where

$$\begin{aligned} Z_T^* = (u,v) = {} & \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} 1\left\{X_t^* \le v\right\} \\ & \times \left(\frac{1}{S}\sum_{s=1}^{S} 1\left\{X_{s,t+\tau}^{\hat{\theta}^*_{T,N,h}} \le u\right\} - 1\left\{X_{t+\tau}^* \le u\right\}\right) \\ & - \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} 1\{X_t \le v\} \\ & \times \left(\frac{1}{S}\sum_{s=1}^{S} 1\left\{X_{s,t+\tau}^{\hat{\theta}^*_{T,N,h}} \le u\right\} - 1\{X_{t+\tau} \le u\}\right). \end{aligned}$$

In the above expression, $\hat{\theta}^*{}_{T,N,h}$ is the bootstrap parameter estimated using the resampled data X_t^* for $t = 1, \ldots, T-\tau$. $X_{s,t+\tau}^{\hat{\theta}^*_{T,N,h}}$, $s = 1, \ldots, S$ and $t = 1, \ldots, T-\tau$ is the simulated date under $\hat{\theta}^*_{T,N,h}$ and X_t^*; $t = 1, \ldots, T-\tau$ is a resampled series constructed using standard block bootstrap methods as described in Sect. 56.3.4. Note that in the original paper, Bhardwaj et al. (2008) propose bootstrap SGMM estimator for conditional distribution of diffusion processes. Corradi and Swanson (2011) extend the test to the case of simulated recursive NPSQML estimator. Regarding the generation of the empirical distribution of Z_T^* (asthmatically the same as Z_T), follow Steps 1–5 in the bootstrap procedure in Sect. 56.3.4. This yields B bootstrap replications (B large) of Z_T^*. One can then compare Z_T with the percentiles of the empirical distribution of Z_T^*, and reject H_0 if Z_T is greater than the $(1-\alpha)$th percentile. Otherwise, do not reject H_0. Tests carried out in this manner are correctly asymptotically sized and have unit asymptotic power.

56.3.1.3 Predictive Density Tests for Multiple Competing Models

In many circumstances, one might want to compare one (benchmark) model (model 1) against multiple competing models (models k, $2 \le k \le m$). In this case, recall in the null in **Hypothesis 3** that no model can outperform the benchmark model. In testing the null, we first choose a particular interval, i.e., $(u_1, u_2) \in U \times U$, where U is a compact set so that the objective is evaluation of predictive densities for a given range of values. In addition, in the recursive setting (not full sample is used to estimate parameters), if we use the recursive NPSQML estimator, say $\hat{\theta}_{1.t,N,h}$ and $\hat{\theta}_{k.t,N,h}$, for models 1 and k, respectively, then the test statistic is defined as

$$D^{Max}_{k,P,S}(u_1,u_2) = \max_{k=2,\ldots,m} D_{k,P,S}(u_1,u_2),$$

where

$$\begin{aligned} D_{k,P,S}(u_1,u_2) = \frac{1}{\sqrt{P}} \sum_{t=R}^{T-\tau} \Bigg(\Bigg[\frac{1}{S} \sum_{i=1}^{S} 1\left\{ u_1 \le X^{\hat{\theta}_{1,t,N,h}}_{1,i,t+\tau}(X_t) \le u_2 \right\} \\ - 1\{u_1 \le X_{t+\tau} \le u_2\} \Bigg]^2 \\ - \Bigg(\Bigg[\frac{1}{S} \sum_{i=1}^{S} 1\left\{ u_1 \le X^{\hat{\theta}_{k,t,N,h}}_{k,i,t+\tau}(X_t) \le u_2 \right\} \\ - \{u_1 \le X_{t+\tau} \le u_2\} \Bigg]^2 \Bigg). \end{aligned}$$

All notations are consistent with previous sections where S is the number of simulated replications used in the estimation of conditional distributions. $X^{\hat{\theta}_{1,t,N,h}}_{1,i,t+\tau}(X_t)$ and $X^{\hat{\theta}_{k,t,N,h}}_{k,i,t+\tau}$, $i = 1, \ldots, S$, $t = 1, \ldots, T-\tau$, are the ith simulated path under $\hat{\theta}_{1,t,N,h}$ and $\hat{\theta}_{k,t,N,h}$. If models 1 and k are nonnested for at least one, $k = 2, \ldots, m$. Under regular conditions and if P, R, S, h are chosen such as $P, R, N \to \infty$ and h, P/N, $h^2P \to 0$, $P/R \to \pi$ where π is finite then

$$\max_{k=2,\ldots,m} \left(D_{k,P,N}(u_1,u_2) - \mu_k(u_1,u_2) \right) \to \max_{k=2,\ldots,m} Z_k(u_1,u_2).$$

where, with an abuse of notation, $\mu_k(u_1, u_2) = \mu_1(u_1, u_2) - \mu_k(u_1, u_2)$, and

$$\mu_j(u_1,u_2) = E\left(\left(\begin{pmatrix} F_{X^{\theta_j^\dagger}_{j,t+\tau}(X_t)}(u_2) - F_{X^{\theta_j^\dagger}_{j,t+\tau}(X_t)}(u_1) \\ -(F_0(u_2|X_t) - F_0(u_1|X_t)) \end{pmatrix} \right)^2 \right),$$

for $j = 1,\ldots, m$, and where $(Z_1(u_1, u_2), \ldots, Z_m(u_1, u_2))$ is an m-dimensional Gaussian random variable the covariance kernels that involves error in parameter estimation. Bootstrap statistics are therefore required to reflect this parameter estimation error issue.[17]

In the implementation, we can obtain the asymptotic critical value using a recursive version of the block bootstrap. The idea is that when forming block

[17]See Corradi and Swanson (2011) for further discussion.

bootstrap samples in the recursive setting, observations at the beginning of the sample are used more frequently than observations at the end of the sample. We can replicate Steps 1–5 in bootstrap procedure in Sect. 56.3.4. It should be stressed that the resampling in the Step 1 is the recursive one. Specifically, begin by resampling b blocks of length l from the full sample, with $lb = T$. For any given τ, it is necessary to jointly resample $X_t, X_{t+1}, \ldots, X_{t+\tau}$. More precisely, let $Z^{t,\tau} = (X_t, X_{t+1}, \ldots, X_{t+\tau})$, $t = 1, \ldots, T-\tau$. Now, resample b overlapping blocks of length l from $Z^{t,\tau}$. This yields $Z^{t,*} = \left(X_t^*, X_{t+1}^*, \ldots, X_{t+\tau}^*\right)$, $t = 1, \ldots, T-\tau$. Use these data to construct bootstrap estimator $\hat{\theta}_{k,t,N,h}$. Recall that N is chosen in Corradi and Swanson (2011) as the number of simulated series used to estimate the parameters ($N = M = S$) and such as N/R, $N/P \to \infty$. Under this condition, simulation error vanishes and there is no need to resample the simulated series.

Corradi and Swanson (2011) show that

$$\frac{1}{\sqrt{P}}\sum_{t=R}^{T}\left(\hat{\theta}^*_{k,t,N,h} - \hat{\theta}_{k,t,N,h}\right)$$

has the same limiting distribution as

$$\frac{1}{\sqrt{P}}\sum_{t=R}^{T}\left(\hat{\theta}_{k,t,N,h} - \theta_k^\dagger\right),$$

conditional on all samples except a set with probability measure approaching zero. Given this, the appropriate bootstrap statistic is

$$\begin{aligned}
D^*_{k,P,S}(u_1,u_2) = \frac{1}{\sqrt{P}}\sum_{t=R}^{T-\tau}\Bigg\{ & \Bigg(\Bigg[\frac{1}{S}\sum_{i=1}^{S} 1\left\{u_1 \le X_{1,i,t+\tau}^{\hat{\theta}_{1,t,N,h}}\left(X_t^*\right) \le u_2\right\} \\
& - 1\left\{u_1 \le X^*_{t+\tau} \le u_2\right\}\Bigg]^2 \\
& - \left(\frac{1}{T}\sum_{j=1}^{T}\left[\frac{1}{S}\sum_{i=1}^{S} 1\left\{u_1 \le X_{1,i,t+\tau}^{\hat{\theta}_{1,t,N,h}}\left(X_j^*\right) \le u_2\right\}\right.\right. \\
& \left.\left. - 1\left\{u_1 \le X_{j+\tau} \le u_2\right\}\right]^2\right. \\
& - \Bigg(\Bigg[\frac{1}{S}\sum_{i=1}^{S} 1\left\{u_1 \le X_{k,i,t+\tau}^{\hat{\theta}_{k,t,N,h}}\left(X_t^*\right) \le u_2\right\} \\
& - 1\left\{u_1 \le X^*_{t+\tau} \le u_2\right\}\Bigg]^2 \\
& - \left(\frac{1}{S}\sum_{i=1}^{S}\left[\frac{1}{S}\sum_{i=1}^{S} 1\left\{u_1 \le X_{k,i,t+\tau}^{\hat{\theta}_{k,t,N,h}}\left(X_j\right) \le u_2\right\}\right.\right. \\
& -1\left\{u_1 \le X_{j+\tau} \le u_2\right\}\Bigg]^2\Bigg)\Bigg)\Bigg\}.
\end{aligned}$$

As the bootstrap statistic is calculated from the last P resampled observations, it is necessary to have each bootstrap term recentered around the (full) sample mean. This is true even in the case there is no need to mimic PEE, i.e., the choice of P, R is such that $P/R \to 0$. In such a case, above statistic can be formed using $\hat{\theta}_{k,t,N,h}$ rather than $\hat{\theta}^*_{k,t,N,h}$.

For any bootstrap replication, repeat B times (B large) bootstrap replications which yield B bootstrap statistics $D^*_{k,P,S}$. Reject H_0 if $D_{k,P,S}$ is greater than the $(1 - \alpha)$th percentile of the bootstrap empirical distribution. For numerical implementation, it is of importance to note that in the case where $P/R \to 0, P, T, R \to \infty$, there is no need to re-estimate $\hat{\theta}^*_{1,t,N,h}\left(\hat{\theta}^*_{k,t,N,h}\right)$. Namely, $\hat{\theta}_{1,t,N,h}\left(\hat{\theta}_{k,t,N,h}\right)$ can be used in all bootstrap experiments.

Of course, the above framework can also be applied using entire simulated distributions rather than predictive densities, by simply estimating parameters once, using the entire sample, as opposed to using recursive estimation techniques, say, as when forming predictions and associated predictive densities.

56.3.2 Multifactor Models

Now, let us turn our attention to multifactor diffusion models of the form $(X(t), V(t))' = (X(t), V^1(t), \ldots, V^d(t))'$, where only the first element, the diffusion process X_t, is observed while $V(t) = (V^1(t), \ldots, V^d(t))'$ is latent. The most popular class of the multifactor models is stochastic volatility model expressed as below:

$$\begin{pmatrix} dX(t) \\ dV(t) \end{pmatrix} = \begin{pmatrix} b_1\left(X(t), \theta^\dagger\right) \\ b_2\left(V(t), \theta^\dagger\right) \end{pmatrix} dt + \begin{pmatrix} \sigma_{11}\left(V(t), \theta^\dagger\right) \\ 0 \end{pmatrix} dW_1(t) + \begin{pmatrix} \sigma_{12}\left(V(t), \theta^\dagger\right) \\ \sigma_{22}\left(V(t), \theta^\dagger\right) \end{pmatrix} dW_2(t), \tag{56.10}$$

where $W_1(t)_{1\times1}$ and $W_2(t)_{1\times1}$ are independent Brownian motions.[18] For instance, many term structure models require the multifactor specification of the above form (see Dai and Singleton (2000)). In a more complicated case, the drift function can also be specified to be a stochastic process which poses even more challenges to testing. As mentioned earlier, the hypotheses (**Hypothesis 1**, **2**, **3**) and the test

[18]Note that the dimension of $X(\cdot)$ can be higher and we can add jumps to the above specification such that it satisfies the regularity conditions outlined in the one-factor case. In addition, Corradi and Swanson (2005) provide a detailed discussion of approximation schemes in the context of stochastic volatility models.

construction strategy for multifactor models are the same as for one-factor model. All theory essentially applies immediately to multifactor cases. In implementation, the key difference is in the simulated approximation scheme facilitating parameter and CDF estimation. $X(t)$ cannot simply be expressed as a function of $d + 1$ driving Brownian motions but instead involves a function of $(W_{jt}, \int_0^t W_{js} dW_{is})$, $i, j = 1, \ldots, d + 1$ (see, e.g., Pardoux and Talay (1985, pp. 30–32) and Corradi and Swanson (2005)).

For illustration, we hereafter focus on the analysis of a stochastic volatility model (56.10) where drift and diffusion coefficients can be written as

$$
\begin{aligned}
b &= \begin{pmatrix} b_1\left(X(t), \theta^\dagger\right) \\ b_2\left(V(t), \theta^\dagger\right) \end{pmatrix} \\
\sigma &= \begin{pmatrix} \sigma_{11}\left(V(t), \theta^\dagger\right) & \sigma_{12}\left(V(t), \theta^\dagger\right) \\ 0 & \sigma_{22}\left(V(t), \theta^\dagger\right) \end{pmatrix}
\end{aligned}
$$

We also examine a three-factor model (i.e., the Chen model as in Eq. 56.5) and a three-factor model with jumps (i.e., CHENJ as in Eq. 56.6). By presenting two- and three-factor models as an extension of our above discussion, we make it clear that specification tests of multiple factor diffusions with $d \geq 3$ can be easily constructed in similar manner.

In distribution estimation, the important challenge for multifactor models lies in the missing variable issue. In particular, for simulation of X_t, one needs initial values of the latent processes $V_1, \ldots, V_d$, which are unobserved. To overcome this problem, it suffices to simulate the process using different random initial values for the volatility process; then construct the simulated distribution using those initial values and average them out. This allows one to integrate out the effect of a particular choice of volatility initial value.

For clarity of exposition, we sketch out a simulation strategy for a general model of d latent variables in Sect. 56.3.3. This generalizes the simulation scheme of three-factor models in Cai and Swanson (2011). As a final remark before moving to the statistic presentation, note that the class of multifactor diffusion processes considered in this chapter is required to match the regular conditions as in previous section (assumption from A1 to A8 in Corradi and Swanson (2011) with A4 being replaced by A4′).

56.3.2.1 Unconditional Distribution Tests

Following the above discussion on test construction, we specialize to the case of two-factor stochastic volatility models. Extension to general multidimensional and multifactor models follows similarly. As the CDF is rarely known in closed form for stochastic volatility models, we rely on simulation technique. With the simulation scheme, estimators, simulated distribution, and bootstrap procedures to be presented in the next sections (see Sects. 56.3.3 and 56.3.4), the test statistics for **Hypothesis 1** turns out to be

$$SV_{T,S,h} = \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(1\{X_t \le u\} - \frac{1}{S}\sum_{t=1}^{S}1\left(X_{t,h}^{\hat{\theta}_{T,N,L,h}} \le u\right)\right).$$

In the above expression, recall that S is the number of simulation paths used in distribution simulation; $\hat{\theta}_{T,N,L,h}$ is a simulated estimator (see Sect. 56.3.3). N is a generic notation throughout this chapter, i.e., $N = L$, the length of each simulation path for SGMM, and $N = M$, the number of random draws (simulated paths) for NPQML estimator. h is the discretization interval used in simulation. Note that $\hat{\theta}_{T,N,L,h}$ is chosen in Corradi and Swanson (2005) to be SGMM estimator using full sample and therefore $N = L = S$.[19] To put it simply, one can write $\hat{\theta}_{T,S,h} = \hat{\theta}_{T,N,L,h}$.

Under the null, choose T, S to satisfy $T, S \to \infty$, $Sh \to 0$, $T/S \to 0$, then

$$SV_{T,S,h}^2 \to \int_U SV^2(u)\pi(u),$$

where Z is a Gaussian process with covariance kernel that reflects both PEE and the time-dependent nature of the data. The relevant bootstrap statistic is

$$\begin{aligned} SV_{T,S,h}^{2*} = &\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(1\{X_t^* \le u\} - 1\{X_t \le u\}\right) \\ &- \frac{1}{\sqrt{T}}\sum_{t=1}^{T}\frac{1}{S}\sum_{t=1}^{S}\left(1\left(X_{t,h}^{\hat{\theta}_{T,N,L,h}} \le u\right) - 1\left(X_{t,h}^{\hat{\theta}_{T,N,L,h}} \le u\right)\right), \end{aligned}$$

where $\hat{\theta}^*_{T,S,h}$ is the bootstrap analog of $\hat{\theta}_{T,S,h}$. Repeat Steps 1–5 in the bootstrap procedure in Sect. 56.3.4 to obtain critical values which are the percentiles of the empirical distribution of Z_T^*. Compare $SV_{T,S,h}$ with the percentiles of the empirical distribution of the bootstrap statistic and reject H_0 if $SV_{T,S,h}$ is greater than the $(1-\alpha)$th percentile thereof. Otherwise, do not reject H_0.

56.3.2.2 Conditional Distribution Tests

To test **Hypothesis 2** for the multifactor models, first we present the test statistic for the case of the stochastic volatility model (X_t, V_t) in Eq. 56.10 (i.e., for two-factor diffusion), and then we discuss testing with the three-factor model (X_t, V_t^1, V_t^2) as in Eq. 56.5. Other multiple factor models can be tested analogously.

[19] As seen in assumption A4′ in Corradi and Swanson (2011) and Sect. 56.3.3 of this chapter, $\hat{\theta}_{T,N,L,h}$ can be other estimators such as the NPSQML estimator. Importantly, $\hat{\theta}_{T,N,L,h}$ satisfies condition A4′ in Corradi and Swanson (2011).

Note that for illustration, we again assume use of the SGMM estimator $\hat{\theta}_{T,N,L,h}$, as in the original work of Bhardwaj et al. (2008) (namely, $\hat{\theta}_{T,N,L,h}$ is the simulated estimator described in Sect. 56.3.3). Specifically, N is chosen as the length of sample path L used in parameter estimation. The associated test statistic is

$$\begin{aligned} SZ_T &= \sup_{u\times v\in U\times V} |SZ_T(u,v)| \\ SZ_T(u,v) &= \frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau} 1\{X_t \le v\} \\ &\quad \times \left(\frac{1}{NS}\sum_{j=1}^{N}\sum_{i=1}^{S} 1\left\{X_{j,i,t+\tau}^{\hat{\theta}_{T,N,L,h}} \le u\right\} - 1\{X_{t+\tau} \le u\}\right), \end{aligned}$$

where $X_{j,i,t+\tau}^{\hat{\theta}_{T,N,L,h}}$ is τ-step ahead simulated skeleton obtained by simulation procedure for multifactor model in Sect. 56.3.3.1.

In a similar manner, the bootstrap statistic analogous to SZ_T is

$$\begin{aligned} SZ_T^* &= \sup_{u\times v\in U\times V} \left|SZ_T^*(u,v)\right|, \\ SZ_T^*(u,v) &= \frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau} 1\left\{X_t^* \le v\right\} \\ &\quad \times \left(\frac{1}{NS}\sum_{j=1}^{N}\sum_{i=1}^{S} 1\left\{X_{j,i,t+\tau}^{\hat{\theta}^*_{T,N,L,h}} \le u\right\} - 1\left\{X_{t+\tau}^* \le u\right\}\right) \\ &\quad - \frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau} 1\{X_t \le v\} \\ &\quad \left(\frac{1}{NS}\sum_{j=1}^{N}\sum_{i=1}^{S} 1\left\{X_{j,i,t+\tau}^{\hat{\theta}_{T,N,L,h}} \le u\right\} - 1\{X_{t+\tau} \le u\}\right) \end{aligned}$$

where $\hat{\theta}^*_{T,N,L,h}$ is the bootstrap estimator described in Sect. 56.3.4. For the three-factor model, the test statistic is defined as

$$\begin{aligned} MZ_T &= \sup_{u\times v\in U\times V} |MZ_T(u,v)|, \\ MZ_T(u,v) &= \frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau} 1\{X_t \le v\} \\ &\quad \times \left(\frac{1}{L^2 S}\sum_{j=1}^{L}\sum_{k=1}^{L}\sum_{i=1}^{S} 1\left\{X_{s,t+\tau}^{\hat{\theta}_{T,N,L,h}} \le u\right\} - 1\{X_{t+\tau} \le u\}\right), \end{aligned}$$

and bootstrap statistics is

$$\begin{aligned}MZ_T^*(u,v) &= \frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}1\{X_t \le v\} \\ &\quad \times\left(\frac{1}{L^2S}\sum_{j=1}^{L}\sum_{k=1}^{L}\sum_{i=1}^{S}1\left\{X_{s,t+\tau}^{\hat{\theta}^*_{t,N,L,h}} \le u\right\} - 1\{X^*_{t+\tau} \le u\}\right) \\ &\quad - \frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau}1\{X_t \le v\} \\ &\quad \times\left(\frac{1}{L^2S}\sum_{j=1}^{L}\sum_{k=1}^{L}\sum_{i=1}^{S}1\left\{X_{s,t+\tau}^{\hat{\theta}_{t,N,L,h}} \le u\right\} - 1\{X_{t+\tau} \le u\}\right)\end{aligned}$$

where $X_{s,t+\tau}^{\hat{\theta}_{T,N,L,h}} = X_{s,t+\tau}^{\hat{\theta}_{T,N,L,h}}\left(X_t, V_j^{1,\hat{\theta}_{T,N,L,h}}, V_k^{2,\hat{\theta}_{T,N,L,h}}\right)$ and $X_{s,t+\tau}^{\hat{\theta}^*_{t,N,L,h}} = X_{s,t+\tau}^{\hat{\theta}^*_{t,N,L,h}}\left(X_t, V_j^{1,\hat{\theta}^*_{t,N,L,h}}, V_k^{2,\hat{\theta}^*_{t,N,L,h}}\right)$.

The first-order asymptotic validity of inference carried out using bootstrap statistics formed as outlined above follows immediately from Bhardwaj et al. (2008). For testing decision, one compares the test statistics $SZ_{T,S,h}$ and $MZ_{T,S,h}$ with the percentiles or the empirical distributions of SZ_T^* and $MZ_{T,S,h}^*$, respectively. Then, reject H_0 if the actual statistic is greater than the $(1-\alpha)$th percentile of the empirical distribution of the bootstrap statistic, as in Sect. 56.3.4. Otherwise, do not reject H_0.

56.3.2.3 Predictive Density Tests for Multiple Competing Models

For illustration, we present the test for the stochastic volatility model (two-factor model). Again, note that extension to other multifactor models follows immediately. In particular, all steps in the construction of the test in the one-factor model case carry through immediately to the stochastic volatility case with the statistic defined as

$$DV_{P,L,S} = \max_{k=2,\dots,m} DV_{k,P,L,S}(u_1,u_2),$$

where

$$\begin{aligned}DV_{k,P,L,S}(u_1,u_2) &= \frac{1}{\sqrt{P}}\sum_{t=R}^{T-\tau} \\ &\left(\left(\frac{1}{SL}\sum_{j=1}^{L}\sum_{i=1}^{S}1\left\{u_1 \le X_{1,t+\tau,i,j}^{\hat{\theta}_{1,t,N,L,h}}\left(X_t, V_{1,j}^{\hat{\theta}_{1,t,N,L,h}}\right) \le u_2\right\} - 1\{u_1 \le X_{t+\tau} \le u_2\}\right)^2\right. \\ &\left. - \left(\frac{1}{SL}\sum_{j=1}^{L}\sum_{i=1}^{S}1\left\{u_1 \le X_{k,t+\tau,i,j}^{\hat{\theta}_{k,t,N,L,h}}\left(X_t, V_{k,j}^{\hat{\theta}_{k,t,N,L,h}}\right) \le u_2\right\} - 1\{u_1 \le X_{t+\tau} \le u_2\}\right)^2\right).\end{aligned}$$

Critical values for these tests can be obtained using a recursive version of the block bootstrap. The corresponding bootstrap test statistic is

$$DV^*_{P,L,S} = \max_{k=2,\ldots,m} DV^*_{k,P,L,S}(u_1, u_2)$$

where

$$\begin{aligned}
DV^*_{k,P,L,S}(u_1,u_2) &= \frac{1}{\sqrt{P}}\sum_{t=R}^{T-\tau} \\
&\Bigg\{\Bigg(\left[\frac{1}{SL}\sum_{j=1}^{L}\sum_{i=1}^{S}1\left\{u_1 \le X_{1,t+\tau,i,j}^{\hat{\theta}^*_{1,t,N,L,h}}\left(X_t^*, V_{1,j}^{\hat{\theta}^*_{1,t,N,L,h}}\right) \le u_2\right\} - 1\left\{u_1 \le X^*_{t+\tau} \le u_2\right\}\right]^2 \\
&-\frac{1}{T}\sum_{l=1}^{T}\left[\frac{1}{SL}\sum_{j=1}^{L}\sum_{i=1}^{S}1\left\{u_1 < X_{1,t+\tau,i,j}^{\hat{\theta}_{1,t,N,L,h}}\left(X_l, V_{1,j}^{\hat{\theta}_{1,t,N,L,h}}\right) \le u_2\right\} - 1\{u_1 \le X_{l+\tau} \le u_2\}\right]^2\Bigg) \\
&-\Bigg(\left[\frac{1}{SL}\sum_{j=1}^{L}\sum_{i=1}^{S}1\left\{u_1 \le X_{k,t+\tau,i,j}^{\hat{\theta}^*_{k,t,N,L,h}}\left(X_t^*, V_{k,j}^{\hat{\theta}^*_{k,t,N,L,h}}\right) \le u_2\right\} - 1\left\{u_1 \le X^*_{t+\tau} \le u_2\right\}\right]^2 \\
&-\frac{1}{T}\sum_{l=1}^{T}\left[\frac{1}{SL}\sum_{j=1}^{L}\sum_{i=1}^{S}1\left\{u_1 \le X_{k,t+\tau,i,j}^{\hat{\theta}_{k,t,N,L,h}}\left(X_l, V_{k,j}^{\hat{\theta}_{k,t,N,L,h}}\right) \le u_2\right\} - 1\{u_1 \le X_{l+\tau} \le u_2\}\right]^2\Bigg)\Bigg\}
\end{aligned}$$

Of note is that we follow Cai and Swanson (2011) by adopting the recursive NPSQML estimator $\hat{\theta}_{1,t,N,L,h}$ and $\hat{\theta}_{k,t,N,L,h}$ for model 1 and k, respectively, as introduced in Sect. 56.3.3.4 with the choice $N = M = S$. $\hat{\theta}^*_{1,t,N,L,h}$ and $\hat{\theta}^*_{k,t,N,L,h}$ are bootstrap analogs of $\hat{\theta}_{1,t,N,L,h}$ and $\hat{\theta}_{k,t,N,L,h}$, respectively (see Sect. 56.3.4). In addition, we do not need to resample the volatility process, although volatility is simulated under both $\hat{\theta}_{k,t,N,L,h}$ and $\hat{\theta}^*_{k,t,N,L,h}$, $k = 1, \ldots, m$.

Repeat Steps 1–5 in the bootstrap procedure in Sect. 56.3.4 to obtain critical values. Compare $DV_{P,\ L,\ S}$ with the percentiles of the empirical distribution of $DV^*_{P,\ L,\ S}$, and reject H_0 if $DV_{P,\ L,\ S}$ is greater than the $(1 - \alpha)$th percentile. Otherwise, do not reject H_0. Again, in implementation, there is no need to re-estimate $\hat{\theta}^*_{k,t,N,L,h}$ for each bootstrap replication if $P/R \to 0$, $P, T, R \to \infty$, *as* parameter estimation error vanishes asymptotically in this case.

56.3.3 Model Simulation and Estimation

56.3.3.1 Simulating Data

Approximation schemes are used to obtain simulated distributions and simulated parameter estimators, which are needed in order to construct the test statistics outlined in previous sections. We therefore devote the first part of this section to a discussion of the Milstein approximation schemes that have been used in Corradi and Swanson (2005), Bhardwaj et al. (2008), and Corradi and Swanson (2011). Let L be the length of each simulation path and h be the discretization interval, $L = Qh$, and θ be a generic parameter in simulation expression. We consider three cases:

The pure diffusion process as in Eq. 56.9:

$$X^{\theta}_{qh} - X^{\theta}_{(q-1)h} = b\left(X^{\theta}_{(q-1)h}, \theta\right)h + \sigma\left(X^{\theta}_{(q-1)h}, \theta\right)\epsilon_{qh} - \frac{1}{2}\sigma\left(X^{\theta}_{(q-1)h}, \theta\right)'\sigma\left(X^{\theta}_{(q-1)h}, \theta\right)h$$
$$+\frac{1}{2}\sigma\left(X^{\theta}_{(q-1)h}, \theta\right)'\sigma\left(X^{\theta}_{(q-1)h}, \theta\right)\epsilon^{2}_{qh},$$

where

$$\left(W_{qh} - W_{(q-1)h}\right) = \epsilon_{qh} \overset{iid}{\sim} N(0, h),$$

$q = 1, \ldots, Q$, *with* $\epsilon_{qh} \sim^{iid} N(0, h)$, and where σ' is the derivative of $\sigma(\cdot)$ *with* respect to its first argument. Hereafter, X^{θ}_{qh} denotes the values of the diffusion at time qh, simulated under generic θ, and with a discrete interval equal to h, and so is a fine-grain analog of $X^{\theta}_{t,h}$.

The pure jump diffusion process without stochastic volatility as in Eq. 56.1:

$$\begin{aligned} X^{\theta}_{(q+1)h} - X^{\theta}_{qh} = {} & b\left(X^{\theta}_{qh}, \theta\right)h + \sigma\left(X^{\theta}_{qh}, \theta\right)\epsilon_{(q+1)h} - \frac{1}{2}\sigma\left(X^{\theta}_{qh}, \theta\right)'\sigma\left(X^{\theta}_{qh}, \theta\right)h \\ & + \frac{1}{2}\sigma\left(X^{\theta}_{qh}, \theta\right)'\sigma\left(X^{\theta}_{qh}, \theta\right)\epsilon^{2}_{(q+1)h} - \lambda\mu_y h \\ & + \sum_{j=1}^{\mathcal{J}} y_j 1\left\{qh \le \mathcal{U}_j \le (q+1)h\right\}. \end{aligned} \tag{56.11}$$

The only difference between this approximation and that used for the pure diffusion is the jump part. Note that the last term on the right-hand side (RHS) of Eq. 56.11 is nonzero whenever we have one (or more) jump realization(s) in the interval $[(q-1)h, qh]$. Moreover, as neither the intensity nor the jump size is state dependent, the jump component can be simulated without any discretization error, as follows. Begin by making a draw from a Poisson distribution with intensity parameter $\hat{\lambda}_{\tau}$, say $\mathcal{J}$. This gives a realization for the number of jumps over the simulation time span. Then draw $\mathcal{J}$ uniform random variables over $[0, L]$, and sort them in ascending order so that $\mathcal{U}_1 \le \mathcal{U}_2 \le \ldots \le \mathcal{U}_{\mathcal{J}}$. These provide realizations for the $\mathcal{J}$ independent draws from ϕ, say $y_1, \ldots, y_{\mathcal{J}}$.

SV models without jumps as in Eq. 56.4 (using a generalized Milstein scheme):

$$
\begin{aligned}
X^{\theta}_{(q+1)h} &= X^{\theta}_{qh} + \widetilde{b}_1\left(X^{\theta}_{qh}, \theta\right)h + \sigma_{11}\left(V^{\theta}_{qh}, \theta\right)\epsilon_{1,(q+1)h} + \sigma_{12}\left(V^{\theta}_{qh}, \theta\right)\epsilon_{2,(q+1)h} \\
&\quad + \frac{1}{2}\sigma_{22}\left(V^{\theta}_{qh}, \theta\right)\frac{\partial\sigma_{12,k}\left(V^{\theta}_{qh}, \theta\right)}{\partial V}\epsilon^2_{2,(q+1)h} + \sigma_{22}\left(V^{\theta}_{qh}, \theta\right)\frac{\partial\sigma_{11}\left(V^{\theta}_{qh}, \theta\right)}{\partial V} \\
&\quad \times \int_{qh}^{(q+1)h}\left(\int_{qh}^{s} dW_{1,\tau}\right)dW_{2.s},
\end{aligned}
\tag{56.12}
$$

$$
\begin{aligned}
V^{\theta}_{(q+1)h} &= V^{\theta}_{qh} + \widetilde{b}_2\left(V^{\theta}_{qh}, \theta\right)h + \sigma_{22}\left(V^{\theta}_{qh}, \theta\right)\epsilon_{2,(q+1)h} \\
&\quad + \frac{1}{2}\sigma_{22}\left(V^{\theta}_{qh}, \theta\right)\frac{\partial\sigma_{22}\left(V^{\theta}_{qh}, \theta\right)}{\partial V}\epsilon^2_{2,(q+1)h},
\end{aligned}
\tag{56.13}
$$

where $h^{-1/2}\epsilon_{i,qh} \sim N(0,1)$, $i = 1, 2, E\left(\epsilon_{1,qh}\epsilon_{2,q'h}\right) = 0$ for all $q \neq q'$, and

$$
\widetilde{b}(V, \theta) = \begin{pmatrix} \widetilde{b}_1(V, \theta) \\ \widetilde{b}_2(V, \theta) \end{pmatrix} = \begin{pmatrix} b_1(V, \theta) - \frac{1}{2}\sigma_{22}(V, \theta)\frac{\partial\sigma_{12}(V, \theta)}{\partial V} \\ b_2(V, \theta) - \frac{1}{2}\sigma_{22}(V, \theta)\frac{\partial\sigma_{22}(V, \theta)}{\partial V} \end{pmatrix}.
$$

The last terms on the RHS of Eq. 56.12 involve stochastic integrals and cannot be explicitly computed. However, they can be approximated up to an error of order $o(h)$ by (see, e.g., Eq. 3.7, pp. 347 in Kloeden and Platen (1999))

$$
\begin{aligned}
&\int_{qh}^{(q+1)h}\left(\int_{qh}^{s} dW_{1,\tau}\right)dW_{2,s} \approx h\left(\frac{1}{2}\xi_1\xi_2 + \sqrt{\rho_p}\left(\mu_{1,p}\xi_2 - \mu_{2,p}\xi_1\right)\right) \\
&+ \frac{h}{2\pi}\sum_{r=1}^{p}\frac{1}{r}\left(\varsigma_{1,r}\left(\sqrt{2}\xi_2 + \eta_{2,r}\right) - \varsigma_{2,r}\left(\sqrt{2}\xi_1 + \eta_{1,r}\right)\right),
\end{aligned}
$$

where for $j = 1, 2$, ξ_j, $\mu_{j,p}$, $\varsigma_{j,r}$, $\eta_{j,r}$ are i.i.d. $N(0, 1)$ random variables, $\rho_p = \frac{1}{12} - \frac{1}{2\pi^2}\sum_{r=1}^{p}\frac{1}{r^2}$, and p is such that as $h \to 0$, $p \to \infty$.

Stochastic Volatility with Jumps

Simulation of sample paths of diffusion processes with stochastic volatility and jumps follows straightforwardly from the previous two cases. Whenever both intensity and jump size are not state dependent, a jump component can be simulated and added to either $X(t)$ or the $V(t)$ in the same manner as above. Extension to general multidimensional and multifactor models both with and without jumps also follows directly.

56.3.3.2 Simulating Distributions

In this section, we sketch out methods used to construct τ-step ahead simulated conditional distributions using simulated data. In applications, simulation techniques are needed when the functional form conditional distribution is unknown. We first illustrate the technique for one-factor models and then discuss multifactor models.

One-factor Models

Consider the one-factor model as in Eq. 56.7. To estimate the simulated CDFs:

Step 1: Obtain $\hat{\theta}_{T,N,h}$ (using the entire sample) or $\hat{\theta}_{t,N,h}$ (recursive estimator) where $\hat{\theta}_{T,N,h}$ and $\hat{\theta}_{t,N,h}$ are estimators as discussed in Sects. 56.3.3.3 and 56.3.3.4.

Step 2: Under $\hat{\theta}_{T,N,h}$ or $\hat{\theta}_{t,N,h},$[20] simulate S paths of length τ, all having the same starting value, X_t. In particular, for each path $i = 1, \ldots, S$ of length τ, generate $X_{i,t+\tau}^{\hat{\theta}_{T,N,h}}(X_t)$ according to a Milstein schemes detailed in previous section, with $\theta = \hat{\theta}_{T,N,h}$ or $\hat{\theta}_{t,N,h}$. The errors used in simulation are $\varepsilon_{qh} \overset{iid}{\sim} N(0,h)$, and $Qh = \tau$. ε_{qh} is assumed to be independent across simulations, so that $E(\varepsilon_{i,qh}\varepsilon_{j,qh}) = 0$, for all $i \neq j$ and $E(\varepsilon_{i,qh}\varepsilon_{i,qh}) = h$, for any i, j. In addition, as the simulated diffusion is ergodic, the effect of the starting value approaches zero at an exponential rate, as $\tau \to \infty$.

Step 3: If $\hat{\theta}_{T,N,h}\left(\hat{\theta}_{t,N,h}\right)$ is used, an estimate for the distribution, at time $t + \tau$, conditional on X_t, with estimator $\hat{\theta}_{T,N,h}\left(\hat{\theta}_{t,N,h}\right)$, is defined as

$$\hat{F}_\tau\left(u \middle| X_t, \hat{\theta}_{T,N,h}\right) = \frac{1}{S}\sum_{i=1}^{S} 1\left\{X_{i,t+\tau}^{\hat{\theta}_{T,N,h}}(X_t) \leq u\right\}.$$

Bhardwaj et al. (2008) show that if the model is correctly specified, then $\frac{1}{S}\sum_{i=1}^{S} 1\left\{X_{i+t+\tau}^{\hat{\theta}_{T,N,h}}(X_t) \leq u\right\}$ provides a consistent of the conditional distribution $F_\tau(u|X_t, \theta^\dagger) = \Pr(X_{t+\tau}^{\theta\dagger} \leq u|X_t^{\theta\dagger} = X_t)$.

Specifically, assume that $T, N, S \to \infty$. Then, for the case of SGMM estimator, if $h \to 0$, $T/N \to 0$, and $h^2T \to 0$, $T^2/S \to \infty$, the following result holds for any Xt, $t \geq 1$, uniformly in u

$$\hat{F}_\tau\left(u \middle| X_t, \hat{\theta}_{T,N,h}\right) - F_\tau\left(u|X_t, \theta^\dagger\right) \xrightarrow{pr} 0,$$

[20] Note that $N = L$ for the SGMM estimator while $N = M = S$ for NSQML estimator.

In addition, if the model is correctly specified (i.e., if $\mu(,) = \mu_0(,)$ and $\sigma(,) = \sigma_0(,)$), then

$$\hat{F}_\tau\left(u\middle|X_t, \hat{\theta}_{T,N,h}\right) - F_{0,\tau}(u|X_t, \theta_0) \xrightarrow{pr} 0.$$

Step 4: Repeat Steps 1–3 for $t = 1, \ldots,$ T $-$ τ. This yields T $-$ τ conditional distributions that are τ-steps ahead which will be used in the construction of the specification tests.

The CDF simulation in the case selection test of multiple models with recursive estimator is similar. For model k, let $\hat{\theta}_{k,t,N,h}$ be the recursive estimator of "pseudo true" $\theta_k^\dagger$ computed using all observations up to varying time t. Then, $X_{k,i,t+\tau}^{\hat{\theta}_{k,t,N,h}}(X_t)$ is generated according to a Milstein schemes as in Sect. 56.3.3.1, with $\theta = \hat{\theta}_{k,t,N,h}$ and the initial value X_t, $Qh = \tau$. The corresponding empirical distribution of the simulated series $X_{k,i,t+\tau}^{\hat{\theta}_{k,t,N,h}}X(t)$ can then be constructed. Under some regularity conditions,

$$\frac{1}{S}\sum_{i=1}^{S} 1\left\{u_1 \le X_{k,i,t+\tau}^{\hat{\theta}_{k,t,N,h}}(X_t) \le u_2\right\} \xrightarrow{pr} F_{X_{k,t+\tau}^{\theta_k^\dagger}(X_t)}(u_2) - F_{X_{k,t+\tau}^{\theta_k^\dagger}(X_t)}(u_1), t = R, \ldots, T-\tau,$$

where $F_{X_{k,t+\tau}^{\theta_k^\dagger}(X_t)}(u)$ is the marginal distribution of $X_{t+\tau}^{\theta_k^\dagger}(X_t)$ implied by k model (i.e., by the model used to simulate the series), conditional on the (simulation) starting value X_t. Furthermore, the marginal distribution of $X_{t+\tau}^{\theta^\dagger}(X_t)$ is the distribution of $X_{t+\tau}$ conditional on the values observed at time t. Thus, $F_{X_{k+t+\tau}^{\theta^\dagger}(X_t)}(u) = F_k^T\left(u\middle|X_t, \theta_k^\dagger\right)$.

Of important note is that in the simulation of $X_{k,i,t+\tau}^{\hat{\theta}_{k,t,N,h}}(X_t)$, $i = 1, \ldots, S$, for each t, $t = R, \ldots, T - \tau$, we must use the same set of randomly drawn errors and similarly the same draws for numbers of jumps, jump times, and jump sizes. Thus, we only allow for the starting value to change. In particular, for each $i = 1, \ldots, S$, we generate $X_{k,i,R+\tau}^{\hat{\theta}_{k,R,N,h}}(X_R), \ldots, X_{k,i,t}^{\hat{\theta}_{k,T-\tau,N,h}}(X_{T-\tau})$. This yields an $S \times P$ matrix of simulated values, where $P = T - R - \tau + 1$ refers to the length of the out-of-sample period. $X_{k,i,R+j+\tau}^{\hat{\theta}_{k,R+j,N,h}}(X_{R+j})$ (at time $R + j + \tau$) can be seen as τ periods ahead value "predicted" by model k using all available information up to time $R + j_{R+j}$, $j = 1, \ldots, P$ (the initial value X_{R+j} and $\hat{\theta}_{k,R+j,N,h}$ estimated using $X_1, \ldots, X_{R+j}$). The key feature of this setup is that it enables us to compare "predicted" τ periods ahead values (i.e., $X_{k,i,R+j+\tau}^{\hat{\theta}_{k,R+j,N,h}}(X_{R+j})$) with actual values that are τ periods ahead (i.e., $X_{R+j+\tau}$), for $j = 1, \ldots, P$. In this manner, simulation-based tests under *ex-ante* predictive density comparison framework can be constructed.

Multifactor Model

Consider the multifactor model with a skeleton $(X_t, V_t^1, \ldots, V_t^d)'$ (e.g., stochastic mean, stochastic volatility models, stochastic volatility of volatility) where only the first element X_t is observed. For simulation of the CDF, the difficulty arises as we do not know the initial values of latent variables $(V_t^1, \ldots, V_t^d)'$ at each point in time. We generalize the simulation plan of Bhardwaj et al. (2008) and Cai and Swanson (2011) to the case of d dimensions. Specifically, to overcome the initial value difficulty, a natural strategy is to simulate a long path of length L for each latent variable $V_t^1, \ldots, V_t^d$ and use them to construct $X_{t+\tau}$ and the corresponding simulated CDF of $X_{t+\tau}$; and finally, we average out the volatility values. Note that there are L^d combinations of the initial values $V_t^1, \ldots, V_t^d$. For illustration, consider the case of stochastic volatility ($d = 1$) and the Chen three-factor model as in Eq. 56.5 ($d = 2$), using recursive estimators.

For the case of stochastic volatility ($d = 1$), i.e., $(X_t, V_t)'$, the steps are as follows:

Step 1: Estimate $\hat{\theta}_{t,N,L,h}$ using recursive SGMM or NSQML estimation methods.

Step 2: Using the scheme in Eq. 56.13 with $\theta = \hat{\theta}_{t,N,L,h}$, generate the path $V_{qh}^{\hat{\theta}_{t,N,L,h}}$ for $q = 1/h, \ldots, Qh$ with $Qh = L$ and hence obtain $V_j^{\hat{\theta}_{t,N,L,h}} j = 1, \ldots L$.

Step 3: Using schemes in Eqs. 56.12, 56.13, simulate $L \times S$ paths of length τ, setting the initial value for the observable state variable to be X_t. For the initial values of unobserved volatility, use $V_{j,qh}^{\hat{\theta}_{t,N,L,h}}$ $j = 1, \ldots, L$ as retrieved in Step 2. Also, keep the simulated random innovations (i.e., $\varepsilon_{1,qh}$, $\varepsilon_{1,qh}$, $\int_{qh}^{(q+1)h}(\int_{qh}^{s} dW_{1,\tau})dW_{2,s}$) to be constant across each j and t. Hence, for each replication i, using initial values X_t and $V_{j,qh}^{\hat{\theta}_{T,N,h}}$, we obtain $X_{j,i,t+\tau}^{\theta_{t,N,L,h}}(X_t)$ which is a τ-step ahead simulated value.

Step 4: Now the estimator of $F_\tau(u|X_t,\theta^\dagger)$ is defined as

$$\hat{F}_\tau\left(u\middle|X_t, \hat{\theta}_{t,N,L,h}\right) = \frac{1}{LS}\sum_{j=1}^{L}\sum_{i=1}^{S} 1\left\{X_{j,i,t+\tau}^{\hat{\theta}_{t,N,h}}(X_t) \le u\right\}.$$

Note that by averaging over the initial value of the volatility process, we have integrated out its effect. In other words, $\frac{1}{S}\sum_{i=1}^{S} 1\left\{X_{j,i,t+\tau}^{\theta_{t,N,h}}(X_t) \le u\right\}$ is an estimate *of* $F_\tau\left(u\middle|X_t, V_{j,h}^{\hat{\theta}_{t,N,h}}, \theta^\dagger\right)$.

Step 5: Repeat Steps 1–4 for $t = 1, \ldots, T - \tau$. This yields $T - \tau$ conditional distributions that are τ-steps ahead which will be used in the construction of the specification tests.

For three-factor model ($d = 2$), i.e., (X_t, V_t^1, V_t^2), consider model (56.5), where $W_t = (W_t^1, W_t^2, W_t^3)$ are mutually independent standard Brownian motions.

Step 1: Estimate $\hat{\theta}_{t,N,L,h}$ using SGMM or NSQML estimation methods.

Step 2: Given the estimated parameter $\hat{\theta}_{t,N,L,h}$, generate the paths $V_{qh}^{1,\hat{\theta}_{t,N,L,h}}$ and $V_{ph}^{2,\hat{\theta}_{t,N,L,h}}$ for q, $p = 1/h, \ldots, Qh$ with $Qh = L$, and hence, obtain $V_j^{1,\hat{\theta}_{T,N,L,h}}$, $V_k^{2,\hat{\theta}_{T,N,L,h}j}$, $k = 1, \ldots, L$.

Step 3: Given the observable X_t and the $L \times L$ simulated latent paths ($V_j^{1,\hat{\theta}_{t,N,L,h}}$ and $V_k^{2,\hat{\theta}_{t,N,L,h}}j$, $k = 1, \ldots, L$) as the start values, we simulate τ-step ahead $X_{t+\tau}^{\hat{\theta}_{t,N,L,h}}\left(X_t, V_j^{1,\hat{\theta}_{t,N,L,h}}, V_k^{2,\hat{\theta}_{t,N,L,h}}\right)$. Since the start values for the two latent variables are $L \times L$ length, so for each X_t we have N^2 path. Now to integrate out the initial effect of latent variables, form the estimate of conditional distribution as

$$\hat{F}_{\tau,s}\left(u\middle|X_t\hat{\theta}\right) = \frac{1}{L^2}\sum_{j=1}^{L}\sum_{k=1}^{L} 1\left\{X_{s,t+\tau}^{\hat{\theta}_{t,N,L,h}}\left(X_t, V_j^{1,\hat{\theta}_{t,N,L,h}}, V_k^{2,\hat{\theta}_{t,N,L,h}}\right) \le u\right\},$$

where s denotes the sth simulation.

Step 4: Simulate $X_{s,t+\tau}^{\hat{\theta}_{t,N,L,h}}$ S times, that is, repeat Step 3 S times, i.e., $s = 1, \ldots, S$. The estimate of $F_\tau(u|X_t,\theta^\dagger)$ is

$$\hat{F}_\tau\left(u\middle|X_t\hat{\theta}\right) = \frac{1}{S}\sum_{i=1}^{S}\hat{F}_{\tau,s}\left(u\middle|X_t, \hat{\theta}_{T,N,h}\right).$$

Step 5: Repeat Steps 1–4 for $t = 1, \ldots, T - \tau$. This yields $T - \tau$ conditional distributions that are t-steps ahead which will be used in the construction of the specification tests.

As a final remark, for the case of multiple competing models, we can proceed similarly. In addition, in the next two subsections, we present the exactly identified simulated (recursive) general method of moments and recursive nonparametric simulated quasi-maximum likelihood estimators that can be used in simulating distributions as well as constructing test statistics described in Sect. 56.3.2. The bootstrap analogs of those estimators will be discussed in Sect. 56.3.4.

56.3.3.3 Estimation: (Recursive) Simulated Generalized Method of Moments (SGMM) Estimators

Suppose that we observe a discrete sample (skeleton) of T observations, say $(X_1, X_2, \ldots, X_T)'$, from the underlying diffusion in Eq. 56.7. The (recursive) SGMM estimator $\hat{\theta}_{t,L,h}$ with $1 \le t \le T$ is specified as

$$\begin{aligned}\hat{\theta}_{t,L,h} &= \arg\min_{\theta\in\Theta}\left(\frac{1}{t}\sum_{j=1}^{t}g(X_j)-\frac{1}{L}\sum_{j=1}^{L}g\left(X_{j,h}^{\theta}\right)\right)'\\ &\quad\times W_t^{-1}\left(\frac{1}{t}\sum_{j=1}^{t}g(X_j)-\frac{1}{L}\sum_{j=1}^{L}g\left(X_{j,h}^{\theta}\right)\right)\\ &= \arg\min_{\theta\in\Theta}G_{t,L,h}(\theta)'W_tG_{t,L,h}(\theta),\end{aligned} \tag{56.14}$$

where g is a vector of p moment conditions, $\Theta \subset \Re^p$ (so that we have as many moment conditions as parameters), and $X_{j,h}^{\theta} = X_{[Qjh/L]}^{\theta}$, with $L = Qh$ is the simulated path under generic parameter θ and with discrete interval h. $X_{j,h}^{\theta}$ is simulated using the Milstein schemes.

Note that in the above expression, in the context of the specification test, $\hat{\theta}_{t,L,h}$ is estimated using the whole sample, i.e., $t = T$. In the out-of-sample context, the recursive SGMM estimator $\hat{\theta}_{t,L,h}$ is estimated recursively using the using sample from 1 up to t.

Typically, the p moment conditions are based on the difference between sample moments of historical and simulated data or between sample moments and model implied moments, whenever the latter are known in closed form. Finally, W_t is the heteroskedasticity and autocorrelation (HAC) robust covariance matrix estimator, defined as

$$W_t^{-1} = \frac{1}{t}\sum_{v=-l_t}^{l_t}w_v\sum_{j=v+1+lt}^{t-l_t}\left(g(X_j)-\frac{1}{t}\sum_{j=1}^{t}g(X_j)\right)\left(g(X_{j-v})-\frac{1}{t}\sum_{j=1}^{t}g(X_j)\right)', \tag{56.15}$$

where $w_v = 1 - v/(l_T + 1)$. Further, the pseudo true value, $\theta^{\dagger}$, is defined to be

$$\theta^{\dagger} = \arg\min_{\theta\in\Theta}G_{\infty}(\theta)'W_0G_{\infty}(\theta),$$

where

$$G_{\infty}(\theta)'W_0G_{\infty}(\theta) = p\lim_{L,T\to\infty,h\to 0}G_{T,L,h}(\theta)'W_TG_{T,L,h}(\theta),$$

and where $\theta^{\dagger} = \theta_0$ if the model is correctly specified.

In the above setup, the exactly identified case is considered rather than the overidentified SGMM. This choice guarantees that $G_{\infty}(\theta^{\dagger}) = 0$ even under misspecification, in the sense that the model differs from the underlying DGP. As pointed out by Hall and Inoue (2003), the root-N consistency does not hold for overidentified SGMM estimators of misspecified models. In addition,

$$\nabla_\theta G_\infty\left(\theta^\dagger\right)' W^\dagger G_\infty\left(\theta^\dagger\right) = 0.$$

However, in the case for which the number of parameters and the number of moment conditions are the same, $\nabla_\theta G_\infty(\theta^\dagger)' W^\dagger$ is invertible, and so the first-order conditions also imply that $G_\infty(\theta^\dagger) = 0$.

Also note that other available estimation methods using moments include the efficient method of moments (EMM) estimator as proposed by Gallant and Tauchen (1996, 1997), which calculates moment functions by simulating the expected value of the score implied by an auxiliary model. In their setup, parameters are then computed by minimizing a chi-square criterion function.

56.3.3.4 Estimation: Recursive Nonparametric Simulated Quasi-maximum Likelihood Estimators

In this section, we outline a recursive version of the NPSQML estimator of Fermanian and Salanié (2004), proposed by Corradi and Swanson (2011). The bootstrap counterpart of the recursive NPSQML estimator will be presented in the next section.

One-Factor Models

Hereafter, let $f(X_t|X_{t-1}, \theta^\dagger)$ be the conditional density associated with the above jump diffusion. If f is known in closed form, we can just estimate $\theta^\dagger$ recursively, using standard QML as[21]

$$\hat{\theta}_t = \arg\max_{\theta \in \Theta} \frac{1}{t} \sum_{j=2}^{t} \ln f\left(X_j \middle| X_{j-1}, \theta\right), t = R, \ldots R + P - 1. \tag{56.16}$$

Note that, similarly to the case of SGMM, the pseudo true value $\theta^\dagger$ is optimal in the sense:

$$\theta^\dagger = \arg\max_{\theta \in \Theta} E(\ln f(X_t|X_{t-1}, \theta)). \tag{56.17}$$

For the case f is not known in closed form, we can follow Kristensen and Shin (2008) and Cai and Swanson (2011) to construct the simulated analog $\hat{f}$ of f and then use it to estimate $\theta^\dagger$. $\hat{f}$ is estimated as function of the simulated sample paths $X^\theta_{t,i}(X_{t-1})$, for $t = 2, \ldots, T-1$, $i = 1, \ldots, M$. First, generate $T-1$ paths of length one for each simulation replication, using X_{t-1} with $t = 1, \ldots, T$ as starting values. Hence, at time t and simulation replication i, we obtain skeletons $X^\theta_{t,i}(X_{t-1})$, for $t = 2, \ldots, T-1$; $i = 1, \ldots M$ where M is the number of simulation paths (number of random draws or $X^\theta_{t,j}(X_{t-1})$ and $X^\theta_{t,l}(X_{t-1})$ are i.i.d.) for each simulation replication.

[21]Note that as model k is, in general, misspecified, $\sum_{t=1}^{T-1} f_k(X_t|X_{t-1},\theta_k)$ is a quasi-likelihood and $f_k(X_t|X_{t-1},\theta_k^\dagger)$ is not necessarily a martingale difference sequence.

M is fixed across all initial values. Then the recursive NPSQML estimator is defined as follows:

$$\hat{\theta}_{t,M,h} = \arg\max_{\theta\in\Theta} \frac{1}{t}\sum_{i=2}^{t} \left(\ln\hat{f}_{M,h}(X_i|X_{i-1},\theta)\tau_M \times \hat{f}_{M,h}(X_i|X_{i-1},\theta)\right), t \geq R,$$

where

$$\hat{f}_{M,h}(X_t|X_{t-1},\theta) = \frac{1}{M\xi_M}\sum_{i=1}^{M} K\left(\frac{X^{\theta}_{t,i,h}(X_{t-1}) - X_t}{\xi_M}\right).$$

Note that with abuse of notation, we define $\hat{\theta}_{t,L,h}$ for SGMM and $\hat{\theta}_{t,M,h}$ for NPSQML estimators where L and M have different interpretations (L is the length of each simulation path and M is number of random draws).

The function $\tau_M\left(\hat{f}_{M,h}(X_t|X_{t-1},\theta)\right)$ is a trimming function. It has some characteristics such as positive and increasing, $\tau_M\left(\hat{f}_{M,h}(X_t,X_{t-1},\theta)\right) = 0$, if $\hat{f}_{M,h}(X_t,X_{t-1},\theta) < \xi_M^{\delta}$ and $\tau_M\left(\hat{f}_{M,h}(X_t,X_{t-1},\theta)\right) = 1$ if $\hat{f}_{M,h}(X_t,X_{t-1},\theta) > 2\xi_M^{\delta}$, for some $\delta > 0$.[22] Note that when the log density is close to zero, the derivative tends to infinity and thus even very tiny simulation errors can have a large impact on the likelihood. The introduction of the trimming parameter into the optimization function ensures the impact of this case to be minimal asymptotically.

Multifactor Models

Since volatility is not observable, we cannot proceed as in the single-factor case when estimating the SV model using NPSQML estimator. Instead, let V_j^{θ} be generated according to Eq. 56.13, setting $qh = j$, and $j = 1, \ldots, L$. The idea is to simulate L different starting values for unobservable volatility, construct the simulated likelihood functions accordingly, and then average them out. For each simulation replication at time t, we simulate L different values of $X_t(X_{t-1}, V_j^{\theta})$ by generating L paths of length one, using fixed observable X_{t-1} and unobservable V_j^{θ}, $j = 1, \ldots, L$ as starting values. Repeat this procedure for any $t = 1, \ldots, T-1$ and for any set $j, j = 1, \ldots, L$ of random errors $\varepsilon_{1,t+(q+1)h,j}$ and $\varepsilon_{2,t+(q+1)h,j}$, $q = 1, \ldots, 1/h$. Note that it is important to use the same set of random errors $\varepsilon_{1,t+(q+1)h,j}$ and $\varepsilon_{2,t+(q+1)h,j}$ across different initial values for volatility. Denote the simulated value

[22] Fermanian and Salanie (2004) suggest using the following trimming function:

$$\tau_N(x) = \frac{4(x-a_N)^3}{a_N^3} - \frac{3(x-a_N)4}{a_N^4},$$

for $a_N \leq x \leq 2a_N$.

at time t and simulation replication i, under generic parameter θ, using X_{t-1}, V_j^θ as starting values as $X_{t,i,h}^\theta(X_{t-1}, V_j^\theta)$. Then

$$\hat{f}_{M,L,h}(X_t|X_{t-1},\theta) = \frac{1}{L}\sum_{j=1}^{L}\frac{1}{M\xi_M}\sum_{i=1}^{M}K\left(\frac{X_{t,i,h}^\theta\left(X_{t-1},V_j^\theta\right) - X_t}{\xi_M}\right),$$

and note that by averaging over the initial values for the unobservable volatility, its effect is integrated out. Finally, define[23]

$$\hat{\theta}_{t,M,L,h} = \arg\min_{\theta\in\Theta}\frac{1}{t}\sum_{s=2}^{t}\left(\ln \hat{f}_{M,L,h}(X_s|X_{s-1},\theta)\tau_M \times \hat{f}_{M,L,h}(X_s|X_{s-1},\theta)\right), t \geq R.$$

Note that in this case, X_t is no longer Markov (i.e., X_t and V_t are jointly Markovian, but X_t is not). Therefore, even in the case of true data generating process, the joint likelihood cannot be expressed as the product of the conditional and marginal distributions. Thus, $\hat{\theta}_{t,M,L,h}$ is necessarily a QML estimator. Furthermore, note that $\nabla_\theta f(X_t|X_{t-1},\theta^\dagger)$ is no longer a martingale difference sequence; therefore, we need to use HAC robust covariance matrix estimators, regardless of whether the model is the "correct" model or not.

56.3.4 Bootstrap Critical Value Procedures

The test statistics presented in Sects. 56.3.1 and 56.3.2 are implemented using critical values constructed via the bootstrap. As mentioned earlier, motivation for using the bootstrap is clear. The covariance kernel of the statistic limiting distributions contains both parameter estimation error and the data-related time-dependence components. Asymptotic critical value cannot thus be tabulated in a usual way. Several methods have been proposed to tackle this issue. One is the block bootstrap procedures which we discuss. Others have been mentioned above.

With regard to the validity of the bootstrap, note that, in the case of dependent observations without PEE, we can tabulate valid critical value using a simple empirical version of the Künsch (1989) block bootstrap. Now, the difficulty in our context lies in accounting for parameter estimation error. Goncalves and White (2002) establish the first-order validity of the block bootstrap for QMLE (or m-estimator) for dependent and heterogeneous data. This is an important result

[23]For discussion of asymptotic properties of $\hat{\theta}_{k,t,M,L,h}$, as well as of regularity conditions, see Corradi and Swanson (2011).

for the class of SGMM and NSQML estimators surveyed in this chapter and allows Corradi and Swanson in CS (2011) and elsewhere to develop asymptotically valid version of the bootstrap that can be applied under generic model misspecification, as assumed throughout this chapter.

For the SGMM estimator, as shown in Corradi and Swanson (2005), the first-order validity of the block bootstrap is valid in the exact identification case, and when $T/S \to 0$. In this case, SGMM is asymptotically equivalent to GMM, and consequently, there is no need to bootstrap the simulated series. In addition, in the exact identification case, GMM estimators can be treated the same way that QMLE estimators are treated. For the NSQML estimator, Corradi and Swanson (2011) point out that the NPSQML estimator is asymptotically equivalent to the QML estimator. Thus, we do not need to resample the simulated observations as the negligible contribution of simulation errors.

Also note that critical values for these tests can be obtained using a recursive version of the block bootstrap. When forming block bootstrap samples in the recursive case, observations at the beginning of the sample are used more frequently than observations at the end of the sample. This introduces a location bias to the usual block bootstrap, as under standard resampling with replacement, all blocks from the original sample have the same probability of being selected. Also, the bias term varies across samples and can be either positive or negative, depending on the specific sample. A first-order valid bootstrap procedure for nonsimulation-based m-estimators constructed using a recursive estimation scheme is outlined in Corradi and Swanson (2007a). Here we extend the results of Corradi and Swanson (2007a) by establishing asymptotic results for cases in which simulation-based estimators are bootstrapped in a recursive setting.

Now the details of bootstrap procedure for critical value tabulation can be outlined in 5 steps as follows:

Step 1: Let $T = bl$, where b denotes the number of blocks and l denotes the length of each block. We first draw a discrete uniform random variable, I_1, that can take values $0, 1, \ldots, T - l$ with probability $1/(T - l + 1)$. The first block is given by $X_{I_1+1}, \ldots, X_{I_1+l}$. We then draw another discrete uniform random variable, say I_2, and a second block of length l is formed, say $X_{I_2+1}, \ldots, X_{I_2+l}$. Continue in the same manner, until you draw the last discrete uniform, say I_b, and so the last block is $X_{I_b+1}, \ldots, X_{I_b+l}$. Let's call the X_t^* the resampled series, and note that $X_1^*, X_2^*, \ldots, X_T^*$ corresponds to $X_{I_1+1}, X_{I_1+2}, \ldots, X_{I_b+l}$. Thus, conditional on the sample, the only random element is the beginning of each block. In particular,

$$X_1^*, \ldots, X_l^*, X_{l+1}^*, \ldots, X_{2l}^*, X_{T-l+1}^*, \ldots, X_T^*,$$

conditional on the sample, can be treated as b i.i.d. blocks of discrete uniform random variables. Note that it can be shown that except a set of probability measure approaching zero,

$$E^*\left(\frac{1}{T}\sum_{t=1}^{T}X_t^*\right)=\frac{1}{T}\sum_{t=1}^{T}X_t+O_P^*(l/T) \tag{56.18}$$

$$Var*\left(\frac{1}{T^{1/2}}\sum_{t=1}^{T}X_t^*\right)=\frac{1}{T}\sum_{t=l}^{T-l}\sum_{i=-l}^{l}\left(X_t-\frac{1}{T}\sum_{t=1}^{T}X_t\right)\left(X_{t+i}-\frac{1}{T}\sum_{t=1}^{T}X_t\right)+O_{P*}\left(l^2/T\right), \tag{56.19}$$

where E^* and Var^* denote the expectation and the variance operators with respect to P^* (the probability law governing the resampled series or the probability law governing the i.i.d. uniform random variables, conditional on the sample) and where $O_{P^*}(l/T)(O_{P^*}(l^2/T))$ denotes a term converging in probability P^* to zero, as $l/T \to 0(l^2/T \to 0)$.

In the case of recursive estimators, we proceed the bootstrap similarly as follows. Begin by resampling b blocks of length l from the full sample, with $lb = T$. For any given τ, it is necessary to jointly resample $X_t, X_{t+1}, \ldots, X_{t+\tau}$. More precisely, let $Z^{t,\tau} = (X_t, X_{t+1}, \ldots, X_{t+\tau})$, $t = 1,\ldots, T - \tau$. Now, resample b overlapping blocks of length l from $Z^{t,\tau}$. This yields $Z^{t,*} = (X_t^*, X_{t+1}^*, \ldots, X_{t+\tau}^*)$, $t = 1, \ldots, T - \tau$.

Step 2: Re-estimate $\hat{\theta}^*_{t,N,h}\left(\hat{\theta}^*_{T,N,L,h}\right)$ using the bootstrap sample $Z^{t,*} = (X_t^*, X_{t+1}^*, \ldots, X_{t+\tau}^*)$, $t = 1, \ldots, T - \tau$ (or full sample $X_1^*, X_2^*, \ldots, X_T^*$). Recall that if we use the entire sample for the estimation, as the specification test in Corradi and Swanson (2005) and Bhardwaj et al. (2008), then $\hat{\theta}^*_{t,N,h}$ is denoted as $\hat{\theta}^*_{T,N,h}$. The bootstrap estimators for SGMM and NPSQML are presented below:

Bootstrap (Recursive) SGMM Estimators

If the full sample is used in the specification test as in Corradi and Swanson (2005) and Bhardwaj et al. (2008), the bootstrap estimator is constructed straightforward as

$$\begin{aligned}\hat{\theta}^*_{T,N,h} = \arg\min_{\theta\in\Theta}&\left(\frac{1}{T}\sum_{j=1}^{T}g\left(X_j^*\right)-\frac{1}{L}\sum_{i=1}^{L}g\left(X_{j,h}^{\theta}\right)\right)'\\ &\times W_T^{*-1}\left(\frac{1}{T}\sum_{j=1}^{T}g\left(X_j^*\right)-\frac{1}{L}\sum_{i=1}^{L}g\left(X_{j,h}^{\theta}\right)\right),\end{aligned}$$

where W_T^{-1} and $g(.)$ are defined in Eq. 56.15 and L is the length of each simulation path.

Note that it is convenient not to resample the simulated series as the simulation error vanishes asymptotically. In implementation, we do not mimic its contribution to the covariate kernel.

In the case of predictive density-type model selection where recursive estimators are needed, define the bootstrap analog as

$$
\begin{aligned}
\hat{\theta}^*_{t,L,h} = \arg\min_{\theta\in\Theta} & \left(\frac{1}{t}\sum_{j=1}^{t}\left(\left(g\left(X_j^*\right) - \frac{1}{T}\sum_{j'=1}^{T} g\left(X_{j'}\right)\right)\right.\right. \\
& \left.\left.- \left(\frac{1}{L}\sum_{i=1}^{L} g\left(X_{j,h}^{\theta}\right) - \frac{1}{L}\sum_{i=1}^{L} g\left(X_{j,h}^{\hat{\theta}_{t,L,h}}\right)\right)\right)\right)' \\
& \times \Omega_t^{*-1}\left(\frac{1}{t}\sum_{j=1}^{t}\left(\left(g\left(X_j^*\right) - \frac{1}{T}\sum_{j'=1}^{T} g\left(X_{j'}\right)\right)\right.\right. \\
& \left.\left.- \left(\frac{1}{L}\sum_{i=1}^{L} g\left(X_{j,h}^{\theta}\right) - \frac{1}{L}\sum_{i=1}^{L} g\left(X_{j,h}^{\hat{\theta}_{t,L,h}}\right)\right)\right)\right) \\
= \arg\min_{\theta\in\Theta} & \, G^*_{t,L,h}(\theta)'\Omega_t^{*-1} G^*_{t,L,h}(\theta),
\end{aligned}
$$

where

$$
\Omega_t^{*-1} = \frac{1}{t}\sum_{v=-l_t}^{l_t} w_{v,t} \sum_{j=v+1+l_t}^{t-l_t}\left[g\left(X_j^*\right) - \frac{1}{T}\sum_{j'=1}^{T} g\left(X_{j'}\right)\right]\left[g\left(X_{j-v}^*\right) - \frac{1}{T}\sum_{j'=1}^{T} g\left(X_{j'}\right)\right].
$$

Note that each bootstrap term is recentered around the (full) sample mean. The intuition behind the particular recentering in bootstrap recursive SGMM estimator is that it ensures that the mean of the bootstrap moment conditions, evaluated at $\hat{\theta}_{t,L,h}$, is zero, up to a negligible term. Specifically, we have

$$
\begin{aligned}
& E*\left(\frac{1}{t}\sum_{j=1}^{t}\left(g\left(X_j^*\right) - \frac{1}{T}\sum_{j'=1}^{T} g\left(X_{j'}\right)\right)\right. \\
& \left.- \left(\frac{1}{L}\sum_{i=1}^{L} g\left(X_{j,h}^{\hat{\theta}^*_{t,L,h}}\right) - \frac{1}{L}\sum_{i=1}^{L} g\left(X_{j,h}^{\hat{\theta}_{t,L,h}}\right)\right)\right) \\
& = E*\left(g\left(X_j^*\right)\right) - \frac{1}{T}\sum_{j'=1}^{T} g\left(X_{j'}\right) \\
& = O(l/T), \text{with}\, l = o\left(T^{1/2}\right),
\end{aligned}
$$

where the $O(l/T)$ term is due to the end block effect (see Corradi and Swanson (2007b) for further discussion).

Bootstrap Recursive NPSQML Estimators

Let $Z^{t,*} = (X_t^*, X_{t+1}^*, \ldots, X_{t+\tau}^*)$, $t = 1, \ldots, T - \tau$. For each simulation replication, generate $T - 1$ paths of length one, using $X_1^*, \ldots, X_{T-1}^*$ as starting values, and so obtaining $X_{t,j}^{\theta}(X_{t-1}^*)$ for $t = 2, \ldots, T - 1$, $i = 1,.., M$. Further, let:

$$\hat{f}^*_{M,h}\left(X_t^* \middle| X_{t-1}^*, \theta\right) = \frac{1}{M\xi_M} \sum_{i=1}^{M} K\left(\frac{X_{t,i,h}^{\theta}\left(X_{t-1}^*\right) - X_t^*}{\xi_M}\right),$$

Now, for $t = R, \ldots, R + P - 1$, define

$$\begin{aligned}
\hat{\theta}^*_{t,M,h} = \arg\max_{\theta \in \Theta} \frac{1}{t} \sum_{l=2}^{t} & \ln \hat{f}_{M,h}\left(X_l^* \middle| X_{l-1}^*, \theta\right) \tau_M \\
& \times \left(\hat{f}_{M,h}\left(X_l^* \middle| X_{l-1}^*, \theta\right)\right) \\
& -\theta' \left(\frac{1}{T} \sum_{l'=2}^{T} \frac{\nabla_\theta \hat{f}_{M,h}(X_{l'} | X_{l'-1}, \theta)}{\hat{f}_{M,h}\left(X_{l'}^* | X_{t-l'}^* \theta\right)} \bigg|_{\theta = \hat{\theta}_{t,M,h}} \right. \\
& \times \tau_M \left(\hat{f}_{M,h}\left(X_{l'} \middle| X_{l'-1}, \hat{\theta}_{t,M,h}\right)\right) \\
& + \tau'_M \left(\hat{f}_{M,h}\left(X_{l'} \middle| X_{l'-1}, \hat{\theta}_{t,M,h}\right)\right) \\
& \times \nabla_\theta \hat{f}_{M,h}(X_{l'} | X_{l'-1}, \theta) \Big|_{\hat{\theta}_{t,M,h}} \\
& \left. \times \ln \hat{f}_{M,h}(X_{l'} | X_{l'-1} \hat{\theta}_{t,M,h}) \right),
\end{aligned}$$

where $\tau'_M(\cdot)$ denotes the derivative of $\tau_M(\cdot)$ with respect to its argument. Note that each term in the simulated likelihood is recentered around the (full) sample mean of the score, evaluated at $\hat{\theta}_{t,M,h}$. This ensures that the bootstrap score has mean zero, conditional on the sample. The recentering term requires computation of $\nabla_\theta \hat{f}_{M,h}\left(X_{l'} \middle| X_{l'-1}, \hat{\theta}_{t,M,h}\right)$, which is not known in closed form. Nevertheless, it can be computed numerically, by simply taking the numerical derivative of the simulated likelihood.

Bootstrap Estimators for Multifactor Model

The SGMM and the bootstrap SGMM estimators in the case of multifactor model are similar as in one-factor model. The difference is that the simulation schemes (56.12) and (56.13) are used instead of Eq. 56.11.

For recursive NPSQML estimators, to construct the bootstrap counterpart $\hat{\theta}^*_{t,M,L,h}$ of $\hat{\theta}_{t,M,L,h}$, since $M/T \to \infty$ and $L/T \to \infty$, the contribution of simulation error is asymptotically negligible. Hence, there is no need to resample the simulated observations or the simulated initial values for volatility. Define

$$\hat{f}_{M,L,h}\left(X_t^*\middle|X_{t-1}^*,\theta\right) = \frac{1}{L}\sum_{j=1}^{L}\frac{1}{M\xi_M}\sum_{i=1}^{M}K\left(\frac{X_{t,i,h}^{\theta}\left(X_{t-1}^*,V_j^{\theta}\right)-X_t^*}{\xi_M}\right).$$

Now, for $t = R, \ldots, R + P - 1$, define

$$\begin{aligned}\hat{\theta}^*_{t,M,L,h} = \arg\max_{\theta\in\Theta}\frac{1}{t}\sum_{l=2}^{t} & \\ & \left(\log\hat{f}_{t,M,L,h}\left(X_l^*\middle|X_{l-1}^*,\theta\right)\tau_M\left(\hat{f}_{t,M,L,h}\left(X_l^*\middle|X_{l-1}^*,\theta\right)\right)\right. \\ & -\theta'\left(\frac{1}{T}\sum_{l'=2}^{T}\frac{\nabla_\theta\hat{f}_{t,M,L,h}(X_{l'}|X_{l'-1},\theta)}{\hat{f}_{t,M,L,h}\left(X_{l'}^*\middle|X_{t-l'}^*,\theta\right)}\Big|_{\theta_{t,M,L,h}}\right. \\ & \times\tau_M = \left(\hat{f}_{t,M,L,h}\left(X_{l'}\middle|X_{l'-1},\hat{\theta}_{t,M,L,h}\right)\right) \\ & +\tau'_M\left(\hat{f}_{t,M,L,h}\left(X_{l'}\middle|X_{l'-1},\hat{\theta}_{t,M,L,h}\right)\right) \\ & \times\nabla_\theta\hat{f}_{t,M,L,h}(X_{l'}|X_{l'-1},\theta)\Big|_{\hat{\theta}_{t,M,L,h}} \\ & \left.\left.\times\ln\hat{f}_{t,M,L,h}(X_{l'}|X_{l'-1},\hat{\theta}_{t,M,L,h})\right)\right),\end{aligned}$$

where $\tau_M'(\cdot)$ denotes the derivative with respect to its argument.

Of note is that each bootstrap term is recentered around the (full) sample mean. This is necessary because the bootstrap statistic is constructed using the last P resampled observations, which in turn have been resampled from the full sample. In particular, this is necessary regardless of the ratio, P/R. In addition, in the case $P/R \to 0$, so that there is no need to mimic parameter estimation error, the bootstrap statistics can be constructed using $\hat{\theta}_{t,M,L,h}$ instead of $\hat{\theta}^*_{t,M,L,h}$.

Step 3: Using the same set of random variables used in the construction of the actual statistics, construct $X_{i,t+\tau,*}^{\hat{\theta}^*_{t,N,h}}$ or $X_{k,i,t+\tau,*}^{\hat{\theta}^*_{t,N,h}}$, $i = 1, \ldots, S$ and $t = 1,\ldots, T - \tau$. Note that we do not need resample the simulated series (as $L/T \to \infty$, simulation error is asymptotically negligible). Instead, simulate the series using bootstrap estimators and using bootstrapped values as starting values.

Step 4: Corresponding bootstrap statistics $V_{T,N,h}^{2*}$ (or $Z_{T,N,h}^*$, $D_{k,P,S}^*$, $SV_{T,N,h}^{2\,*}$, $SZ_{T,N,h}^*$, $SD_{k,P,S}^*$ depending on the types of tests) which are built on $\hat{\theta}^*_{t,N,h}\left(\hat{\theta}^*_{t,N,L,h}\right)$ then are followed correspondingly. For the numerical implementation, again, of important note is that in the case where we pick the choice $P/R \to 0$, P, T,

$R \to \infty$, there is no need to re-estimate $\hat{\theta}^*_{t,N,h}\left(\hat{\theta}^*_{t,N,L,h}\right)$. $\hat{\theta}^*_{t,N,h}\left(\hat{\theta}^*_{t,N,L,h}\right)$ can be used in all the bootstrap replications.

Step 5: Repeat the bootstrap Steps 1–4 B times, and generate the empirical distribution of the B bootstrap statistics.

56.4 Summary of Empirical Applications of the Tests

In this section, we briefly review some empirical applications of the methods discussed above. We start with unconditional distribution test, as in Corradi and Swanson (2005) and then give a specific empirical example using the conditional distribution test from Bhardwaj et al. (2008). Finally, we briefly discuss on conditional distribution specification test applied to multiple competing models. The list of the diffusion models considered is provided in Table 56.1.

Note that specification testing of the first model – a simplified version of the CIR model (we refer to this model as Wong) – is carried out using the unconditional distribution test. With the cumulative distribution function known in closed form as in Eq. 56.8, the test statistic can be straightforwardly calculated. It is also convenient to use GMM estimation in this case as the first two moments are known in closed form, i.e., $\alpha - \lambda$ and $\alpha/2(\alpha - \beta)$, respectively. Corradi and Swanson (2005) examine Hypothesis 1 using simulated data. Their Monte Carlo experiments suggest that the test is useful, even for samples as small as 400 observations.

Hypothesis 2 is tested in Bhardwaj et al. (2008) and Cai and Swanson (2011). For illustration, we focus on the results in Bhardwaj et al. (2008) where CIR, SV, and SVJ models are empirically tested using the 1-month Eurodollar deposit rate (as a proxy for short rate) for the sample period January 6, 1971–September 30, 2005, which yields 1,813 weekly observations. Note that one might apply these tests to other datasets including the monthly federal funds rate, the weekly 3-month T-bill rate, the weekly US dollar swap rate, the monthly yield on zero-coupon bonds with different maturities, and the 6-month LIBOR. Some of these variables have been examined elsewhere, for example, in Ait-Sahalia (1999), Andersen et al. (2004), Dai and Singleton (2000), Diebold and Li (2006), and Piazzesi (2001).

The statistic needed to apply the test discussed in Sect. 56.3.1.2 is

$$Z_T = \sup_{v \in V} |Z_T(v)|,$$

Table 56.1 Specification test hypotheses of continuous time spot rate process[a]

Model	Specification	Reference and data	Hypothesis
Wong	$dr(t) = (\alpha - \lambda - r(t))dt + \sqrt{\alpha r(t)}dW_r(t)$	Corradi and Swanson (2005)	H1
		Simulated data	
		Bhardwaj et al. (2008)	
		Eurodollar rate	H2
		(1971–2005)	
CIR	$dr(t) = k_r(\bar{r} - r(t))dt + \sqrt{V(t)}dW_r(t),$	Cai and Swanson (2011)	
		Eurodollar Rate	H2, H3
		(1971–2008)	
		Cai and Swanson (2011)	
CEV	$dr(t) = k_r(\bar{r} - r(t))dt + \sigma_r r(t)^{\rho}dW_r(t)$	Eurodollar rate	H2, H3
		(1971–2008)	
SV	$dr(t) = k_r(\bar{r} - r(t))dt + \sqrt{V(t)}dW_r(t),$	Bhardwaj et al. (2008)	H2
	$dV(t) = k_r(\bar{v} - V(t))dt + \sigma_v\sqrt{V(t)}dW_v(t),$	Cai and Swanson (2011)	H2, H3
SVJ	$dr(t) = k_r(\bar{r} - r(t))dt + \sqrt{V(t)}dW_r(t) + J_u dq_u - J_d dq_d,$	Bhardwaj et al. (2008)	H2
	$dV(t) = k_v(\bar{v} - V(t))dt + \sigma_v\sqrt{V(t)}dW_v(t),$	Cai and Swanson (2011)	H2, H3
CHEN	$dr(t) = \kappa_r(\theta(t) - r(t))dt + \sqrt{V(t)}dW_r,$	Cai and Swanson (2011)	
	$dV(t) = \kappa_v(\bar{v} - V(t))dt + \sigma_v\sqrt{V(t)}dW_v(t),$	Eurodollar rate	H2, H3
	$d\theta(t) = \kappa_\theta\left(\bar{\theta} - \theta(t)\right)dt + \sigma_\theta\sqrt{\theta(t)}dW_\theta(t),$	(1971–2008)	
CHENJ	$dr(t) = \kappa_r(\theta(t) - r(t))dt + \sqrt{V(t)}dW_r(t) + J_u dq_u - J_d dq_d,$	Cai and Swanson (2011)	
	$dV(t) = \kappa_v(\bar{v} - V(t))dt + \sigma_v\sqrt{V(t)}dW_v(t),$	Eurodollar rate	H2, H3
	$d\theta(t) = \kappa_\theta\left(\bar{\theta} - \theta(t)\right)dt + \sigma_\theta\sqrt{\theta(t)}dW_\theta(t),$	(1971–2008)	

[a]Note that the third column, "Reference and data," provides the referenced papers and data used in empirical applications. In the fourth column, H1, H2, and H3 denote Hypothesis 1, Hypothesis 2, and Hypothesis 3, respectively. The hypotheses are presented corresponding to the references in the third column. For example, for CIR model, H2 corresponds to Bhardwaj et al. (2008) and H2 and H3 correspond to Cai and Swanson (2011)

where

$$
\begin{aligned}
Z_T(v) = {} & \frac{1}{\sqrt{T-\tau}}\sum_{t=1}^{T-\tau} 1\{X_t \le v\} \\
& \times \left(\frac{1}{S}\sum_{s=1}^{S} 1\left\{\underline{u} \le X_{s,t+\tau}^{\hat{\theta}_{T,N,h}} \le \overline{u}\right\} - 1\left\{\underline{u} \le X_{t+\tau} \le \overline{u}\right\}\right),
\end{aligned}
$$

and

$$Z_T^* = \sup_{v \in V} \left| Z_T^*(v) \right|,$$

Where

$$\begin{aligned} Z_T^*(v) = & \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} 1\left\{ X_t^* \leq v \right\} \\ & \times \left(\frac{1}{S} \sum_{s=1}^{S} 1\left\{ \underline{u} \leq X_{s,t+\tau,*}^{\hat{\theta}_{T,N,h}^*} \leq \overline{u} \right\} - 1\left\{ \underline{u} \leq X_{t+\tau}^* \leq \overline{u} \right\} \right) \\ & - \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} 1\{ X_t \leq v \}. \\ & \times \left(\frac{1}{S} \sum_{s=1}^{S} 1\left\{ \underline{u} \leq X_{s,t+\tau}^{\hat{\theta}_{T,N,h}} \leq \overline{u} \right\} - 1\left\{ \underline{u} \leq X_{t+\tau} \leq \overline{u} \right\} \right). \end{aligned}$$

For the case of stochastic volatility models, similarly we have

$$SZ_T = \sup_{v \in V} |SZ_T(v)|,$$

where

$$\begin{aligned} SZ_T(v) = & \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} 1\{ X_t \leq v \} \\ & \times \left(\frac{1}{LS} \sum_{j=1}^{L} \sum_{s=1}^{S} 1\left\{ \underline{u} \leq X_{j,s,t+\tau}^{\hat{\theta}_{T,N,h}} \leq \overline{u} \right\} - 1\left\{ \underline{u} \leq X_{t+\tau} \leq \overline{u} \right\} \right), \end{aligned}$$

and its bootstrap analog

$$SZ_T^* = \sup_{v \in V} \left| SZ_T^*(v) \right|,$$

where

$$\begin{aligned} SZ_T^*(v) = & \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} 1\left\{ X_t^* \leq v \right\} \times \left(\frac{1}{LS} \sum_{j=1}^{L} \sum_{s=1}^{S} 1\left\{ \underline{u} \leq X_{j,s,t+\tau,*}^{\hat{\theta}_{i,T,N,h}^*} \leq \overline{u} \right\} \right. \\ & \left. - 1\left\{ \underline{u} \leq X_{t+\tau}^* \leq \overline{u} \right\} \right) - \frac{1}{\sqrt{T-\tau}} \sum_{t=1}^{T-\tau} 1\{ X_t \leq v \} \\ & \times \left(\frac{1}{LS} \sum_{j=1}^{L} \sum_{s=1}^{S} 1\left\{ \underline{u} \leq X_{j,s,t+\tau}^{\hat{\theta}_{i,T,N,h}} \leq \overline{u} \right\} - 1\left\{ \underline{u} \leq X_{t+\tau} \leq \overline{u} \right\} \right). \end{aligned}$$

Bhardwaj et al. (2008) carry out these tests using τ-*step* ahead confidence intervals. They set $\tau = \{1, 2, 4, 12\}$ which corresponds to 1-week, 2-week, 1-month, and one-quarter ahead intervals and set $\left(\underline{u}\overline{u}\right) = \left(\overline{X} \pm 0.5\sigma_X, \overline{X} \pm \sigma_X\right)$, covering 46.3 % and 72.4 % coverage, respectively. $\overline{X}$ and σ_X are the mean and variance of an initial sample of data. In addition, $S = \{10T, 20T\}$ and $l = \{5, 10, 20, 50\}$.

For illustrative purposes, we report one case from Bhardwaj et al. (2008). The test is implemented by setting $S = 10T$ and $l = 25$ for the calculation of both Z_T and SZ_T. In Table 56.2, single, double, and triple starred entries represent rejection using 20 %, 10 %, and 5 % size tests, respectively. Not surprisingly, the findings are consistent with some other papers in the specification test literature such as Aït-Sahalia (1996) and Bandi (2002). Namely, the CIR model is rejected using 5 % size tests in almost all cases. When considering SV and SVJ models, smaller confidence intervals appear to lead to more model rejections. Moreover, results are somewhat mixed when evaluating the SVJ model, with a slightly higher frequency of rejection than in the case of SV models.

Finally, turning to **Hypothesis 3**, Cai and Swanson (2011) use an extended version of the above dataset, i.e., the 1-month Eurodollar deposit rate from January 1971 to April 2008 (1,996 weekly observations). Specifically, they examine whether the Chen model is the "best" model amongst multiple alternative models including those outlined in Table 56.1. The answer is "yes." In this example, the test was implemented using $D_{k,p,N}(u1,\ u2)$, as described in Sects. 56.3.1 and 56.3.2, where $P = T/2$ and predictions are constructed using recursively estimated models and the simulation sample length used to address latent variable initial values is set at $L = 10T$. The choice of other inputs to the test such as τ and interval $\left(\underline{u}, \overline{u}\right)$ is the same as in Bhardwaj et al. (2008). The number of replications S, the block length l, and number of bootstrap replications are $S = 10T$, $l = 20$, and $B = 100$.

Cai and Swanson (2011) also compare the Chen model with the so-called smooth transition autoregression (STAR) model defined as follows:

$$r_t = (\theta_1 + \beta_1 r_{t-1})G(\gamma, z_t, c) + (\theta_1 + \beta_2 r_{t-1})(1 - G(\gamma, z_t, c)) + u_t,$$

where u_t is a disturbance term; θ_1, β_1, γ, β_2, and c are constants; $G(\cdot)$ is the logistic CDF $\left(\text{i.e.}, G, (\gamma, z_t, c), =, \frac{1}{1+\exp(\gamma(z_t - c))}\right)$ and the number of lags; and p is selected via the use of Schwarz information criterion. Test statistics and predictive density-type "mean square forecast error" (MSFEs) values are again calculated as in Sects. 56.3.1 and 56.3.2.[24] Their results indicate that at a 90 % level of confidence, one cannot reject the null hypothesis that the Chen model generates predictive densities at least as accurate as the STAR model, regardless of forecast horizon and confidence interval width. Moreover, in almost all cases, the Chen model has lower MSFE, and the magnitude of the MSFE differential between the Chen model and STAR model rises as the forecast horizon increases. This confirms their in-sample findings that the Chen model also wins when carrying out in-sample tests.

[24]See Table 6 in Cai and Swanson (2011) for complete details.

Table 56.2 Empirical illustration of specification testing – *CIR*, *SV*, *SVJ* models[a]

	$(\underline{u}, \overline{u})$	CIR			SV			SVJ		
		Z_T	5 % CV	10 % CV	SZ_T	5 % CV	10 % CV	SZ_T	5 % CV	10 % CV
					$l = 25$					
1	$\overline{X} \pm 0.5\sigma_X$	0.5274^{***}	0.2906	0.3545	0.9841^{***}	0.8729	0.9031	1.1319	1.8468	2.1957
	$\overline{X} \pm \sigma_X$	0.4289^{***}	0.2658	0.3178	0.6870	0.6954	0.7254	1.2272^{*}	1.1203	1.3031
2	$\overline{X} \pm 0.5\sigma_X$	0.6824^{***}	0.4291	0.4911	0.4113	1.3751	1.4900	0.9615^{*}	0.8146	1.1334
	$\overline{X} \pm \sigma_X$	0.4897^{*}	0.4264	0.5182	0.3682	1.1933	1.2243	1.2571	1.3316	1.4096
4	$\overline{X} \pm 0.5\sigma_X$	0.8662^{**}	0.7111	0.8491	1.2840	2.3297	2.6109	1.5012^{*}	1.1188	1.6856
	$\overline{X} \pm \sigma_X$	0.8539^{*}	0.7512	0.9389	1.0472	2.2549	2.2745	0.9901^{*}	0.9793	1.0507
12	$\overline{X} \pm 0.5\sigma_X$	1.1631^{*}	1.0087	1.3009	1.7687	4.9298	5.2832	2.4237^{*}	2.0818	3.0640
	$\overline{X} \pm \sigma_X$	1.0429	1.4767	2.0222	1.7017	5.2601	5.6522	1.4522	1.7400	2.1684

[a]Tabulated entries are test statistics and 5 %, 10 %, and 20 % level critical values. Test intervals are given in the second column of the table, for $\tau = 1, 2, 4, 12$. All tests are carried out using historical 1-month Eurodollar deposit rate data for the period January 1971–September 2005, measured at a weekly frequency. *, **, and *** denote rejection at the 20 %, 10 %, and 5 % levels, respectively. Additionally, $\overline{X}$ and σ_X are the mean and standard deviation of the historical data. See above for complete details

56.5 Conclusion

This chapter reviews a class of specification and model selection-type tests developed by Corradi and Swanson (2005), Bhardwaj et al. (2008), and Corradi and Swanson (2011) for continuous time models. We begin with outlining the setup used to specify the types of diffusion models considered in this chapter. Thereafter, diffusion models in finance are discussed, and testing procedures are outlined. Related testing procedures are also discussed, both in contexts where models are assumed to be either correctly specified under the null hypothesis or generically misspecified under both the null and alternative test hypotheses. In addition to discussing tests of correct specification and test for selecting amongst alternative competing models, using both in-sample methods and via comparison of predictive accuracy, methodology is outlined allowing for parameter estimation, model and data simulation, and bootstrap critical value construction.

Several extensions that are left to future research are as follows. First, it remains to construct specification tests that do not integrate out the effects of latent factors. Additionally, it remains to examine the finite sample properties of the estimators and bootstrap methods discussed in this chapter.

Acknowledgments The authors thank the editor, Cheng-Few Lee, for many useful suggestions given during the writing of this chapter. Duong and Swanson would like to thank the Research Council at Rutgers University for research support.

References

Ait-Sahalia, Y. (1999). Transition densities for interest rate and others nonlinear diffusions. *Journal of Finance*, *LIV*, 1361–1395.

Aït-Sahalia, Y. (1996). Testing continuous time models of the spot interest rate. *Review of Financial Studies, 9*, 385–426.

Aït-Sahalia, Y. (2002). Maximum likelihood estimation of discretely sampled diffusions: A closed form approximation approach. *Econometrica, 70*, 223–262.

Aït-Sahalia, Y. (2007). Estimating continuous-time models using discretely sampled data. In R. Blundell, P. Torsten, & W. K. Newey (Eds.), *Econometric Society World Congress invited lecture*, *in Advances in economics and econometrics*, *theory and applications*, *Ninth World Congress*, Econometric Society Monographs. Cambridge University Press.

Aït-Sahalia, Y., Fan, J., & Peng, H. (2009). Nonparametric transition-based tests for diffusions. *Journal of the American Statistical Association, 104*, 1102–1116.

Altissimo, F., & Mele, A. (2009). Simulated nonparametric estimation of dynamic models with application in finance. *Review of Economic Studies, 76*, 413–450.

Andersen, T. G., & Lund, J. (1997). Estimating continuous time stochastic volatility models of short term interest rates. *Journal of Econometrics, 72*, 343–377.

Andersen, T. G., Benzoni, L., & Lund, J. (2004). Stochastic volatility, mean drift, and jumps in the short-term interest rate (Working paper). Northwestern University.

Andrews, D. W. K. (1993). An introduction to econometric applications of empirical process theory for dependent random variables. *Econometric Reviews, 12*, 183–216.

Andrews, D. W. K. (1997). A conditional Kolmogorov test. *Econometrica, 65*, 1097–1128.

Bai, J. (2003). Testing parametric conditional distributions of dynamic models. *Review of Economics and Statistics, 85*, 531–549.

Bandi, F. (2002). Short-term interest rate dynamics: A spatial approach. *Journal of Financial Economics, 65*, 73–110.
Beckers, S. (1980). The constant elasticity of variance model and US implications for option pricing. *Journal of Finance, 35*, 661–673.
Bhardwaj, G., Corradi, V., & Swanson, N. R. (2008). A simulation based specification test for diffusion processes. *Journal of Business and Economic Statistics, 26*, 176–193.
Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy, 81*, 637–654.
Bontemps, C., & Meddahi, N. (2005). Testing normality: A GMM approach. *Journal of Econometrics, 124*, 149–186.
Brennan, M. J., & Schwartz, E. S. (1979). A continuous-time approach to the pricing bonds. *Journal of Banking and Finance, 3*, 133–155.
Cai, L., & Swanson, N. R. (2011). An empirical assessment of spot rate model stability. *Journal of Empirical Finance, 18*, 743–764.
Chan, K. C., Karolyi, G. A., Longstaff, F. A., & Sanders, A. B. (1992). An empirical comparison of alternative models of the short-term interest rate. *Journal of Finance, 47*, 1209–1227.
Chen, L. (1996). *Stochastic mean and stochastic volatility – A three-factor model of term structure of interest rates and its application to the pricing of interest rate derivatives*. Oxford, UK: Blackwell.
Clements, M. P., & Smith, J. (2000). Evaluating the forecast densities of linear and nonlinear models: Applications to output growth and unemployment. *Journal of Forecasting, 19*, 255–276.
Clements, M. P., & Smith, J. (2002). Evaluating multivariate forecast densities: A comparison of two approaches. *International Journal of Forecasting, 18*, 397–407.
Corradi, V., & Swanson, N. R. (2005). A bootstrap specification test for diffusion processes. *Journal of Econometrics, 124*, 117–148.
Corradi, V., & Swanson, N. R. (2006). Predictive density and conditional confidence interval accuracy tests. *Journal of Econometrics, 135*, 187–228.
Corradi, V., & Swanson, N. R. (2007a). Nonparametric bootstrap procedures for predictive inference based on recursive estimation schemes. *International Economic Review February, 2007*(48), 67–109.
Corradi, V., & Swanson, N. R. (2007b). Nonparametric bootstrap procedures for predictive inference based on recursive estimation schemes. *International Economic Review, 48*, 67–109.
Corradi, V., & Swanson, N. R. (2011). Predictive density construction and accuracy testing with multiple possibly misspecified diffusion models. *Journal of Econometrics, 161*, 304–324.
Corradi, V. & Swanson, N. R. (2002). A consistent test for nonlinear out of sample predictive accuracy. *Journal of Econometrics*, *110*, 353–381.
Courtadon, G. (1982). The pricing of options on default-free bonds. *Journal of Financial and Quantitative Analysis, 17*, 75–100.
Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985). A theory of the term structure of interest rates. *Economerica, 53*, 385–407.
Dai, Q., & Singleton, K. J. (2000). Specification analysis of affine term structure models. *Journal of Finance, 55*, 1943–1978.
Diebold, F. X., & Li, C. (2006). Forecasting the term structure of government bond yields. *Journal of Econometrics, 130*, 337–364.
Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics, 13*, 253–263.
Diebold, F. X., Gunther, T., & Tay, A. S. (1998a). Evaluating density forecasts with applications to finance and management. *International Economic Review, 39*, 863–883.
Diebold, F. X., Tay, A. S., & Wallis, K. D. (1998b). Evaluating density forecasts of inflation: The survey of professional forecasters. In R. F. Engle & H. White (Eds.), *Festschrift in honor of C.W.J. Granger*. Oxford: Oxford University Press.

Diebold, F. X., Hahn, J., & Tay, A. S. (1999). Multivariate density forecast evaluation and calibration in financial risk management: High frequency returns on foreign exchange. *Review of Economics and Statistics, 81*, 661–673.

Duan, J. C. (2003). A specification test for time series models by a normality transformation (Working paper). University of Toronto.

Duffie, D., & Kan, R. (1996). A yield-factor model of interest rates. *Mathematical Finance, 6*, 379–406.

Duffie, D., & Singleton, K. (1993). Simulated moment estimation of Markov models of asset prices. *Econometrica, 61*, 929–952.

Duffie, D., Pan, J., & Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica, 68*, 1343–1376.

Duong, D., & Swanson, N. R. (2010). Empirical evidence on jumps and large fluctuations in individual stocks (Working paper). Rutgers University.

Duong, D., & Swanson, N. R. (2011). Volatility in discrete and continuous time models: A survey with new evidence on large and small jumps. *Missing Data Methods: Advances in Econometrics, 27B*, 179–233.

Emanuel, D., & Mcbeth, J. (1982). Further results on the constant elasticity of variance call option pricing model. *Journal of Financial Quantitative Analysis, 17*, 533–554.

Fermanian, J. D., & Salanié, B. (2004). A non-parametric simulated maximum likelihood estimation method. *Econometric Theory, 20*, 701–734.

Gallant, A. R., & Tauchen, G. (1996). Which moments to match. *Econometric Theory, 12*, 657–681.

Gallant, A. R., & Tauchen, G. (1997). Estimation of continuous time models for stock returns and interest rates. *Macroeconomic Dynamics, 1*, 135–168.

Goncalves, S., & White, H. (2002). The bootstrap of the mean for dependent heterogeneous arrays. *Econometric Theory, 18*, 1367–1384.

Hall, A. R., & Inoue, A. (2003). The large sample behavior of the generalized method of moments estimator in misspecified models. *Journal of Econometrics, 114*, 361–394.

Hansen, B. E. (1996). Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica, 64*, 413–430.

Heston, S. L. (1993). A closed form solution for option with stochastic volatility with applications to bond and currency options. *Review of Financial Studies, 6*, 327–344.

Hong, Y. (2001). Evaluation of out-of-sample probability density forecasts with applications to S&P 500 stock prices (Working paper). Cornell University.

Hong, Y. M., & Li, H. T. (2005). Nonparametric specification testing for continuous time models with applications to term structure interest rates. *Review of Financial Studies, 18*, 37–84.

Hong, Y. M., Li, H. T., & Zhao, F. (2007). Can the random walk model be beaten in out-of-sample density forecasts? Evidence from intraday foreign exchange rates. *Journal of Econometrics, 141*, 736–776.

Hull, J., & White, A. (1987). The pricing of options on assets with stochastic volatility. *Journal of Finance, 42*, 281–300.

Inoue. (2001). Testing for distributional change in time series. *Econometric Theory, 17*, 156–187.

Jiang, G. J. (1998). Nonparametric modelling of US term structure of interest rates and implications on the prices of derivative securities. *Journal of Financial and Quantitative Analysis, 33*, 465–497.

Karlin, S., & Taylor, H. M. (1981). *A second course in stochastic processes*. San Diego: Academic.

Kloeden, P. E., & Platen, E. (1999). *Numerical solution of stochastic differential equations*. New York: Springer.

Kolmogorov, A. N. (1933). Sulla Determinazione Empirica di una Legge di Distribuzione. *Giornale dell'Istitutodegli Attuari, 4*, 83–91.

Kristensen, D., & Shin, Y. (2008). Estimation of dynamic models with nonparametric simulated maximum likelihood (Creates Research paper 2008-58). University of Aarhus and Columbia University.

Künsch, H. R. (1989). The Jackknife and the bootstrap for general stationary observations. *Annals of Statistics, 17*, 1217–1241.
Marsh, T., & Rosenfeld, E. (1983). Stochastic processes for interest rates and equilibrium bond prices. *Journal of Finance, 38*, 635–646.
Meddahi, N. (2001). An Eigenfunction approach for volatility modeling (Working paper). University of Montreal.
Merton, C. R. (1973). Theory of rational option pricing. *The Bell Journal of Economics and Management Science, 4*, 141–183.
Nelson, D. B. (1990). ARCH as diffusion approximations. *Journal of Econometrics, 45*, 7–38.
Pardoux, E., & Talay, D. (1985). Discretization and simulation of stochastic differential equations. *Acta Applicandae Mathematicae, 3*, 23–47.
Piazzesi, M. (2001). Macroeconomic jump effects and the yield curve (Working paper). UCLA.
Piazzesi, M. (2004). *Affine term structure models*. Manuscript prepared for the Handbook of Financial Econometrics. University of California at Los Angeles.
Piazzesi, M. (2005). Bond yields and the federal reserve. *Journal of Political Economy, 113*, 311–344.
Politis, D. N., Romano, J. P., & Wolf, M. (1999). *Subsampling*. New York: Springer.
Pritsker, M. (1998). Nonparametric density estimators and tests of continuous time interest rate models. *Review of Financial Studies, 11*, 449–487.
Rosenblatt, M. (1952). Remarks on a multivariate transformation. *Annals of Mathematical Statistics, 23*, 470–472.
Singleton, K. J. (2006). *Empirical dynamic asset pricing – Model specification and econometric assessment*. Princeton: Princeton University Press.
Smirnov, N. (1939). On the estimation of the discrepancy between empirical curves of distribution for two independent samples. *Bulletin Mathematique de l'Universite' de Moscou*, *2*, fasc. 2.
Sullivan, R., Timmermann, A., & White, H. (1999). Data-snooping, technical trading rule performance, and the bootstrap. *The Journal of Finance, 54*(5), 1647–1691.
Sullivan, R., Timmermann, A., & White, H. (2001). Dangers of data-driven inference: The case of calendar effects in stock returns. *Journal of Econometrics, 105*, 249–286.
Thompson, S. B. (2008). Identifying term structure volatility from the LIBOR-swap curve. *Review of Financial Studies, 21*, 819–854.
van der Vaart, A. (1998). Asymptotic statistics. Cambridge, MA: Cambridge University Press.
Vasicek, O. A. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics, 5*, 177–188.
Vuong, Q. (1989). Likelihood ratio tests for model selection and non-nested hypotheses. *Econometrica, 57*, 307–333.
White, H. (1982). Maximum likelihood estimation of misspecified models. *Econometrica, 50*, 1–25.
White, H. (2000). A reality check for data snooping. *Econometrica, 68*, 1097–1126.
Wong, E. (1964). The construction of a class of stationary Markov processes. In R. Bellman (Ed.), *Sixtenn symposia in applied mathematics: Stochastic processes in mathematical physics and engineering*. Providence: American Mathematical Society.

Assessing the Performance of Estimators Dealing with Measurement Errors

57

Heitor Almeida, Murillo Campello, and Antonio F. Galvao

Contents

This article is based on Almeida, H., M. Campello, and A. Galvao, 2010, Measurement Errors in Investment Equations, Review of Financial Studies, 23, pages 3279–3382.

H. Almeida (✉)
University of Illinois at Urbana-Champaign, Champaign, IL, USA
e-mail: halmeida@illinois.edu

M. Campello
Cornell University, Ithaca, NY, USA
e-mail: campello@cornell.edu

A.F. Galvao
University of Iowa, Iowa City, IA, USA
e-mail: antonio-galvao@uiowa.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_57,

Abstract

We describe different procedures to deal with measurement error in linear models and assess their performance in finite samples using Monte Carlo simulations and data on corporate investment. We consider the standard instrumental variable approach proposed by Griliches and Hausman (Journal of Econometrics 31:93–118, 1986) as extended by Biorn (Econometric Reviews 19:391–424, 2000) [OLS-IV], the Arellano and Bond (Review of Economic Studies 58:277–297, 1991) instrumental variable estimator, and the higher-order moment estimator proposed by Erickson and Whited (Journal of Political Economy 108:1027–1057, 2000, Econometric Theory 18:776–799, 2002). Our analysis focuses on characterizing the conditions under which each of these estimators produce unbiased and efficient estimates in a standard "errors-in-variables" setting. In the presence of fixed effects, under heteroscedasticity, or in the absence of a very high degree of skewness in the data, the EW estimator is inefficient and returns biased estimates for mismeasured and perfectly measured regressors. In contrast to the EW estimator, IV-type estimators (OLS-IV and AB-GMM) easily handle individual effects, heteroscedastic errors, and different degrees of data skewness. The IV approach, however, requires assumptions about the autocorrelation structure of the mismeasured regressor and the measurement error. We illustrate the application of the different estimators using empirical investment models. Our results show that the EW estimator produces inconsistent results when applied to real-world investment data, while the IV estimators tend to return results that are consistent with theoretical priors.

Keywords

Investment equations • Measurement error • Monte Carlo simulations • Instrumental variables • GMM • Bias • Fixed effects • Heteroscedasticity • Skewness • High-order moments

57.1 Introduction

OLS estimators are the workhorse of empirical research in many fields in applied economics. Researchers see a number of advantages in these estimators. Most notably, they are easy to implement and the results they generate are easy to replicate. Another appealing feature of OLS estimators is that they easily accommodate the inclusion of individual (e.g., firm and time) idiosyncratic effects. Despite their popularity, however, OLS estimators are weak in dealing with the problem of errors in variables. When the independent (right-hand side) variables of an empirical model are mismeasured, coefficients estimated via standard OLS are inconsistent (*attenuation bias*). This poses a problem since, in practice, it is hard to think of any empirical proxies in applied research whose measurement is not a concern.

In this chapter, we describe three estimators that deal with the problem of mismeasurement, namely, the standard instrumental variable approach extended by Biorn (2000) [OLS-IV], the Arellano and Bond (1991) instrumental variable estimator [AB-GMM], and the higher-order moment estimator proposed by Erickson and Whited (2000, 2002) [henceforth, EW]. We also assess the performance of these estimators in finite samples using Monte Carlo simulations and illustrate their application using data on corporate investment.

While we provide a formal presentation in the next section, it is useful to discuss the intuition behind the estimation approaches we analyze. All approaches share the attractive property that they do not require the researcher to look for instruments outside the model being considered.[1] They differ, however, on how identification is achieved.

Both instrumental variable approaches rely on assumptions about the serial correlation of the latent variable and the innovations of the model (the model's disturbances and the measurement error). There are two conditions that must hold to ensure identification. First, the true value of the mismeasured regressor must have some degree of autocorrelation. In this case, lags of the mismeasured regressor are natural candidates for the instrumental set since they contain information about the current value of the mismeasured regressor.[2] This condition is akin to the standard requirement that the instrument be correlated with the variable of interest. The other necessary condition is associated with the *exclusion restriction* that is standard in IV methods and relates to the degree of serial correlation of the innovations. A standard assumption guaranteeing identification is that the measurement-error process is independently and identically distributed. This condition ensures that past values of the observed variables are uncorrelated with the current value of the measurement error, validating the use of lags of observed variables as instruments. Under certain conditions, one can also allow for autocorrelation in the measurement error. Examples in which identification works are when autocorrelation is constant over time or when it evolves according to a moving average process.[3] The first assumption ensures identification because it means that while lagged values of the measurement error are correlated with its current value, any past shocks to the measurement-error process do not persist over time. The moving average assumption allows for shocks to persist over time, but it imposes restrictions on the instrumental set. In particular, as we show below, it precludes the

[1]Naturally, if extraneous instruments are available, they can help solve the identification problem. See Rauh (2006) for the use of discontinuities in pension contributions as a source of variation in cash flows in an investment model. Bond and Cummins (2000) use information contained in financial analysts' forecasts to instrument for investment demand.

[2]Lags of the well-measured variable may also be included in the instrument set if they are believed to also contain information about the mismeasured one.

[3]See, among others, Biorn (2000), Wansbeek (2001), and Xiao et al. (2008).

use of shorter lags of observed variables in the instrument set, as the information contained in short lags may be correlated with the current value of the measurement error.

The EW estimator is based on high-order moments of residuals obtained by "partialling out" perfectly measured regressors from the dependent, observed mismeasured, and latent variables, as well as high-order moments of the innovations of the model. The key idea is to create a set of auxiliary equations as a function of these moments and cross-moments. Implementation then requires a high degree of skewness in the distribution of the partialled out latent variable. Our analysis also shows that the presence of individual fixed effects and heteroscedasticity also impacts the performance of the EW estimator, particularly so if they are both present in the data process.

We perform a series of Monte Carlo simulations to compare the performance of the EW and IV estimators in finite samples. Emulating the types of environments commonly found by empirical researchers, we set up a panel data model with individual fixed effects and potential heteroscedasticity in the errors. Monte Carlo experiments enable us to study those estimators in a "controlled environment," where we can investigate the role played by each element (or assumption) of an estimator in evaluating its performance. Our simulations compare the EW and IV (OLS-IV and AB-GMM) estimators in terms of bias and root mean squared error (RMSE), a standard measure of efficiency.

We consider several distributional assumptions to generate observations and errors. Experimenting with multiple distributions is important because researchers often find a variety of distributions in real-world applications and because one ultimately does not observe the distribution of the mismeasurement term. Since the EW estimator is built around the notion of skewness of the relevant distributions, we experiment with three skewed distributions (lognormal, chi-square, and F-distribution), using the standard normal (non-skewed) as a benchmark. The simulations also allow for significant correlation between mismeasured and well-measured regressors (as in Erickson and Whited 2000, 2002), so that the attenuation bias of the mismeasured regressor affects the coefficient of the well-measured regressor.

Our simulation results can be summarized as follows. First, we examine the identification test proposed by Erickson and Whited (2002). This is a test that the data contain a sufficiently high degree of skewness to allow for the identification of their model. We study the *power* of the EW identification test by generating data that *do not* satisfy its null hypothesis of non-skewness. In this case, even for the most skewed distribution (lognormal), the test rejects the null hypothesis only 47 % of the time – this is far less than desirable, given that the null is false. The power of the test becomes even weaker after we treat the data for the presence of fixed effects in the true model ("within transformation"). In this case, the rejection rate under the lognormal distribution drops to 43 %. The test's power declines even further when we consider alternative skewed distributions (chi-square and F-distributions). The upshot of this first set of experiments is that the EW model too often rejects data that are generated to fit its identifying assumptions. These findings may help explain some of the difficulties previous researchers have reported when attempting to implement the EW estimator.

We then study the bias and efficiency of the EW and IV estimators. Given that the true model contains fixed effects, it is appropriate to apply the within transformation to the data. However, because most empirical implementations of the EW estimator have used data in "level form" (i.e., not treated for the presence of fixed effects),[4] we also experiment with cases in which we do not apply the within transformation. EW propose three estimators that differ according to the number of moments used: GMM3, GMM4, and GMM5. We consider all of them in our experiments.

In a first round of simulations, we impose error homoscedasticity. When we implement the EW estimator with the data in level form, we find that the coefficients returned are significantly biased even when the data have a high degree of skewness (i.e., under the lognormal case, which is EW's preferred case). Indeed, for the mismeasured regressors the EW biases are in excess of 100 % of their true value. As should be expected, the performance of the EW estimator improves once the within transformation is used. In the case of the lognormal distribution, the EW estimator bias is relatively small. In addition, deviations from the lognormal assumption tend to generate significant biases for the EW estimator.

In a second round of simulations, we allow for heteroscedasticity in the data. We focus our attention on simulations that use data that are generated using a lognormal distribution after applying the within transformation (the best case scenario for the EW estimator). Heteroscedasticity introduces heterogeneity to the model and consequently to the distribution of the partialled out dependent variable, compromising identification in the EW framework. The simulations show that the EW estimator is biased and inefficient for both the mismeasured and well-measured regressors. In fact, biases emerge even for very small amounts of heteroscedasticity, where we find biases of approximately 40 % for the mismeasured regressor. Paradoxically, biases "switch signs" depending on the degree of heteroscedasticity that is allowed for in the model. For instance, for small amounts of heteroscedasticity, the bias of the mismeasured regressor is negative (i.e., the coefficient is biased downwards). However, the bias turns positive for a higher degree of heteroscedasticity. Since heteroscedasticity is a naturally occurring phenomenon in corporate data, our simulations imply that empirical researchers might face serious drawbacks when using the EW estimator.

Our simulations also show that, in contrast to the EW estimator, the bias in the IV estimates is small and insensitive to the degree of skewness and heteroscedasticity in the data. Focusing on the OLS-IV estimator, we consider the case of time-invariant correlation in the error structure and use the second lag of the

[4]Examples are Whited (2001, 2006), Hennessy (2004), and Colak and Whited (2007).

observed mismeasured variable as an instrument for its current (differenced) value.[5] We also allow the true value of the mismeasured regressor to have a moderate degree of autocorrelation. Our results suggest that the OLS-IV estimator renders fairly unbiased estimates. In general, that estimator is also distinctly more efficient than the EW estimator.

We also examine the OLS-IV estimator's sensitivity to the autocorrelation structures of the mismeasured regressor and the measurement error. First, we consider variations in the degree of autocorrelation in the process for the true value of the mismeasured regressor. Our simulations show that the IV bias is largely insensitive to variations in the autoregressive (AR) coefficient (except for extreme values of the AR coefficient). Second, we replace the assumption of time-invariant autocorrelation in the measurement error with a moving average (MA) structure. Our simulations show that the OLS-IV bias remains small if one uses long enough lags of the observable variables as instruments. In addition, provided that the instrument set contains suitably long lags, the results are robust to variations in the degree of correlation in the MA process. As we discuss below, understanding these features (and limitations) of the IV approach is important given that the researcher will be unable to pin down the process followed by the measurement-error process.

To illustrate the performance of these alternative estimators on real data, in the final part of our analysis, we estimate empirical investment models under the EW and IV frameworks. Concerns about measurement errors have been emphasized in the context of the empirical investment model introduced by Fazzari et al. (1988), where a firm's investment is regressed on a proxy for investment demand (Tobin's q) and cash flows. Theory suggests that the correct proxy for a firm's investment demand is captured by *marginal* q, but this quantity is unobservable and researchers use instead its measurable proxy, *average* q. Since the two variables are not the same, a measurement problem naturally arises (Hayashi 1982; Poterba 1988). Following Fazzari et al. (1988), investment-cash flow sensitivities became a standard metric in the literature that examines the impact of financing imperfections on corporate investment (Stein 2003). These empirical sensitivities are also used for drawing inferences about efficiency in internal capital markets (Lamont 1997; Shin and Stulz 1998), the effect of agency on corporate spending (Hadlock 1998; Bertrand and Mullainathan 2005), the role of business groups in capital allocation (Hoshi et al. 1991), and the effect of managerial characteristics on corporate policies (Bertrand and Schoar 2003; Malmendier and Tate 2005).

Theory does not pin down exact values for the expected coefficients on q and cash flow in an investment model. However, two conditions would seem reasonable in practice. First, given that the estimator is addressing measurement error in q that may be "picked up" by cash flow (joint effects of attenuation bias and regressor

[5]The results for the Arellano–Bond GMM estimator are similar to those of the OLS-IV estimator. To save space and because the OLS-IV estimator is easier to implement, we focus on this estimator.

covariance), we should expect the q coefficient to go up and the cash flow coefficient to go down, when compared to standard (likely biased) OLS estimates. Second, we would expect the q and cash flow coefficients to be nonnegative after addressing the problem of mismeasurement. If the original q-theory of investment holds and the estimator does a good job of addressing mismeasurement, then the cash flow coefficient would be zero. Alternatively, the cash flow coefficient could be positive because of financing frictions.[6]

Using data from Compustat from 1970 to 2005, we estimate investment equations in which investment is regressed on proxies for q and cash flow. Before doing so, we perform standard tests to check for the presence of individual fixed effects and heteroscedasticity in the data. In addition, we perform the EW identification test to check whether the data contain a sufficiently high degree of skewness.

Our results are as follows. First, our tests reject the hypotheses that the data do not contain firm-fixed effects and that errors are homoscedastic. Second, the EW identification tests indicate that the data fail to display sufficiently high skewness. These initial tests suggest that the EW estimator is not suitable for standard investment equation applications. In fact, we find that, when applied to the data, the EW estimator returns coefficients for q and cash flow that are highly unstable across different years. Moreover, following the EW procedure for panel models (which comprises combining yearly cross-sectional coefficients into single estimates), we obtain estimates for q and cash flow that do not satisfy the conditions discussed above. In particular, EW estimators do not reduce the cash flow coefficient relative to that obtained by standard OLS, while the q coefficient is never statistically significant. In addition, those estimates are not robust with respect to the number of moments used: EW's GMM3, GMM4, and GMM5 models procedure results that are inconsistent with one another. These results suggest that the presence of heteroscedasticity and fixed effects in real-world investment data hampers identification when using the EW estimator.

In contrast to EW, the OLS-IV procedure yields estimates that are fairly sensible. The q coefficient goes up by a factor of 3–5, depending on the set of instruments used. At the same time, the cash flow coefficient goes down by about two-thirds of the standard OLS value. Similar conclusions apply to the AB-GMM estimator. We also examine the robustness of the OLS-IV to variations in the set of instruments used in the estimation, including sets that feature only longer lags of the variables in the model. The OLS-IV coefficients remain fairly stable after such changes. These results suggest that real-world investment data likely satisfies the assumptions that are required for identification of IV estimators.

The remainder of the chapter is structured as follows. We start the next section discussing in detail the EW estimator, clarifying the assumptions that are needed for its implementation. Subsequently, we show how alternative IV models deal with

[6]See Hubbard (1998) and Stein (2003) for comprehensive reviews. We note that the presence of financing frictions does not necessarily imply that the cash flow coefficient should be positive. See Chirinko (1993) and Gomes (2001) for arguments suggesting that financing frictions are not sufficient to generate positive cash flow coefficients.

the errors-in-variables problem. In Sect. 57.3, we use Monte Carlo simulations to examine the performance of alternative estimators in small samples and when we relax the assumptions that are required for identification. In Sect. 57.4, we take our investigation to actual data, estimating investment regressions under the EW and IV frameworks. Section 57.5 concludes the chapter.

57.2 Dealing with Mismeasurement: Alternative Estimators

57.2.1 The Erickson–Whited Estimator

In this section, we discuss the estimator proposed in companion papers by Erickson and Whited (2000, 2002). Those authors present a two-step generalized method of moments (GMM) estimator that exploits information contained in the high-order moments of residuals obtained from perfectly measured regressors (similar to Cragg 1997). We follow EW and present the estimator using notation of cross-section estimation. Let (y_i, z_i, x_i), $i = 1,\ldots,n$, be a sequence of observable vectors, where $x_i \equiv (x_{i1},\ldots,x_{iJ})$ and $z_i \equiv (1, z_{i1},\ldots, z_{iL})$. Let $(u_i, \chi_i, \varepsilon_i)$ be a sequence of unobservable vectors, where $\chi_i \equiv (\chi_{i1},\ldots,\chi_{iJ})$ and $\varepsilon_i \equiv (\varepsilon_{i1},\ldots,\varepsilon_{iJ})$. Consider the following model:

$$y_i = z_i a + \chi_i \beta + u_i \tag{57.1}$$

where y_i is the dependent variable, z_i is a perfectly measured regressor, χ_i is a mismeasured regressor, u_i is the innovation of the model, and $\alpha \equiv (\alpha_0, \alpha_i, \ldots,\alpha_L)'$ and $\beta \equiv (\beta_1, \ldots,\beta_J)'$. The measurement error is assumed to be additive such that

$$x_i = \chi_i + \varepsilon_i \tag{57.2}$$

where x_i is the observed variable and ε_i is the measurement error. The observed variables are y_i, z_i, and x_i; and by substituting Eq. 57.2 in Eq. 57.1, we have

$$y_i = z_i \alpha + x_i \beta + v_i,$$

where $v_i = u_i - \varepsilon_i \beta$. In the new regression, the observable variable x_i is correlated with the innovation term v_i, causing the coefficient of interest, β, to be biased.

To compute the EW estimator, it is necessary to first partial out the effect of the well-measured variable, z_i, in Eqs. 57.1 and 57.2 and rewrite the resulting expressions in terms of residual populations:

$$y_i - z_i \mu_y = \eta_i \beta + u_i \tag{57.3}$$

$$x_i - z_i \mu_x = \eta_i + \varepsilon_i, \tag{57.4}$$

where $(\mu_y, \mu_x, \mu_\chi) = [E(z_i'z_i)]^{-1}E[z_i'(y_i, x_i, \chi_i)]$ and $\eta_i \equiv \chi_i - z_i\mu_\chi$. For the details of this derivation, see Erickson and Whited (2002, p. 779). One can then consider a two-step estimation approach, where the first step is to substitute least squares estimates $(\hat{\mu}_y, \hat{\mu}_x) \equiv \left(\sum_{i=1}^{n} z_i'z_i\right)^{-1}\sum_{i=1}^{n} z_i'(y_i, x_i)$ into Eqs. 57.3 and 57.4 to obtain a lower dimensional errors-in-variables model. The second step consists of estimating β using GMM using high-order sample moments of $y_i - z_i\hat{\mu}_y$ and $x_i - z_i\hat{\mu}_x$. Estimates of α are then recovered via $\mu_y = \alpha + \mu_x\beta$. Thus, the estimators are based on equations giving the moments of $y_i - z_i\mu_y$ and $x_i - z_i\mu_x$ as functions of β and the moments of $(u_i, \varepsilon_i, \eta_i)$.

To give a concrete example of how the EW estimator works, we explore the case of $J = 1$. The more general case is discussed below. By substituting

$$\hat{\mu}_y, \hat{\mu}_x \equiv \left(\sum_{i=1}^{n} z_i'z_i\right)^{-1}\sum_{i=1}^{n} z_i{}'(y_i, x_i)$$

into Eqs. 57.1 and 57.2, one can estimate β, $E(u_i^2)$, $E(\varepsilon_i^2)$, and $E(\eta_i^2)$ via GMM. Estimates of the lth element of a are obtained by substituting the estimate of β and the lth elements of $\hat{\mu}_y$ and $\hat{\mu}_x$ into

$$\alpha_l = \mu_{yl} - \mu_{xl}\beta, \quad l \neq 0.$$

There are three second-order moment equations:

$$E\left[\left(y_i - z_i\mu_y\right)^2\right] = \beta^2 E\left(\eta_i^2\right) + E\left(u_i^2\right) \tag{57.5}$$

$$E\left[\left(y_i - z_i\mu_y\right)(x_i - z_i\mu_x)\right] = \beta E\left(\eta_i^2\right) \tag{57.6}$$

$$E\left[(x_i - z_i\mu_x)^2\right] = E\left(\eta_i^2\right) + E\left(\varepsilon_1^2\right). \tag{57.7}$$

The left-hand side quantities are consistently estimable, but there are only three equations with which to estimate four unknown parameters on the right-hand side. The third-order product moment equations, however, consist of two equations in two unknowns:

$$E\left[\left(y_i - z_i\mu_y\right)^2(x_i - z_i\mu_x)\right] = \beta^2 E\left(\eta_i^3\right), \tag{57.8}$$

$$E\left[\left(y_i - z_i\mu_y\right)(x_i - z_i\mu_x)^2\right] = \beta E\left(\eta_i^3\right). \tag{57.9}$$

It is possible to solve these two equations for β. Crucially, a solution exists if the identifying assumptions $\beta \neq 0$ and $E(\eta_i^3) \neq 0$ are true, and one can test the contrary hypothesis (i.e., $\beta \neq 0$ and/or $E(\eta_i^3) = 0$) by testing whether their sample counterparts are significatively different from zero.

Given β, Eqs. 57.5, 57.6, 57.7, and 57.9 can be solved for the remaining right-hand side quantities. One obtains an overidentified equation system by combining Eqs. 57.5, 57.6, 57.7, 57.8, and 57.9 with the fourth-order product moment equations, which introduce only one new quantity, $E(\eta_i^4)$:

$$E\left[\left(y_i - z_i\mu_y\right)^3\left(x_i - z_i\mu_x\right)\right] = \beta^3 E\left(\eta_i^4\right) + 3E\left(\eta_i^2\right)E\left(u_i^2\right), \tag{57.10}$$

$$\begin{aligned} E\left[\left(y_i - z_i\mu_y\right)^2\left(x_i - z_i\mu_x\right)^2\right] = \beta^2\left[E\left(\eta_i^4\right) + E\left(\eta_i^2\right)E\left(u_i^2\right)\right] \\ + E\left(u_i^2\right)\left[E\left(\eta_i^2\right) + E\left(\varepsilon_i^2\right)\right], \end{aligned} \tag{57.11}$$

$$E\left[\left(y_i - z_i\mu_y\right)\left(x_i - z_i\mu_x\right)^3\right] = \beta\left[E\left(\eta_i^4\right) + 3E\left(\eta_i^2\right)E\left(\varepsilon_i^2\right)\right]. \tag{57.12}$$

The resulting eight-equation system Eqs. 57.5, 57.6, 57.7, 57.8, 57.9, 57.10, 57.11, and 57.12 contains the six unknowns (β, $E(u_i^2)$, $E(\varepsilon_i^2)$, $E(\eta_i^2)$, $E(\eta_i^3)$, $E(\eta_i^4)$). It is possible to estimate this vector by numerically minimizing a quadratic form that minimizes asymptotic variance.

The conditions imposed by EW imply restrictions on the residual moments of the observable variables. Such restrictions can be tested using the corresponding sample moments. EW also propose a test for residual moments that is based on several assumptions.[7] These assumptions imply testable restrictions on the residuals from the population regression of the dependent and proxy variables on the perfectly measured regressors. Accordingly, one can develop Wald-type partially adjusted statistics and asymptotic null distributions for the test. Empirically, one can use the Wald test statistic and critical values from a chi-square distribution to test whether the last moments are equal to zero. This is an identification test, and if in a particular application one cannot reject the null hypothesis, then the model is unidentified and the EW estimator may not be used. We study the finite sample performance of this test and its sensitivity to different data-generating processes in the next section.

It is possible to derive more general forms of the EW estimator. In particular, the EW estimators are based on the equations for the moments of $y_i - z_i\mu_y$ and $x_i - z_i\mu_x$ as functions of β and the moments u_i, ε_i, and η_i. To derive these equations, write Eq. 57.3 as $y_i - z_i\mu_y = \sum_{j=1}^{J} \eta_{ij}\beta_j + u_i$, where J is the number of well-measured regressors and the jth equation in Eq. 57.4 as $x_{ij} - z_i\mu_{xj} = \eta_{ij} + \varepsilon_{ij}$, where μ_{xj} is the jth column of and μ_x and $(\eta_{ij}, \varepsilon_{ij})$ is the jth row of $(\eta_{ij}', \varepsilon_{ij}')$. Next write

[7] First, the measurement errors, the equation error, and all regressors have finite moments of sufficiently high order. Second, the regression error and the measurement error must be independent of each other and of all regressors. Third, the residuals from the population regression of the unobservable regressors on the perfectly measured regressors must have a nonnormal distribution.

$$E\left[\left(y_i - z_i\mu_y\right)^{r_0}\prod_{j=1}^{J}\left(x_i - z_i\mu_x\right)^{r_j}\right] = E\left[\left(\sum_{j=1}^{J}\eta_i\beta + u_i\right)^{r_0}\prod_{j=1}^{J}\left(\eta_i + \varepsilon_i\right)^{r_j}\right], \tag{57.13}$$

where $(r_0, r_1,\ldots, r_J)$ are nonnegative integers. After expanding the right-hand side of Eq. 57.13, using the multinomial theorem, it is possible to write the above moment condition as

$$E[g_i(\mu)] = c(\theta),$$

where $\mu = vec(\mu_y, \mu_x)$, $g_i(\mu)$ is a vector of distinct elements of the form $\left(y_i - z_i\mu_y\right)^{r_0}\prod_{j=1}^{J}\left(x_i - z_i\mu_x\right)^{r_j}$, $c(\theta)$ contains the corresponding expanded version of the right-hand side of Eq. 57.13, and θ is a vector containing the elements of β and the moments of $(u_i, \varepsilon_i, \eta_i)$. The GMM estimator θ is defined as

$$\hat{\theta} = \arg \min_{t\in\Theta} \left(\overline{g}_i(\hat{\mu}) - c(t)\right)'\hat{W}\left(\overline{g}_i(\hat{\mu}) - c(t)\right),$$

where $\bar{g}_i(s) \equiv \sum_{t=1}^{n} g_i(s)$ for all s, and $\hat{W}$ is a positive definite matrix. Assuming a number of regularity conditions,[8] the estimator is consistent and asymptotically normal.

It is important to notice that the estimator proposed by Erickson and Whited (2002) was originally designed for cross-sectional data. To accommodate a panel-like structure, Erickson and Whited (2000) propose transforming the data before the estimation using the within transformation or differencing. To mimic a panel structure, the authors propose the idea of combining different cross-sectional GMM estimates using a minimum distance estimator (MDE).

The MDE estimator is derived by minimizing the distance between the auxiliary parameter vectors under the following restrictions:

$$f\left(\beta, \hat{\theta}\right) = H\beta - \hat{\theta} = 0,$$

where the $R \cdot K \times K$ matrix H imposes $(R - 1) \cdot K$ restrictions on θ. The $R\,K \times 1$ vector $\hat{\theta}$ contains the R stacked auxiliary parameter vectors, and β is the parameter of interest. Moreover, H is defined by an $R \cdot K \times K$ – dimensional stacked identity matrix.

---

[8] More specifically, these conditions are as follows: $(z_i, \chi_i, u_i, \varepsilon_i)$ is an independent and identically distributed sequence; u_i and the elements of z_i, χ_i, and ε_i, have finite moments of every order; (u_i, ε_i) is independent of (z_i, χ_i), and the individual elements in (u_i, ε_i) are independent of each other; $E(u_i) = 0$ and $E(\varepsilon_i) = 0$; $E[(z_i, \chi_i)'(z_i, \chi_i)]$ is positive definite; every element of β is nonzero; and the distribution of η satisfies $E[(\eta_i c)^3] \neq 0$ for every vector of constants $c = (c_1,\cdots,c_J)$ having at least one nonzero element.

The MDE is given by the minimization of

$$D(\beta) = f\left(\beta, \hat{\theta}\right)' \hat{V}\left[\hat{\theta}\right]^{-1} f\left(\beta, \hat{\theta}\right), \tag{57.14}$$

where $\hat{V}[\hat{\theta}]$ is the common estimated variance–covariance matrix of the auxiliary parameter vectors.

In order to implement the MDE, it is necessary to determine the covariances between the cross-sections being pooled. EW propose to estimate the covariance by using the covariance between the estimators' respective influence functions.[9] The procedure requires that each cross-section have the same sample size, that is, the panel needs to be balanced.

Thus, minimization of D in Eq. 57.14 leads to

$$\hat{\beta} = \left(H' \hat{V}\left[\hat{\theta}\right]^{-1} H\right)^{-1} H' \hat{V}\left[\hat{\theta}\right]^{-1} \hat{\theta},$$

with variance–covariance matrix:

$$\hat{V}\left[\hat{\beta}\right] = \left(H' \hat{V}\left[\hat{\theta}\right]^{-1} H\right)^{-1}.$$

H is a vector in which R is the number of GMM estimates available (for each time period) and $K = 1$, $\hat{\theta}$ is a vector containing all the EW estimates for each period, and β is the MDE of interest. In addition, $V\left[\hat{\theta}\right]$ is a matrix carrying the estimated variance–covariance matrices of the GMM parameter vectors.

57.2.2 An OLS-IV Framework

In this section, we revisit the work of Griliches and Hausman (1986) and Biorn (2000) to discuss a class of OLS-IV estimators that can help address the errors-in-variables problem.

Consider the following single-equation model:

$$y_{it} = \gamma_i + \chi_{it}\beta + u_{it}, \quad i = 1, \ldots N, \quad t = 1, \ldots, T, \tag{57.15}$$

where u_{it} is independently and identically distributed, with mean zero and variance σ_u^2, and $\text{Cov}(\chi_{it}, u_{is}) = \text{Cov}(\gamma_i, u_{is}) = 0$ for any t and s, but $\text{Cov}(\gamma_i, \chi_{it}) \neq 0$, y is an observable scalar, χ is a $1 \times K$ vector, and β is $K \times 1$ vector. Suppose we do not observe χ_{it} itself, but rather the error-ridden measure:

[9] See Erickson and Whited (2002) Lemma 1 for the definition of their proposed influence function.

$$x_{it} = \chi_{it} + \varepsilon_{it}, \tag{57.16}$$

where $\mathrm{Cov}(\chi_{it}, \varepsilon_{it}) = \mathrm{Cov}(\gamma_i, \varepsilon_{it}) = \mathrm{Cov}(u_{is}, \varepsilon_{it}) = 0$, $\mathrm{Var}(\varepsilon_{it}) = \sigma_\varepsilon^2$, $\mathrm{Cov}(\varepsilon_{it}, \varepsilon_{it-1}) = \gamma_\varepsilon \sigma_\varepsilon^2$, and ε is a $1 \times K$ vector. If we have a panel data with $T > 3$, by substituting Eq. 57.16 in Eq. 57.15, we can take first differences of the data to eliminate the individual effects γ_i and obtain

$$y_{it} - y_{it-1} = (x_{it} - x_{it-1})\beta + [(u_{it} - \varepsilon_{it}\beta) - (u_{it-1} - \varepsilon_{it-1}\beta)] \tag{57.17}$$

Because of the correlation between the mismeasured variable and the innovations, the coefficient of interest is known to be biased.

Griliches and Hausman (1986) propose an instrumental variable approach to reduce the bias. If the measurement error ε_{it} is i.i.d. across i and t, and x is serially correlated, then, for example, x_{it-2}, x_{it-3}, or $(x_{it-2} - x_{it-3})$ are valid as instruments for $(x_{it} - x_{it-1})$. The resulting instrumental variable estimator is consistent even though T is finite and N might tend to infinity.

As emphasized by Erickson and Whited (2000), for some applications, the assumption of i.i.d. measurement error can be seen as too strong. Nonetheless, it is possible to relax this assumption to allow for autocorrelation in the measurement errors. While other alternatives are available, here we follow the approach suggested by Biorn (2000).[10]

Biorn (2000) relaxes the i.i.d. condition for innovations in the mismeasured equation and proposes alternative assumptions under which consistent IV estimators of the coefficient of the mismeasured regressor exists. Under those assumptions, as we will show, one can use the lags of the variables already included in the model as instruments. A notable point is that the consistency of these estimators is robust to potential correlation between individual heterogeneity and the latent regressor.

Formally, consider the model described in Eqs. 57.15 and 57.16 and assume that $(\chi_{it}, u_{it}, \varepsilon_{it}, \gamma_i)$ are independent across individuals. For the necessary orthogonality assumptions, we refer the reader to Biorn (2000), since these are quite standard. More interesting are the assumptions about the measurement errors and disturbances. The standard Griliche–Hausman's assumptions are
(A1) $E(\varepsilon'_{it}\varepsilon_{i\theta}) = 0_{KK}$, $t \neq \theta$,
(A2) $E(u_{it}u_{i\theta}) = 0$, $t \neq \theta$
which impose non-autocorrelation on innovations. It is possible to relax these assumptions in different ways. For example, we can replace (A1) and (A2) with
(B1) $E(\varepsilon'_{it}\varepsilon_{i\theta}) = 0_{KK}$, $|t - \theta| > \tau$,
(B2) $E(u_{it}u_{i\theta}) = 0$, $|t - \theta| > \tau$,

[10] A more recent paper by Xiao et al. (2008) also shows how to relax the classical Griliches–Hausman assumptions for measurement error models.

This set of assumptions is weaker since (B1) and (B2) allow for a vector moving average (MA) structure up to order τ ($\geq$1) for the innovations. Alternatively, one can use the following assumptions:

(C1) $E(\varepsilon_{it}'\varepsilon_{t\theta})$ is invariant to t, θ, $t \neq \theta$.

(C2) $E(u_{it}u_{i\theta})$ is invariant to t, θ, $t \neq \theta$.

Assumptions (C1) and (C2) allow for a different type of autocorrelation, more specifically they allow for any amount of autocorrelation that is time invariant. Assumptions (C1) and (C2) will be satisfied if the measurement errors and the disturbances have individual components, say $\varepsilon_{it} = \varepsilon_{1i} + \varepsilon_{2it}$, $u_{it} = u_{1i} + u_{2it}$, where ε_{1i}, ε_{2it}, u_{1i}, u_{2it} i.i.d. Homoscedasticity of ε_{it} and/or u_{it} across i and t need not be assumed; the model accommodates various forms of heteroscedasticity.

Biorn also considers assumptions related to the distribution of the latent regressor vector χ_{it}:

(D1) $\mathrm{E}(\chi_{it})$ is invariant to t.

(D2) $\mathrm{E}(\gamma_i\chi_{it})$ is invariant to t.

Assumptions (D1) and (D2) hold when χ_{it} is stationary. Note that χ_{it} and i need not be uncorrelated.

To ensure identification of the slope coefficient vector when panel data are available, it is necessary to impose restrictions on the second-order moments of the variables (χ_{it}, u_{it}, ε_{it}, γ_i). For simplicity, Biorn assumes that this distribution is the same across individuals and that the moments are finite. More specifically, $C(\chi_{it}, \chi_{it}) = \sum_{t\theta}^{\chi\chi}$, $E(\chi_{it}\eta_i) = \sum_t^{\chi n}$, $E(\varepsilon_{it}'\varepsilon_{it}) = \sum_{t\theta}^{\varepsilon\varepsilon}$, $E(u_{it}u_{it}) = \sigma_{t\theta}^{uu}$, $E(\eta_i^2) = \sigma^{\eta\eta}$, where C denotes the covariance matrix operator. Then, it is possible to derive the second-order moments of the observable variables and show that they only depend on these matrices and the coefficient β.[11] In this framework, there is no need to use assumptions based on higher-order moments.

Biorn proposes several strategies to estimate the slope parameter of interest. Under the OLS-IV framework, he proposes estimation procedures of two kinds:

- *OLS-IV A*: The equation is transformed to differences to remove individual heterogeneity and is estimated by OLS-IV. Admissible instruments for this case are the level values of the regressors and/or regressands for other periods.
- *OLS-IV B*: The equation is kept in level form and is estimated by OLS-IV. Admissible instruments for this case are differenced values of the regressors and/or regressands for other periods.

Using moment conditions from the OLS-IV framework, one can define the estimators just described. In particular, using the mean counterpart and the moment conditions, one can formally define the *OLS-IV A* and *OLS-IV B* estimators.

[11]Formally, one can show that $C(x_{it}, x_{i\theta}) = \sum_{t\theta}^{\chi\chi} + \sum_{t\theta}^{\varepsilon\varepsilon}$, $E(x_{it}, y_{i\theta}) = \sum_{t\theta}^{\chi\chi}\beta + \sum_t^{\chi\eta}$, and $E(y_{it}, y_{i\theta}) = \beta' \sum_{t\theta}^{\chi\chi}\beta + \sum_t^{xn}\beta + \beta^{H\prime}(\sum_\theta^{xn})' + \sigma_{t\theta}^{uu} + \sigma^{\eta\eta}$.

In particular, the estimator for *OLS-IV A* can be defined as

$$\hat{\beta}_{xp(t\theta)} = \left[\sum_{i=1}^{N} x'_{ip}(\Delta x_{it\theta})\right]^{-1}\left[\sum_{i=1}^{N} x'_{ip}(\Delta y_{it\theta})\right],$$

where (t, θ, p) are indices. Let the dimension of β be defined by K. If $K = 1$, it is possible to define the following estimator for a given (t, θ, p):

$$\hat{\beta}_{yp(t\theta)} = \left[\sum_{i=1}^{N} y_{ip}(\Delta x_{it\theta})\right]^{-1}\left[\sum_{i=1}^{N} y_{ip}(\Delta y_{it\theta})\right].$$

If $K > 1$, the latter estimator is infeasible, but it is possible to modify the former estimator by replacing one element in x'_{ip} by y_{ip}.

The estimator for *OLS-IV B* (equation in level and instruments in difference) can be defined as

$$\hat{\beta}_{x(pq)t} = \left[\sum_{i=1}^{N} (\Delta x_{ipq})' x_{it}\right]^{-1}\left[\sum_{i=1}^{N} (\Delta x_{ipq})' y_{it}\right].$$

As in the previous case, if the dimension of β, K is equal to 1, it is possible to define the following estimator for (t, p, q):

$$\hat{\beta}_{y(pq)t} = \left[\sum_{i=1}^{N} (\Delta y_{ipq}) x_{it}\right]^{-1}\left[\sum_{i=1}^{N} (\Delta y_{ipq}) y_{it}\right].$$

If $K > 1$, the latter estimator is infeasible, but it is possible to modify the former estimator by replacing one element in Δx_{ip} by Δy_{ip}.

For some applications, it might be useful to impose weaker conditions on the autocorrelation of measurement errors and disturbances. In this case, it is necessary to restrict slightly further the conditions on the instrumental variables. More formally, if one replaces assumptions (A1) and (A2), or (C1) and (C2), by the weaker assumptions (B1) and (B2), then it is necessary to ensure that the IV set has a lag of at least $\tau - 2$ and/or lead of at least $\tau + 1$ periods of the regressor in order to "clear" the τ period memory of the MA process. Consistency of these estimators is discussed in Biorn (2000).[12]

To sum up, there are two simple ways to relax the standard assumption of i.i.d. measurement errors. Under the assumption of time-invariant autocorrelation, the set of instruments can contain the same variables used under Griliches–Hausman.

[12]In particular, if $|t - p|$, $|\theta - p| > \tau$, then (B1) and rank $(E[\chi'_{ip}(\Delta\chi_{it\theta})]) = K$ for some $p \neq t \neq 0$ ensure consistency of *OLS–IV B*, $\hat{\beta}_{xp(t\theta)}$, and (B2) and the same rank condition ensure consistency of $\hat{\beta}_{xy(t\theta)}$. In the same way, if $|p - t|$, $q - t > \tau$, (B1), (D1), (D2), and rank $(E[(\Delta\chi_{ipq})'\chi_{it})]) = K$ for some $p \neq q \neq t$ ensure consistency of *OLS-IV B*, $\hat{\beta}_{x(pq)t}$, and (B2), (D1), (D2), and the same rank condition ensure consistency of $\hat{\beta}_{y(pq)t}$.

For example, if one uses the *OLS-IV A* estimator (equation in differences and instruments in levels), then twice-lagged levels of the observable variables can be used as instruments. Under a moving average structure for the innovations in the measurement error [assumptions (B1) and (B2)], identification requires the researcher to use longer lags of the observable variables as instruments. For example, if the innovations follow an MA(1) structure, then consistency of the *OLS-IV A* estimator requires the use of instruments that are lagged three periods and longer. Finally, identification requires the latent regressor to have some degree of autocorrelation (since lagged values are used as instruments). Our Monte Carlo simulations will illustrate the importance of these assumptions and will evaluate the performance of the OLS-IV estimator under different sets of assumptions about the structure of the errors.

57.2.3 GMM Estimator

Within the broader instrumental variable approach, we also consider an entire class of GMM estimators that deal with mismeasurement. These GMM estimators are close to the OLS-IV estimator discussed above but may attain appreciable gains in efficiency by combining numerous orthogonality conditions [see Biorn (2000) for a detailed discussion]. GMM estimators that use all the available lags at each period as instruments for equations in first differences were proposed by Holtz-Eakin et al. (1988) and Arellano and Bond (1991). We provide a brief discussion in turn.

In the context of a standard investment model, Blundell et al. (1992) use GMM allowing for correlated firm-specific effects, as well as endogeneity (mismeasurement) of q. The authors use an instrumental variable approach on a first-differenced model in which the instruments are weighted optimally so as to form the GMM estimator. In particular, they use q_{it-2} and twice-lagged investments as instruments for the first-differenced equation for firm i in period t. The Blundell, Bond, Devereux, and Schiantarelli estimators can be seen as an application of the GMM instrumental approach proposed by Arellano and Bond (1991), which was originally applied to a dynamic panel.

A GMM estimator for the errors-in-variables model of Eq. 57.17 based on IV moment conditions takes the form

$$\hat{\beta} = \left[\left(\Delta x' Z\right) V_N^{-1} \left(Z' \Delta x\right)\right]^{-1} \left(\Delta x' Z\right) V_N^{-1} \left(Z' \Delta y\right),$$

where Δx is the stacked vector of observations on the first difference of the mismeasured variable and Δy is the stacked vector of observations on the first difference of the dependent variable. As in Blundell et al. (1992), the instrument matrix Z has the following form[13]:

[13] In models with exogenous explanatory variables, Z_i may consist of sub-matrices with the block diagonal (exploiting all or part of the moment restrictions), concatenated to straightforward one-column instruments.

$$Z_i = \begin{pmatrix} x_1 & 0 & 0 & \cdots & 0 & \cdots & 0 \\ 0 & x_1 & x_2 & \cdots & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & x_1 & \cdots & x_{T-2} \end{pmatrix}.$$

According to standard GMM theory, an optimal choice of the inverse weight matrix V_N is a consistent estimate of the covariance matrix of the orthogonality conditions $E(Z_i'\Delta v_i \Delta v_i' Z_i)$, where Δv_i are the first-differenced residuals of each individual. Accordingly, a one-step GMM estimator uses $\hat{V} = \sum_{i-1}^{N} Z_i' DD' Z_i$, where D is the first-difference matrix operator. A two-step GMM estimator uses a robust choice $\widetilde{V} = \sum_{i=1}^{N} Z\Delta\hat{v}_i \Delta\hat{v}_i{\prime} Z_i$ where $\Delta\hat{v}_i$ are one-step GMM residuals.

Biorn (2000) proposes estimation of linear, static regression equations from panel data models with measurement errors in the regressors, showing that if the latent regressor is autocorrelated or nonstationary, several consistent OLS-IV and GMM estimators exist, provided some structure is imposed on the disturbances and measurement errors. He considers alternative GMM estimations that combine all essential orthogonality conditions. The procedures are very similar to the one described just above under non-autocorrelation in the disturbances. In particular, the required assumptions when allowing autocorrelation in the errors are very similar to those discussed in the previous section. For instance, when one allows for an MA(τ) structure in the measurement error, for instance, one must ensure that the variables in the IV matrix have a lead or lag of at least $\tau + 1$ periods to the regressor.

We briefly discuss the GMM estimators proposed by Biorn (2000). First, consider estimation using the equation in differences and instrumental variables in levels. After taking differences of the model, there are $(T-1)+(T+1)$ equations that can be stacked for individual i as

$$\begin{bmatrix} \Delta y_{i21} \\ \Delta y_{i32} \\ \vdots \\ \Delta y_i, T, T-1 \\ \Delta y_{i31} \\ \Delta y_{i42} \\ \vdots \\ \Delta y_i, T, T-2 \end{bmatrix} = \begin{bmatrix} \Delta x_{i21} \\ \Delta x_{i32} \\ \vdots \\ \Delta x_i T, T-1 \\ \Delta x_{i31} \\ \Delta x_{i42} \\ \vdots \\ \Delta x_i, T, T-2 \end{bmatrix} \beta + \begin{bmatrix} \Delta\varepsilon_{i21} \\ \Delta\varepsilon_{i32} \\ \vdots \\ \Delta\varepsilon_i, T, T-1 \\ \Delta\varepsilon_{i31} \\ \Delta\varepsilon_{i42} \\ \vdots \\ \Delta\varepsilon_i, T, T-2 \end{bmatrix},$$

or compactly

$$\Delta y_i = \Delta X_i \beta + \Delta \in_i.$$

The IV matrix is the $((2T - 3) \times KT(T - 2))$ diagonal matrix with the instruments in the diagonal defined by Z. Let

$$\Delta y = \left[(\Delta y_1)', \ldots, (\Delta y_N)'\right]', \Delta \in = \left[(\Delta \in_1)', \ldots, (\Delta \in_N)'\right]'$$

$$\Delta X = \left[(\Delta X_1)', \ldots, (\Delta X_N)'\right]', \quad Z = \left[Z_1', \ldots, Z_N'\right]'.$$

The GMM estimator that minimizes $[N^{-1}(\Delta\in)'Z](N^{-2}V)^{-1}[N^{-1}Z'(\Delta\in)]$ for $V = Z'Z$ can be written as

$$\hat{\beta}_{Dx} = \left[\left[\sum_i (\Delta X_i)' Z_i\right]\left[\sum_i Z_i' Z_i\right]^{-1}\left[\sum_i Z_i'(\Delta X_i)\right]\right]^{-1} \times \left[\left[\sum_i (\Delta X_i)' Z_i\right]\left[\sum_i Z_i' Z_i\right]^{-1}\left[\sum_i Z_i'(\Delta y_i)\right]\right].$$

If $\Delta\in$ has a non-scalar covariance matrix, a more efficient GMM estimator, $\widetilde{\beta}_{Dx}$, can be obtained setting $V = V_{Z(\Delta\in)} = E[Z'(\Delta\in)(\Delta\in)'Z]$ and estimating $\hat{V}_{z(\Delta\in)}$ by

$$\frac{\hat{V}_{Z(\Delta\in)}}{N} = \frac{1}{N}\sum_i Z'\left(\widehat{\Delta\in}\right)\left(\widehat{\Delta\in}\right)'Z,$$

where $\widehat{\Delta\in}_i = \Delta y_i - (\Delta X_i)\hat{\beta}_{Dx}$. This procedure assumes that (A1) and (A2) are satisfied. However, as Biorn (2000) argues, one can replace them by (B1) or (B2) and then ensure that the variables in the IV matrix have a lead or lag of at least $\tau + 1$ periods to the regressor, to "get clear of" the t period memory of the MA(τ) process. The procedure described below is also based on the same set of assumptions and can be extended similarly.[14]

The procedure for estimation using equation in levels and IVs in difference is similar. Consider the T stacked level equations for individual i:

$$\begin{bmatrix} y_{i1} \\ \vdots \\ y_{iT} \end{bmatrix} = \begin{bmatrix} c \\ \vdots \\ c \end{bmatrix} + \begin{bmatrix} x_{i1} \\ \vdots \\ x_{iT} \end{bmatrix}\beta + \begin{bmatrix} \in_{i1} \\ \vdots \\ \in_{iT} \end{bmatrix},$$

or more compactly,

$$y_i = e_T c + X_i\beta + \in,$$

where e_T denotes a $(T \times 1)$ vector of ones. Let the $(T \times T(T-2)K)$ diagonal matrix of instrument be denoted by ΔZ_i. This matrix has the instruments in difference in the main diagonal. In addition,

[14]See Propositions 1* and 2* in Biorn (2000) for a formal treatment of the conditions.

define:

$$y = \left[y_1', \ldots, y_N'\right]', \in = \left[\in_1', \ldots, \in_N'\right]'$$
$$X = \left[X_1', \ldots, X_N'\right]', \Delta Z = \left[(\Delta Z_1)', \ldots, (\Delta Z_N)'\right]'.$$

The GMM estimator that minimizes $[N^{-1}\in'(\Delta Z_i)']|(N^{-2}V_\Delta)^{-1}[N^{-1}(\Delta Z)'\in]$ for $V_\Delta = (\Delta Z)'(\Delta Z)$ is

$$\begin{aligned}\hat{\beta}_{Lx} &= \left[\left[\sum_i X_i'(\Delta Z_i)\right]\left[\sum_i (\Delta Z_i)'\Delta Z_i\right]^{-1}\left[\sum_i (\Delta Z_i)'X_i\right]\right]^{-1}\\ &\times\left[\left[\sum_i X_i{}'(\Delta Z_i)\right]\left[\sum_i (\Delta Z_i)'\Delta Z_i\right]^{-1}\left[\sum_i (\Delta Z_i)'y_i\right]\right].\end{aligned}$$

If $\in$ has a non-scalar covariance matrix, a more efficient GMM estimator, $\widetilde{\beta}_{Lx}$, can be obtained setting $V_\Delta = V_{(\Delta Z)\in} = E[(\Delta Z)\in\in'(\Delta Z)]$ and estimating $\hat{V}_{(\Delta Z)\in}$ by

$$\frac{\hat{V}_{(\Delta Z)\in}}{N} = \frac{1}{N}\sum_i (\Delta Z)'\hat{\in}\hat{\in}'(\Delta Z),$$

where $\hat{\in} = y_i - X_i\hat{\beta}_{Lx}$.

Finally, let us briefly contrast the OLS-IV and AB-GMM estimators. The advantages of GMM over IV are clear: if heteroscedasticity is present, the GMM estimator is more efficient than the IV estimator, while if heteroscedasticity is not present, the GMM estimator is no worse asymptotically than the IV. Implementing the GMM estimator, however, usually comes with a high price. The main problem, as Hayashi (2000, p. 215) points out, concerns the estimation of the optimal weighting matrix that is at the core of the GMM approach. This matrix is a function of fourth moments, and obtaining reasonable estimates of fourth moments requires very large sample sizes. Problems also arise when the number of moment conditions is high, that is, when there are "too many instruments." This latter problem affects squarely the implementation of the AB-GMM, since it relies on large numbers of lags (especially in long panels). The upshot is that the efficient GMM estimator can have poor small sample properties [see Baum et al. (2003) for a discussion]. These problems are well documented and remedies have been proposed by, among others, Altonji and Segal (1996) and Doran and Schmidt (2006).

57.3 Monte Carlo Analysis

We use Monte Carlo simulations to assess the finite sample performance of the EW and IV estimators discussed in Sect. 57.2. Monte Carlo simulations are an ideal experimental tool because they enable us to study those two estimators in a controlled setting, where we can assess and compare the importance of elements that are key to estimation performance. Our simulations use several distributions to generate observations. This is important because researchers will often find a variety of distributions in real-world applications and because one ultimately does not see the distribution of the mismeasurement term. Our Monte Carlos compare the EW, OLS-IV, and AB-GMM estimators presented in Sect. 57.2 in terms of bias and RMSE.[15] We also investigate the properties of the EW identification test, focusing on the empirical size and power of this test.

57.3.1 Monte Carlo Design

A critical feature of panel data models is the observation of multiple data points from the same individuals over time. It is natural to consider that repeat samples are particularly useful in that individual idiosyncrasies are likely to contain information that might influence the error structure of the data-generating process.

We consider a simple data-generating process to study the finite sample performance of the EW and OLS-IV estimators. The response variable y_{it} is generated according to the following model:

$$y_{it} = \gamma_i + \beta\chi_{it} + z_{it}'\alpha(1 + \rho w_{it})u_{it}, \tag{57.18}$$

where γ_i captures the individual-specific intercepts, β is a scalar coefficient associated with the mismeasured variable χ_{it}, $\alpha = (\alpha_1,\alpha_2,\alpha_3)'$ is 3×1 vector of coefficients associated with the 3×1 vector of perfectly measured variables $z_{it} = (z_{it1}, z_{it2}, z_{it3})$, u_{it} is the error in the model, and ρ modulates the amount of heteroscedasticity in the model. When $\rho = 0$, the innovations are homoscedastic. When $\rho > 0$, there is heteroscedasticity associated with the variable w_{it}, and this correlation is stronger as the coefficient gets larger. The model in Eq. 57.18 is flexible enough to allow us to consider two different variables as w_{it}: (1) the individual-specific intercept γ_i and (2) the well-measured regressor z_{it}.

We consider a standard additive measurement error

$$x_{it} = \chi_{it} + v_{it}, \tag{57.19}$$

where χ_{it} follows an AR(1) process:

[15]The mean squared error (MSE) of an estimator $\hat{\theta}$ incorporates a component measuring the variability of the estimator (precision) and another measuring its bias (accuracy). An estimator with good MSE properties has small combined variance and bias. The MSE of $\hat{\theta}$ can be defined as $\mathrm{Var}\left(\hat{\theta}\right) + \left[\mathrm{Bias}\left(\hat{\theta}\right)\right]^2$. The root mean squared error (RMSE) is simply the square root of the MSE. This is an easily interpretable statistic, since it has the same unit as the estimator $\hat{\theta}$. For an approximately unbiased estimator, the RMSE is just the square root of the variance, that is, the standard error.

$$(1 - \phi L)\chi_{it} = \in_{it}. \tag{57.20}$$

In all simulations, we set $\chi^{*}_{i,-50} = 0$ and generate χ_{it} for $t = -49, -48, \ldots, T$, such that we drop the first 50 observations. This ensures that the results are not unduly influenced by the initial values of the χ_{it} process.

Following Biorn (2000), we relax the assumption of i.i.d. measurement error. Our benchmark simulations will use the assumption of time-invariant autocorrelation [(C1) and (C2)]. In particular, we assume that $u_{it} = u_{1i} + u_{2it}$ and $v_{it} = v_{1i} + v_{2it}$. We draw all the innovations (u_{1i},u_{2it},v_{1i},v_{2it},) from a lognormal distribution; that is, we exponentiate two normal distributions and standardize the resulting variables to have unit variances and zero means (this follows the approach used by EW). In Sect. 57.3.6, we analyze the alternative case in which the innovations follow an MA structure.

The perfectly measured regressor is generated according to

$$z_{it} = \mu_i + \in_{it}. \tag{57.21}$$

And the fixed effects, μ_i and γ_i, are generated as

$$\begin{aligned} \mu_i &= e_{1i} \\ \gamma_i &= e_{2i} + \frac{1}{\sqrt{T}}\sum_{t=1}^{T} W_{it}, \end{aligned} \tag{57.22}$$

where W_{it} is the sum of the explanatory variables. Our method of generating μ_i and γ_i ensures that the usual random effects estimators are inconsistent because of the correlation that exists between the individual effects and the error term or the explanatory variables. The variables (e_{1i}, e_{2i}) are fixed as standard normal distributions.[16]

We employ four different schemes to generate the disturbances ($\in_{it}$, ε_{it}). Under Scheme 1, we generate them under a normal distribution, $N(0,\sigma_u^2)$. Under Scheme 2, we generate them from a lognormal distribution, $LN(0,\sigma_u^2)$. Under Scheme 3, we use a chi-square with 5 degrees of freedom, χ^2_5. Under Scheme 4, we generate the innovations from a $F_{m,n}$-distribution with $m = 10$ and $n = 40$. The latter three distributions are right-skewed so as to capture the key distributional assumptions behind the EW estimator. We use the normal (non-skewed) distribution as a benchmark.

Naturally, in practice, one cannot determine how skewed – if at all – is the distribution of the partially out latent variable. One of our goals is to check how this assumption affects the properties of the estimators we consider. Figure 57.1 provides a visual illustration of the distributions we employ. By inspection, at least, the three skewed distributions we study appear to be plausible candidates for the distribution governing mismeasurement, assuming EW's prior that measurement error must be markedly rightly skewed.

[16] Robustness checks show that the choice of a standard normal does not influence our results.

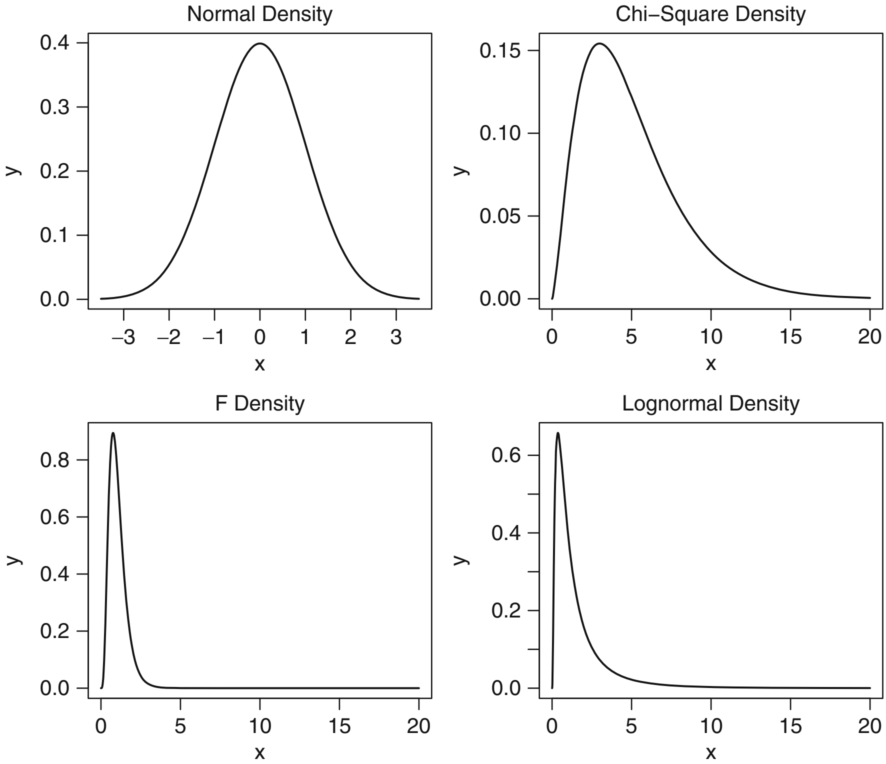


Fig. 57.1

As in Erickson and Whited (2002), our simulations allow for cross-sectional correlation among the variables in the model. We do so because this correlation may aggravate the consequences of mismeasurement of one regressor on the estimated slope coefficients of the well-measured regressors. Notably, this source of correlation is emphasized by EW in their argument that the inferences of Fazzari et al. (1988) are flawed in part due to the correlation between q and cash flows. To introduce this correlation in our application, for each period in the panel, we generate $(\chi_i, z_{i1}, z_{i2}, z_{i3})$ using the correspondent error distribution and then multiply the resulting vector by $[\mathrm{var}(\chi_i, z_{i1}, z_{i2}, z_{i3})]^{1/2}$ with diagonal elements equal to 1 and off-diagonal elements equal to 0.5.

In the simulations, we experiment with $T = 10$ and $N = 1{,}000$. We set the number of replications to 5,000 and consider the following values for the remaining parameters:

$$(\beta, \alpha_1, \alpha_2, \alpha_3) = (1, -1, 1, -1)$$
$$\phi = 0.6, \sigma_u^2 = \sigma_{e1}^2 = \sigma_{e2}^2 = 1,$$

where the set of slope coefficients β, α_i is set similarly to EW.

Notice that the parameter ϕ controls the amount of autocorrelation of the latent regressor. As explained above, this autocorrelation is an important requirement for the identification of the IV estimator. While we set $\phi = 0.6$ in the following experiments, we also conduct simulations in which we check the robustness of the results with respect to variations in ϕ between 0 and 1 (see Sect. 57.3.6).

57.3.2 The EW Identification Test

We study the EW identification test in a simple panel data setup. In the panel context, it is important to consider individual fixed effects. If the data contain fixed effects, according to Erickson and Whited (2000), a possible strategy is to transform the data first and then apply their high-order GMM estimator. Accordingly, throughout this section, our estimations consider data presented in two forms: "level" and "within." The first refers to data in their original format, without the use of any transformation; estimations in level form ignore the presence of fixed effects.[17] The second applies the within transformation to the data – eliminating fixed effects – before the model estimation.

We first compute the empirical size and power of the test. Note that the null hypothesis is that the model is incorrectly specified, such that $\beta = 0$ and/or $E[\eta_i^3] = 0$. The *empirical size* is defined as the number of rejections of the null hypothesis when the null is true – ideally, this should hover around 5 %. In our case, the empirical size is given when we draw the innovations $(\in_{it}, \varepsilon_{it})$ from a non-skewed distribution, which is the normal distribution since it generates $E[\eta_i^3] = 0$. The empirical power is the number of rejections when the null hypothesis is false – ideally, this should happen with very high probability. In the present case, the empirical power is given when we use skewed distributions: lognormal, chi-square, and F-distribution.

Our purpose is to investigate the validity of the skewness assumption once we are setting $\beta \neq 0$. Erickson and Whited (2002) also restrict every element of β to be nonzero. We conduct a Monte Carlo experiment to quantify the second part of this assumption. It is important to note that we can compute the $E[(\eta_i)^3]$ since, in our controlled experiment, we generate χ_i and therefore observe it.

Since the EW test is originally designed for cross-sectional data, the first difficulty the researcher faces when implementing a panel test is aggregation. Following EW, our test is computed for each year separately. We report the average of empirical rejections over the years.[18] To illustrate the size and power of the test for the panel data case, we set the time series dimension of the

[17]To our knowledge, all but one of the empirical applications of the EW model use the data in level form. In other words, firm-fixed effects are ignored outright in panel setting estimations of parameters influencing firm behavior.

[18]The results using the median are similar.

Table 57.1 The performance of the EW identification test

Distribution	Null is	Data form	Frequency of rejection
Normal	True	Level	0.05
		Within	0.05
Lognormal	False	Level	0.47
		Within	0.43
χ_3^2	False	Level	0.14
		Within	0.28
$F_{10,\,40}$	False	Level	0.17
		Within	0.28

This table shows the performance of the EW identification test for different distributional assumptions displayed in column 1. The tests are computed for the data in levels and after applying a within transformation. Column 4 shows the frequencies at which the null hypothesis that the model is not identified is rejected at the 5 % level of significance

panel to $T = 10$. Our tests are performed over 5,000 samples of cross-sectional size equal to 1,000. We use a simple homoscedastic model with $\rho = 0$, with the other model parameters given as above.

Table 57.1 reports the empirical size and power of the statistic proposed by EW for testing the null hypothesis $H_0 : E(\dot{y}_i^2 \dot{x}_i) = E(\dot{y}_i \dot{x}_i^2) = 0$. This hypothesis is equivalent to testing $H_0 : \beta = 0$ and/or $E(\eta_i^3) = 0$. Table 57.1 reports the frequencies at which the statistic of test is rejected at the 5 % level of significance for, respectively, the normal, lognormal, chi-square, and F-distributions of the data-generating process. Recall that when the null hypothesis is true, we have the size of the test, and when the null is false, we have the power of the test.

The results reported in Table 57.1 imply an average size of approximately 5 % for the test. In particular, the first two rows in the table show the results in the case of a normal distribution for the residuals (implying that we are operating under the null hypothesis). For both the level and within cases, the empirical sizes match the target significance level of 5 %.

When we move to the case of skewed distributions (lognormal, chi-square, and F), the null hypothesis is not satisfied by design, and the number of rejections delivers the empirical power of the test. In the case when the data is presented in levels and innovations are drawn from a lognormal distribution (see row 2), the test rejects about 47 % of the time the null hypothesis of no skewness. Using within data, the test rejects the null hypothesis 43 % of the time. Not only are these frequencies low, but comparing these results, one can see that the within transformation slightly reduces the power of the test.

The results associated with the identification test are more disappointing when we consider other skewed distributions. For example, for the F-distribution, we obtain only 17 % of rejections of the null hypothesis in the level case and only 28 % for the within case. Similarly, poor statistical properties for the model identification test are observed in the chi-square case.

57.3.3 Bias and Efficiency of the EW, OLS-IV, and AB-GMM Estimators

In this section we present simulation results that assess the finite sample performance of the estimators discussed in Sect. 57.2. The simulations compare the estimators in terms of bias and efficiency under several distributional assumptions. In the next subsection, we consider the cross-sectional setting, focusing on the properties of the EW estimator. Subsequently, we examine the panel case in detail, comparing the performance of the EW, OLS-IV, and AB-GMM estimators in terms of bias and efficiency.

57.3.3.1 The Cross-Sectional case

We generate data using a simple model as in Eqs. 57.18, and 57.19 with $T = 1$, such that there are no fixed effects, no autocorrelation($\phi = 0$), and no heteroscedasticity ($\rho = 0$). The other parameters are $(\beta, \alpha_1, \alpha_2, \alpha_3) = (1, -1, 1, -1)$. Table 57.2 shows the results for bias and RMSE for four different distributions: lognormal, chi-square, F-distribution, and standard normal. For each distribution we estimate the model using three different EW estimators: EW-GMM3, EW-GMM4, and EW-GMM5. These estimators are based on the respective third, fourth, and fifth moment conditions. By combining the estimation of 4 parameters, under 4 different distributions, for all 3 EW estimators – a total of 48 estimates – we aim at establishing robust conclusions about the bias and efficiency of the EW approach.

Panel A of Table 57.2 presents the results for bias and RMSE when we use the lognormal distribution to generate innovations ($\in_i$ ε_i) that produce χ_i and z_i. Under this particular scenario, point estimates are approximately unbiased, and the small RMSEs indicate that coefficients are relatively efficiently estimated.

Panels B and C of Table 57.2 present the results for the chi-square and F-distribution, respectively. The experiments show that coefficient estimates produced by the EW approach are generally very biased. For example, Panel B shows that the β coefficient returned for the EW-GMM4 and EW-GMM5 estimators is biased downwards by approximately 35 %. Panel C shows that for EW-GMM3, the β coefficient is biased upwards about 35 %. Paradoxically, for EW-GMM4 and EW-GMM5, the coefficients are biased downwards by approximately 25 %. The coefficients returned for the perfectly measured regressors are also noticeably biased. And they, too, switch bias signs in several cases. Panels B and C show that the EW RMSEs are very high. Notably, the RMSE for EW-GMM4 under the chi-square distribution is 12.23, and under F-distribution, it is 90.91. These RMSE results highlight the lack of efficiency of the EW estimator. Finally, Panel D presents the results for the normal distribution case, which has zero skewness. In this case, the EW estimates are severely biased and the RMSEs are extremely high. The estimated coefficient for the mismeasured variable using EW-GMM3 has a bias of 1.91 (about three times larger than its true value) and an RMSE of 2305.

These results reveal that the EW estimators only have acceptable performance in the case of very strong skewness (lognormal distribution). They relate to the last section in highlighting the poor identification of the EW framework, even in the

Table 57.2 The EW estimator: cross-sectional data

		β	α_1	α_2	α_3
Panel A. Lognormal distribution					
EW-GMM3	Bias	0.0203	−0.0054	−0.0050	−0.0056
	RMSE	0.1746	0.0705	0.0704	0.0706
EW-GMM4	Bias	0.0130	−0.0034	−0.0033	−0.0040
	RMSE	0.2975	0.1056	0.1047	0.1083
EW-GMM5	Bias	0.0048	−0.0013	−0.0013	−0.0019
	RMSE	0.0968	0.0572	0.0571	0.0571
Panel B. Chi-square distribution					
EW-GMM3	Bias	−0.0101	0.0092	0.0101	−0.0060
	RMSE	61.9083	16.9275	16.1725	14.7948
EW-GMM4	Bias	−0.3498	0.0938	0.0884	0.0831
	RMSE	12.2386	3.2536	2.9732	3.1077
EW-GMM5	Bias	−0.3469	0.0854	0.0929	0.0767
	RMSE	7.2121	1.8329	1.8720	1.6577
Panel C. F-distribution					
EW-GMM3	Bias	0.3663	−0.1058	−0.0938	−0.0868
	RMSE	190.9102	53.5677	52.4094	43.3217
EW-GMM4	Bias	−0.2426	0.0580	0.0649	0.0616
	RMSE	90.9125	24.9612	24.6827	21.1106
EW-GMM5	Bias	−0.2476	0.0709	0.0643	0.0632
	RMSE	210.4784	53.5152	55.8090	52.4596
Panel D. Normal distribution					
EW-GMM3	Bias	1.9179	−0.6397	−0.5073	−0.3512
	RMSE	2305.0309	596.1859	608.2098	542.2125
EW-GMM4	Bias	−1.0743	0.3012	0.2543	0.2640
	RMSE	425.5931	111.8306	116.2705	101.4492
EW-GMM5	Bias	3.1066	−1.0649	−0.9050	−0.5483
	RMSE	239.0734	60.3093	65.5883	58.3686

This table shows the bias and the RMSE associated with the estimation of the model in Eqs. 57.17, 57.18, 57.19, 57.20, and 57.21 using the EW estimator in simulated cross-sectional data. β is the coefficient on the mismeasured regressor, and α_1 to α_3 are the coefficients on the perfectly measured regressors. The table shows the results associated with GMM3, GMM4, and GMM5 for all the alternative distributions. These estimators are based on the respective third, fourth, and fifth moment conditions

most basic cross-sectional setup. Crucially, for the other skewed distributions we study, the EW estimator is significantly biased for both the mismeasured and the well-measured variables. In addition, the RMSEs are quite high, indicating low efficiency.

57.3.3.2 The Panel Case

We argue that a major drawback of the EW estimator is its limited ability to handle individual heterogeneity – fixed effects and error heteroscedasticity – in panel data. This section compares the impact of individual heterogeneity on the EW, OLS-IV, and AB-GMM estimators in a panel setting. In the first round of experiments, we assume error homoscedasticity by setting the parameter ρ in Eq. 57.18 equal to zero. We shall later allow for changes in this parameter.

Although the EW estimations are performed on a period-by-period basis, one generally wants a single coefficient for each of the variables in an empirical model. To combine the various (time-specific) estimates, EW suggest the minimum distance estimator (MDE) described below. Accordingly, the results presented in this section are for the MDE that combines the estimates obtained from each of the ten time periods considered. For example, EW-GMM3 is the MDE that combines the ten different cross-sectional EW-GMM3 estimates in our panel.

The OLS-IV is computed after differencing the model and using the second lag of the observed mismeasured variable x, x_{t-2}, as an instrument for Δx_t. The AB-GMM estimates (Arellano and Bond 1991) use all the orthogonality conditions, with all available lags of x's as instrumental variables. We also concatenate the well-measured variables z's in the instruments' matrix. The AB-GMM estimator is also computed after differencing Eq. 57.18. To highlight the gains of these various estimators vis-à-vis the standard (biased) OLS estimator, we also report the results of simulations for OLS models using equation in first difference without instruments.

We first estimate the model using data in level form. While the true model contains fixed effects (and thus it is appropriate to use the within transformation), it is interesting to see what happens in this case since most applications of the EW estimator use data in level form, and as shown previously, the EW identification test performs slightly better using data in this form.

The results are presented in Table 57.3. The table makes it clear that the EW method delivers remarkably biased results when ignoring the presence of fixed effects. Panel A of Table 57.3 reports the results for the model estimated with the data under strong skewness (lognormal). In this case, the coefficients for the mismeasured regressor are very biased, with biases well in excess of 100 % of the true coefficient for the EW-GMM3, EW-GMM4, and EW-GMM5 estimators. The biases for the well-measured regressors are also very strong, all exceeding 200 % of the estimates' true value. Panels B and C report results for models under chi-square and F-distributions, respectively. The EW method continues to deliver very biased results for all of the estimates considered. For example, the EW-GMM3 estimates that are returned for the mismeasured regressors are biased downwardly by about 100 % of their true values – those regressors are deemed irrelevant when they are not. Estimates for the well-measured regressors are positively biased by approximately 200 % – they are inflated by a factor of 3. The RMSEs reported in Panels A, B, and C show that the EW methodology produces very inefficient estimates even when one assumes pronounced skewness in the data. Finally, Panel D reports the results for the normal distribution. For the non-skewed data case, the EW framework can produce estimates for the mismeasured regressor that are downwardly biased by about 90 % of their true parameter values for all models. At the same time, that estimator induces an upward bias of larger than 200 % for the well-measured regressors.

Table 57.3 The EW estimator: panel data in levels

		β	α_1	α_2	α_3
Panel A. Lognormal distribution					
EW-GMM3	Bias	−1.6450	2.5148	2.5247	2.5172
	RMSE	1.9144	2.5606	2.5711	2.5640
EW-GMM4	Bias	−1.5329	2.5845	2.5920	2.5826
	RMSE	1.9726	2.6353	2.6443	2.6354
EW-GMM5	Bias	−1.3274	2.5468	2.5568	2.5490
	RMSE	1.6139	2.5944	2.6062	2.5994
Panel B. Chi-square distribution					
EW-GMM3	Bias	−1.0051	2.2796	2.2753	2.2778
	RMSE	1.1609	2.2887	2.2841	2.2866
EW-GMM4	Bias	−0.9836	2.2754	2.2714	2.2736
	RMSE	1.0540	2.2817	2.2776	2.2797
EW-GMM5	Bias	−0.9560	2.2661	2.2613	2.2653
	RMSE	1.0536	2.2728	2.2679	2.2719
Panel C. F-distribution					
EW-GMM3	Bias	−0.9926	2.2794	2.2808	2.2777
	RMSE	1.1610	2.2890	2.2904	2.2870
EW-GMM4	Bias	−0.9633	2.2735	2.2768	2.2720
	RMSE	1.0365	2.2801	2.2836	2.2785
EW-GMM5	Bias	−0.9184	2.2670	2.2687	2.2654
	RMSE	2.0598	2.2742	2.2761	2.2725
Panel D. Normal distribution					
EW-GMM3	Bias	−0.8144	2.2292	2.228	2.2262
	RMSE	0.9779	2.2363	2.2354	2.2332
EW-GMM4	Bias	−0.9078	2.2392	2.2363	2.2351
	RMSE	0.9863	2.2442	2.2413	2.2400
EW-GMM5	Bias	−0.8773	2.2262	2.2225	2.2217
	RMSE	0.9846	2.2316	2.2279	2.2269

This table shows the bias and the RMSE associated with the estimation of the model in Eqs. 57.17, 57.18, 57.19, 57.20, and 57.21 using the EW estimator in simulated panel data. The table reports results from data in levels (i.e., without applying the within transformation). β is the coefficient on the mismeasured regressor, and α_1, α_2, α_3 are the coefficients on the perfectly measured regressors. The table shows the results for the EW estimator associated with EW-GMM3, EW-GMM4, and EW-GMM5 for all the alternative distributions. These estimators are based on the respective third, fourth, and fifth moment conditions

Table 57.4 reports results for the case in which we apply the within transformation to the data. Here, we introduce the OLS, OLS-IV, and AB-GMM estimators. We first present the results associated with the set up that is most favorable for the EW estimations, which is the lognormal case in Panel A. The EW estimates for the lognormal case are relatively unbiased for the well-measured regressors (between 4 % and 7 % deviation from true parameter values). The same applies for the mismeasured regressors. Regarding the OLS-IV, Panel A shows that coefficient estimates are unbiased in all models considered. AB-GMM estimates are also approximately

Table 57.4 OLS, OLS-IV, AB-GMM, and EW estimators: panel data after within transformation

		β	α_1	α_2	α_3
Panel A. Lognormal distribution					
OLS	Bias	−0.7126	0.1553	0.1558	0.1556
	RMSE	0.7131	0.1565	0.1570	0.1568
OLS-IV	Bias	0.0065	−0.0019	−0.0014	−0.0015
	RMSE	0.1179	0.0358	0.0357	0.0355
AB-GMM	Bias	−0.0248	0.0080	0.0085	0.0081
	RMSE	0.0983	0.0344	0.0344	0.0340
EW-GMM3	Bias	−0.0459	0.0185	0.0184	0.0183
	RMSE	0.0901	0.0336	0.0335	0.0335
EW-GMM4	Bias	−0.0553	0.0182	0.0182	0.0183
	RMSE	0.1405	0.0320	0.0321	0.0319
EW-GMM5	Bias	−0.0749	0.0161	0.0161	0.0161
	RMSE	0.1823	0.0303	0.0297	0.0297
Panel B. Chi-square distribution					
OLS	Bias	−0.7126	0.1555	0.1553	0.1556
	RMSE	0.7132	0.1565	0.1563	0.1567
OLS-IV	Bias	0.0064	−0.0011	−0.0017	−0.001
	RMSE	0.1149	0.0348	0.0348	0.0348
AB-GMM	Bias	−0.0231	0.0083	0.0077	0.0081
	RMSE	0.0976	0.0339	0.0338	0.0342
EW-GMM3	Bias	−0.3811	0.0982	0.0987	0.0982
	RMSE	0.4421	0.1133	0.1136	0.1133
EW-GMM4	Bias	−0.3887	0.0788	0.0786	0.0783
	RMSE	0.4834	0.0927	0.0923	0.0919
EW-GMM5	Bias	−0.4126	0.0799	0.0795	0.0798
	RMSE	0.5093	0.0926	0.0921	0.0923
Panel C. F-distribution					
OLS	Bias	−0.7123	0.1554	0.1549	0.1555
	RMSE	0.7127	0.1565	0.1559	0.1566
OLS-IV	Bias	0.0066	−0.0013	−0.0023	−0.001
	RMSE	0.1212	0.0359	0.0362	0.0361
AB-GMM	Bias	−0.0232	0.0079	0.0072	0.0085
	RMSE	0.0984	0.0343	0.0342	0.0344
EW-GMM3	Bias	−0.3537	0.0928	0.0916	0.0917
	RMSE	0.4239	0.1094	0.1086	0.1095
EW-GMM3	Bias	−0.3906	0.0802	0.0790	0.0791
	RMSE	0.4891	0.0939	0.0930	0.0932
EW-GMM3	Bias	−0.4188	0.0818	0.0808	0.0813
	RMSE	0.5098	0.0939	0.0932	0.0935

(*continued*)

Table 57.4 (continued)

		β	α_1	α_2	α_3
Panel D. Normal distribution					
OLS	Bias	−0.7119	0.1553	0.1554	0.1551
	RMSE	0.7122	0.1563	0.1564	0.1562
OLS-IV	Bias	0.0060	−0.0011	−0.0012	−0.0014
	RMSE	0.1181	0.0353	0.0355	0.0358
AB-GMM	Bias	−0.0252	0.0086	0.0085	0.0084
	RMSE	0.0983	0.0344	0.0339	0.0343
EW-GMM3	Bias	−0.7370	0.1903	0.1904	0.1895
	RMSE	0.7798	0.2020	0.2024	0.2017
EW-GMM4	Bias	−0.8638	0.2141	0.2137	0.2137
	RMSE	0.8847	0.2184	0.218	0.2182
EW-GMM5	Bias	−0.8161	0.1959	0.1955	0.1955
	RMSE	0.8506	0.2021	0.2018	0.2017

This table shows the bias and the RMSE associated with the estimation of the model in Eqs. 57.17, 57.18, 57.19, 57.20, 57.21 using the OLS, OLS-IV, AB-GMM, and EW estimators in simulated panel data. The table reports results from the estimators on the data after applying the within transformation. β is the coefficient on the mismeasured regressor, and α_1, α_2, α_3 are the coefficients on the perfectly measured regressors. The table shows the results for the EW estimator associated with EW-GMM3, EW-GMM4, and EW-GMM5 for all the alternative distributions. These estimators are based on the respective third, fourth, and fifth moment conditions

unbiased, while standard OLS estimates are very biased. In terms of efficiency, the RMSEs of the EW-GMM3 are somewhat smaller than those of the OLS-IV and AB-GMM for the well-measured and mismeasured regressors. However, for the mismeasured regressor, both OLS-IV and AB-GMM have smaller RMSEs than EW-GMM4 and EW-GMM5.

Panel B of Table 57.4 presents the results for the chi-square distribution. One can see that the EW yields markedly biased estimates in this case. The bias in the mismeasured regressor is approximately 38 % (downwards), and the coefficients for the well-measured variable are also biased (upwards). In contrast, the OLS-IV and AB-GMM estimates for both well-measured and mismeasured regressors are approximately unbiased. In terms of efficiency, as expected, the AB-GMM presents slightly smaller RMSEs than the OLS-IV estimator. These IV estimators' RMSEs are much smaller than those associated with the EW estimators.

Panels C and D of Table 57.4 show the results for the F and standard normal distributions, respectively. The results for the F-distribution in Panel C are essentially similar to those in Panel B: the instrumental variable estimators are approximately unbiased while the EW estimators are very biased. Finally, Panel D shows that deviations from a strongly skewed distribution are very costly in terms of bias for the EW estimator, since the bias for the mismeasured regressor is larger than 70 %, while for the well measured, it is around 20 %. A comparison of RMSEs shows that the IV estimators are more efficient in both the F and normal cases. In all, our simulations show that standard IV methods almost universally dominate the EW estimator in terms of bias and efficiency.

We reiterate that the bias and RMSE of the IV estimators in Table 57.4 are all relatively invariant to the distributional assumptions, while the EW estimators are all very sensitive to those assumptions. In short, this happens because the EW relies on the high-order moment conditions as opposed to the OLS and IV estimators.

57.3.4 Heteroscedasticity

One way in which individual heterogeneity may manifest itself in the data is via error heteroscedasticity. Up to this point, we have disregarded the case in which the data has a heteroscedastic error structure. However, most empirical applications in corporate finance entail the use of data for which heteroscedasticity might be relevant. It is important that we examine how the EW and the IV estimators are affected by heteroscedasticity.[19]

The presence of heteroscedasticity introduces heterogeneity in the model and consequently in the distribution of the partialled out dependent variable. This compromises identification in the EW framework. Since the EW estimator is based on equations giving the moments of $(y_i - z_i\mu_y)$ and $(y_i - z_i\mu_\chi)$ as functions of β and moments of $(u_i, \varepsilon_i, \eta_i)$, the heteroscedasticity associated with the fixed effects (α_i) or with the perfectly measured regressor (z_{it}) distorts the required moment conditions associated with $(y_i - z_i\mu_y)$, yielding biased estimates. These inaccurate estimates enter the minimum distance estimator equation and consequently produce incorrect weights for each estimate along the time dimension. As our simulations of this section demonstrate, this leads to biased MDE estimates, where the bias is a function of the amount of heteroscedasticity.

We examine the biases imputed by heteroscedasticity by way of graphical analysis. The graphs we present below are useful in that they synthesize the outputs of numerous tables and provide a fuller visualization of the contrasts we draw between the EW and OLS-IV estimators. The graphs depict the sensitivity of those two estimators with respect to heteroscedasticity as we perturb the coefficient p in Eq. 57.18.

In our simulations, we alternatively set $w_{it} = \gamma_i$ or $w_{it} = z_{it}$. In the first case, heteroscedasticity is associated with the individual effects. In the second, heteroscedasticity is associated with the well-measured regressor. Each of our figures describes the biases associated with the mismeasured and the well-measured regressors for each of the OLS-IV, EW-GMM3, EW-GMM4, and EW-GMM5 estimators.[20] In order to narrow our discussion, we only present results for the highly skewed distribution case (lognormal distribution) and for data that is treated for fixed effects using the within transformation. As Sect. 57.3.3 shows, this is the only case in which the EW estimator returns relatively unbiased estimators for the parameters of interest. In all the other cases (data in levels and for data generated by

[19]We focus on the OLS-IV estimator hereinafter for the purpose of comparison with the EW estimator.

[20]Since estimation biases have the same features across all well-measured regressors of a model, we restrict attention to the first well-measured regressor of each of the estimated models.

chi-square, F, and normal distributions), the estimates are strongly biased even under the assumption of homoscedasticity.[21]

Figure 57.2 presents the simulation results under the assumption that $w_{it} = \gamma_i$ as we vary the amount of heteroscedasticity by changing the parameter ρ,[22] the results for the mismeasured coefficients show that biases in the EW estimators are generally small for ρ equal to zero (this is the result reported in Sect. 57.3.3). However, as this coefficient increases, the bias quickly becomes large. For example, for $\rho = 0.45$, the biases in the coefficient of the mismeasured variable are, respectively, -11 %, -20 %, and -43 %, for the EW-GMM3, EW-GMM4, and EW-GMM5 estimators. Notably, those biases, which are initially negative, turn positive for moderate values of ρ. As heteroscedasticity increases, some of the biases diverge to positive infinite. The variance of the biases of the EW estimators is also large. The results regarding the well-measured variables using EW estimators are analogous to those for the mismeasured one. Biases are substantial even for small amounts of heteroscedasticity, they switch signs for some level of heteroscedasticity, and their variances are large. In sharp contrast, the same simulation exercises show that the OLS-IV estimates are approximately unbiased even under heteroscedasticity. While the EW estimator may potentially allow for some forms of heteroscedasticity, it is clear that it is not well equipped to deal with this problem in more general settings.

57.3.5 Identification of the EW Estimator in Panel Data

Our Monte Carlo experiments show that the EW estimator has a poor handle of individual fixed effects and that biases arise for deviations from the assumption of strict lognormality. Biases in the EW framework are further magnified if one allows for heteroscedasticity in the data (even under lognormality). The biases arising from the EW framework are hard to measure and sign, ultimately implying that it can be very difficult to replicate the results one obtains under that framework.

To better understand these results, we now discuss in more mathematical details the identification of the EW estimator for the panel data case for both the model in level and after the within transformation. Extending the EW estimator to panel data seems to be a nontrivial task. EW have proposed to break the problem for each time series, estimate a cross-section model for each t, and after that combine the estimates using a minimum distance estimator. In what follows we show that this procedure might affect the identification condition.

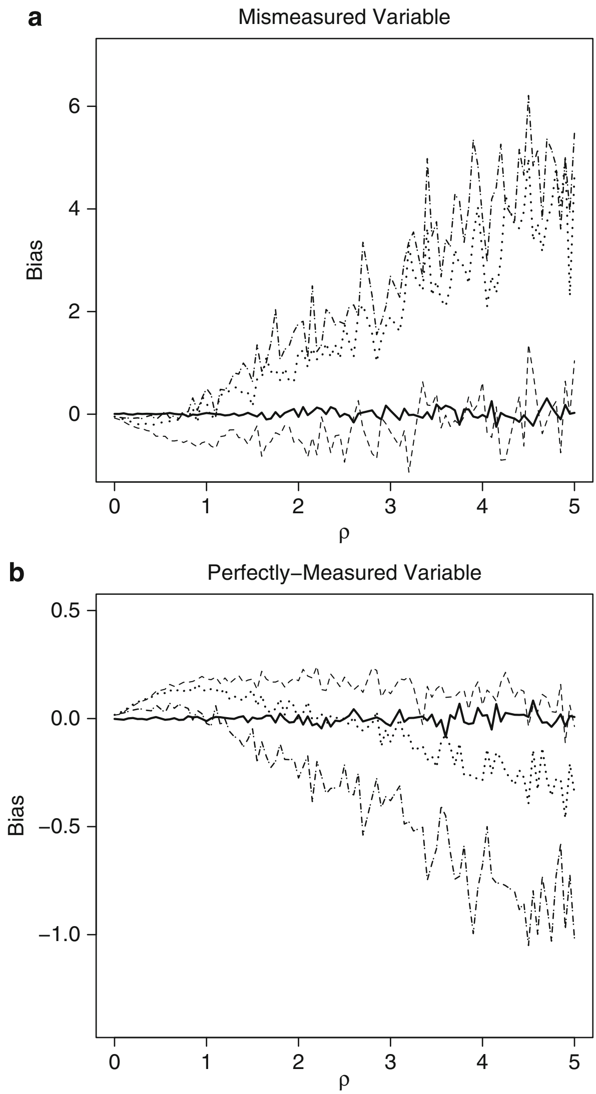
Consider the following model:

$$y_{it} = \alpha_i + \beta \chi_{it} + u_{it}, \quad i = 1, \ldots, N; t = 1, \ldots, T, \tag{57.23}$$

[21]Our simulation results (available upon request) suggest that introducing heteroscedasticity makes the performance of the EW estimator even worse in these cases.

[22]The results for $w_{it} = z_{it}$ are quite similar to those we get from setting $w_{it} = \gamma_i$. We report only one set of graphs to save space.

Fig. 57.2



where u_{it} is independent and identically distributed, with mean zero and variance σ_u^2. Assume that the $\mathrm{Var}(x_{it}) = \sigma_\chi^2$. The independent variable and unobserved effects are exogenous, that is, $\mathrm{Cov}(\chi_{it}, u_{is}) = \mathrm{Cov}(\alpha_i, u_{it}) = 0$ for any t and s. However, $\mathrm{Cov}(\alpha_i, \chi_{it}) \neq 0$. Now, assume that we do not observe the true variable χ_{it}, but rather a mismeasured variable, that is, you observe the following variable with an error:

$$x_{it} = \chi_{it} + e_{it}, \quad i = 1, \ldots, N \ldots, N; t = 1, \ldots T, \tag{57.24}$$

where $\mathrm{Cov}(x_{it}, e_{is}) = \mathrm{Cov}(\alpha_i, e_{is}) = \mathrm{Cov}(u_{it}, e_{is}) = 0$, and $\mathrm{Var}(e_{it}) = \sigma_e^2$, $\mathrm{Cov}(e_{it}, e_{it-1}) = \gamma\sigma_e^2$.

In addition, assume here that there is no variable z_{it} ($\alpha = 0$ in Eq. 57.18) to simplify the argument.

57.3.5.1 Model in Level

As mentioned before, EW propose to fix a particular time series and estimate the model using the cross-section data. Without loss for generality, fix $T = 1$. Thus, Eqs. 57.23 and 57.24 become

$$y_{i1} = \alpha_i + \beta\chi_{i1} + u_{i1}, \quad i = 1, \ldots, N, \tag{57.25}$$

and

$$x_{i1} = \chi_{i1} + e_{i1}, \quad i = 1, \ldots, N. \tag{57.26}$$

However, the unobserved individual-specific intercepts, α_i, are still present in Eq. 57.25 and in addition $\mathrm{Cov}(\alpha_i, \chi_{i1}) \neq 0$. Therefore, one can see that it is impossible to estimate β consistently since α_i's are unobserved. This argument is easily extended for every $t = 1, \ldots, T$. Thus, the estimator for each fixed t is inconsistent, and consequently the minimum distance estimator is inconsistent by construction. Therefore, we conclude that the EW minimum distance estimator produces inconsistent estimates for panel data model with fixed effects.

57.3.5.2 Model After Within Transformation

Given the inconsistency of the model in levels presented in the last section, one strategy is to previously transform the data to eliminate the fixed effects. One suggestion is to use the within transformation in the data before estimation.

In order to analyze the model after the transformation, let's assume that $T = 2$ for simplification. Using the within transformation in Eqs. 57.23 and 57.24, we obtain

$$y_{it} - \overline{y}_{it} = \beta(\chi_{it} - \overline{\chi}_i) + (u_{it} - \overline{u}_i),$$

and

$$x_{it} - \overline{x}_i = (\chi_{it} - \overline{\chi}_i) + (e_{it} - \overline{e}_i),$$

where $\bar{y}_i = \frac{1}{2}(y_{i1} + y_{i2})$, $\bar{\chi}_i = \frac{1}{2}(\chi_{i1} + \chi_{i2})$, and so on.

Now, again EW propose to use a particular time series and estimate the model using the cross-section data. Let's use $t = 1$ for ease of exposition. The model can be written as

$$y_{il} - \overline{y}_i = \beta(\chi_{i1} - \overline{\chi}_i) + (u_{i1} - \overline{u}_i)$$

and

$$x_{i1} - \overline{x}_i = (\chi_{i1} - \overline{\chi}_i) + (e_{it} - \overline{e}_i).$$

Now substituting the definition of the deviations and rearranging, we have

$$y_{i1} - \frac{1}{2}(y_{i1} + y_{i2}) = \beta\left(\chi_{i1} - \frac{1}{2}(\chi_{i1} + \chi_{i2})\right) + \left(u_{i1} - \frac{1}{2}(u_{i1} + u_{i2})\right),$$
$$y_{il} + y_{i2} = \beta(\chi_{i1} + \chi_{i2}) + (u_{i1} + u_{i2}),$$

and

$$x_{i1} - \frac{1}{2}(x_{i1} + x_{i2}) = \left(\chi_{i1} - \frac{1}{2}(\chi_{i1} + \chi_{i2})\right) + \left(e_{i1} - \frac{1}{2}(e_{i1} + e_{i2})\right),$$
$$x_{i1} + x_{i2} = (\chi_{i1} + \chi_{i2}) + (e_{i1} + e_{i2}).$$

Finally, our model can be described as

$$y_{i1} + y_{i2} = \beta(\chi_{i1} + \chi_{i2}) + (u_{i1} + u_{i2}),$$

and

$$x_{i1} + x_{i2} = (\chi_{i1} + \chi_{i2}) + (e_{i1} + e_{i2}).$$

Let's now define $Y_i = y_{i1} + y_{i2}$, $X_i = x_{i1} + x_{i2}$, $U_i = u_{i1} + u_{i2}$, $v_i = \chi_{i1} + \chi_{i2}$, and $E_i = e_{i1} + e_{i2}$. So, the model could be rewritten as

$$Y_i = \beta v_i + U_i$$

and

$$X_i = v_i + E_i.$$

Notice that the requirements for identification now are on the high-order moments of (v, U, E). However, note that $v_i = \chi_{i1} + \chi_{i2}$, which is a sum of two random variables. As it is well known from the econometrics literature, convolution of random variables is in general a nontrivial object.

One example of why the identification condition may worsen considerably is the following. Consider a model where χ_{i1} and χ_{i2} are independent chi-square distributions with 2 degrees of freedom. The skewness of the chi-square with k degrees of freedom is $\sqrt{8/k}$. Note that the sum of two independent chi-squares with k degrees of freedom is a chi-square with $2k$ degrees of freedom. Therefore, the skewness of the $v_i = \chi_{i1} + \chi_{i2}$ drops from two for the model using one distribution to 1.41 for the model using the summation of both χ_{i1} and χ_{i2}.

From this simple analysis one could conclude that the identification conditions required for EW estimator can deteriorate considerably when using the within transformation to eliminate the fixed effects in panel data. Thus, the required conditions to achieve unbiased estimates with EW are very strong.

57.3.6 Revisiting the OLS-IV Assumptions

Our Monte Carlo simulations show that the OLS-IV estimator is consistent even when one allows for autocorrelation in the measurement-error structure. We have assumed, however, some structure on the processes governing innovations. In this section, we examine the sensitivity of the OLS-IV results with respect to our assumptions about measurement-error correlation and the amount of autocorrelation in the latent regressor. These assumptions can affect the quality of the instruments and therefore should be examined in some detail.

We first examine conditions regarding the correlation of the measurement errors and disturbances. The assumption of time-invariant autocorrelation for measurement errors and disturbances implies that past shocks to measurement errors do not affect the current level of the measurement error. One way to relax this assumption is to allow for the measurement-error process to have a moving average structure. This structure satisfies Biorn's assumptions (B1) and (B2). In this case, Proposition 1 in Biorn (2000) shows that for an MA(τ), the instruments should be of order of at most $t - \tau - 2$. Intuitively, the set of instruments must be "older" than the memory of the measurement-error process. For example, if the measurement error is MA(1), then one must use third- and longer-lagged instruments to identify the model.

To analyze this case, we conduct Monte Carlo simulations in which we replace the time-invariant assumption for innovation u_{it} and v_{it} with an MA(1) structure for the measurement-error process. The degree of correlation in the MA process is set to $\theta = 0.4$. Thus, the innovation in Eqs. 57.18 and 57.19 has the following structure:

$$u_{it} = u_{1it} - \theta u_{1it-1} \;\text{ and }\; v_{it} = v_{1it} = v_{1it} - \theta v_{1it-1},$$

with $|\theta| \leq 1$, and u_{1it} and v_{1it} are i.i.d. lognormal distributions. The other parameters in the simulation remain the same.

The results are presented in Table 57.5. Using MA(1) in the innovations and the third lag of the latent regressor as an instrument (either on its own or in combination with the fourth lag), the bias of the OLS estimator is very small (approximately 2–3 %). The bias increases somewhat when we use only the fourth lag. While the fourth is an admissible instrument in this case, using longer lags decreases the implied autocorrelation in the latent regressor [which follows an

AR(1) process by Eq. 57.20]. This effect decreases somewhat the quality of the instruments. Notice also that when we do not eliminate short lags from the instrument set, the identification fails. For example, the bias is 60 % when we use the second lag as an instrument. These results thus underscore the importance of using long enough lags in this MA case. Table 57.5 also reports results based on an MA (2) structure. The results are qualitatively identical to those shown in the MA (1) case. Once again, the important condition for identification is to use long enough lags (no less than four lags in this case).[23]

The second condition underlying the use of the OLS-IV is that the latent regressor is not time invariant. Accordingly, the degree of autocorrelation in the process for the latent regressor is an important element of the identification strategy. We assess the sensitivity of the OLS-IV results to this condition by varying the degree of autocorrelation through the autoregressive coefficient in the AR(1) process for the latent regressor. In these simulations, we use a time-invariant autocorrelation condition for the measurement error, but the results are very similar for the MA case.

Figure 57.3 shows the results for the bias in the coefficients of interest for the well-measured and mismeasured variables, using the second lag of the mismeasured variable as an instrument. The results show that the OLS-IV estimator performs well for a large range of the autoregressive coefficient. However, as expected, when the ϕ coefficient is very close to zero or one, we have evidence of a weak instrument problem. For example, when $\phi = 1$, then $\Delta\chi_{it}$ is uncorrelated with any variable dated at time $t-2$ or earlier. These simulations show that, provided that one uses adequately lagged instruments, the exact amount of autocorrelation in the latent variable is not a critical aspect of the estimation.

The simulations of this section show how the performance of the OLS-IV estimator is affected by changes in assumptions concerning measurement errors and latent regressors. In practical applications, it is important to verify whether the results obtained with OLS-IV estimators are robust to the elimination of short lags from the instrumental set. This robustness check is particularly important given that the researcher will be unable to pin down the process followed by the measurement error. Our empirical application below incorporates this suggestion. In addition, identification relies on some degree of autocorrelation in the process for the latent regressor. While this condition cannot be directly verified, we can perform standard tests of instrument adequacy that rely on "first-stage" test statistics calculated from the processes for the observable variables in the model.

Another important assumption in the OLS is non-autocorrelation in both u_{it} and v_{it}. For example, these innovations cannot follow an autoregressive process.

[23]We note that if the instrument set uses suitably long lags, then the OLS-IV results are robust to variations in the degree of correlation in the MA process. In unreported simulations under MA(1), we show that the OLS bias is nearly invariant to the parameter θ.

Table 57.5 Moving average structures for the measurement-error process

Instrument	MA(1)	MA(2)
X_{it-2}	−0.593	−0.368
	(0.60)	(0.38)
X_{it-3}	0.028	−0.707
	(0.30)	(0.71)
X_{it-3}, X_{it-4}	0.025	0.077
	(0.63)	(1.01)
$X_{it-3}, X_{it-4}, X_{it-5}$	−0.011	−0.759
	(0.30)	(0.76)
X_{it-4}	−0.107	−0.144
	(1.62)	(2.01)
X_{it-4}, X_{it-5}	−0.113	−0.140
	(0.58)	(0.59)
X_{it-5}	−0.076	−0.758
	(0.31)	(0.76)

This table shows the bias in the well-measured coefficient for OLS-IV using moving average structure for the measurement-error process. Numbers in parentheses are the RMSE

When this is the case, the IV strategy of using lags of mismeasured variable as valid instruments is invalid (see Biorn 2000).

57.3.7 Distributional Properties of the EW and OLS-IV Estimators

A natural question is whether our simulation results are rooted in the lack of accuracy of the asymptotic approximation of the EW method. Inference in models with mismeasured regressors is based on asymptotic approximations; hence inference based on estimators with poor approximations might lead to wrong inference procedures. For instance, we might select wrong critical values for a test under poor asymptotic approximations and make inaccurate statements under such circumstances. In this section, we use the panel data simulation procedure of Sect. 3.3.2 to study and compare the accuracy of the asymptotic approximation of the EW and IV methods. To save space, we restrict our attention to the mismeasured regressor coefficient for the EW-GMM5 and OLS-IV cases. We present results where we draw the data from the lognormal, chi-square, and F-distributions. The EW-GMM5 estimator is computed after the within transformation and the OLS-IV uses second lags as instruments.

One should expect both the IV and EW estimators to have asymptotically normal representations, such that when we normalize the estimator by subtracting the true parameter and divide by the standard deviation, this quantity behaves asymptotically as a normal distribution. Accordingly, we compute the empirical density and the distribution functions of the normalized sample estimators and their normal approximations. These functions are plotted in Fig. 57.4. The true normal density

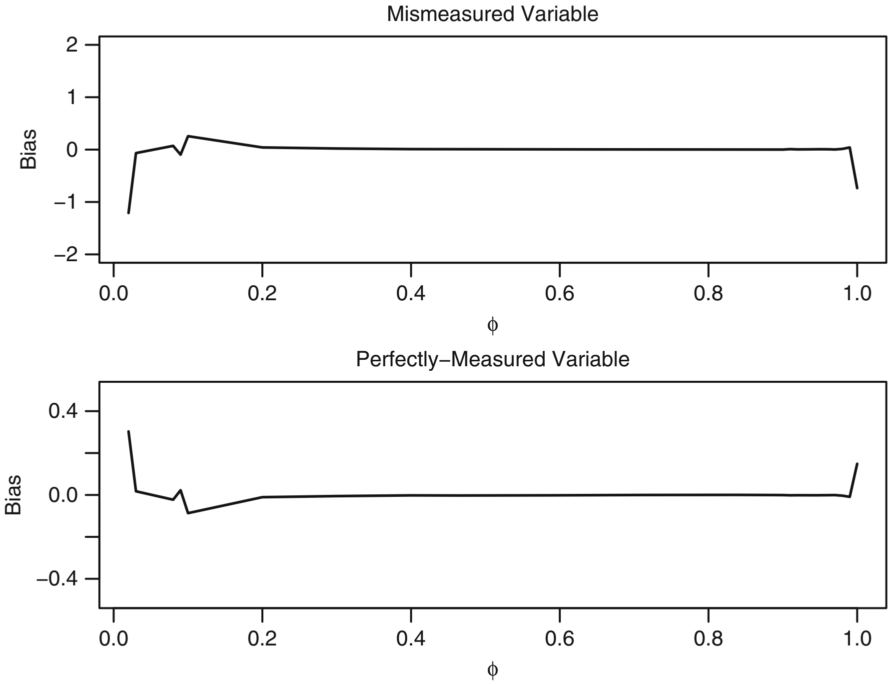


Fig. 57.3

and distribution functions (drawn in red) serve as benchmarks. The graphs in Fig. 57.4 depict the accuracy of the approximation. We calculate the density of the estimators using a simple Gaussian Kernel estimator and also estimate the empirical cumulative distribution function.[24]

Consider the lognormal distribution (first panel). In that case, the OLS-IV (black line) displays a very precise approximation to the normal curve in terms of both density and distribution. The result for the OLS-IV is robust across all of the distributions considered (lognormal, chi-square, and F). These results are in sharp contrast to those associated with the EW-GMM5 estimator. This estimator presents a poor asymptotic approximation for all distributions examined. For the lognormal case, the density is not quite centered at zero, and its shape does not fit the normal distribution. For the chi-square and F-distributions, Fig. 57.4 shows that the shapes of the density and distribution functions are very unlike the normal case, with the center of the distribution located far away from zero. These results imply that inference procedures using the EW estimator might be asymptotically invalid in simple panel data with fixed effects, even when the relevant distributions present high skewness.

[24]The empirical cumulative distribution function F_n is a step function with jumps i/n at observation values, where i is the number of tied observations at that value.

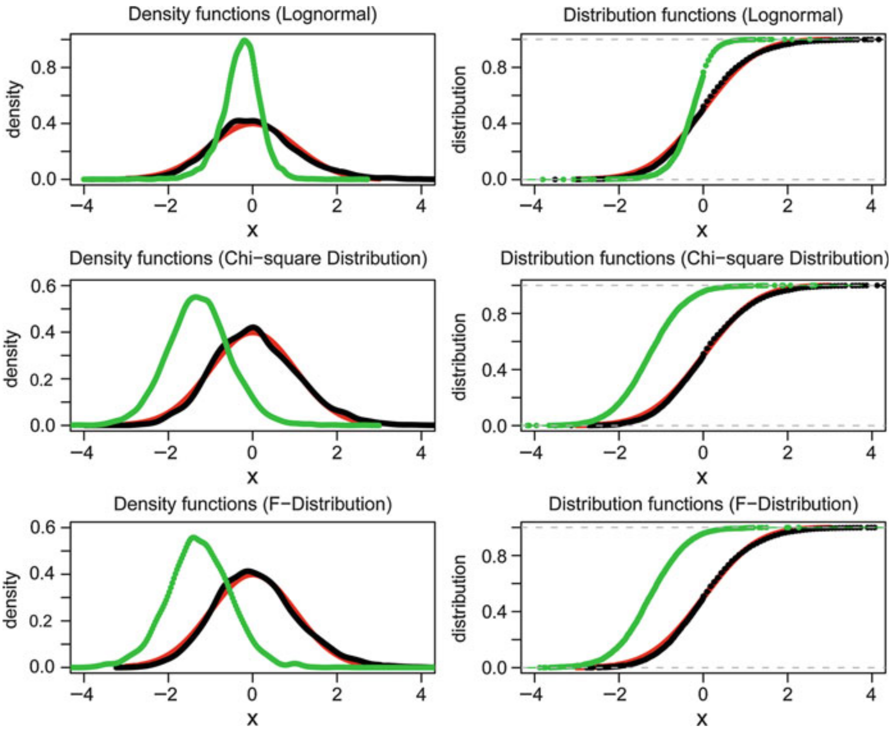


Fig. 57.4

57.4 Empirical Application

We apply the EW and OLS-IV estimators to Fazzari et al. (1988) investment equation. This is the most well-known model in the corporate investment literature, and we use this application as a way to illustrate our Monte Carlo-based results. In the Fazzari, Hubbard, and Petersen model, a firm's investment spending is regressed on a proxy for investment demand (Tobin's q) and the firm's cash flow. Theory suggests that the correct proxy for the firm's investment demand is *marginal* q, but this quantity is unobservable and researchers use instead its measurable proxy, *average* q. Because average q measures marginal q imperfectly, a measurement problem naturally arises. Erickson and Whited (2002) uses the Fazzari, Hubbard, and Petersen model to motivate the adoption of their estimator in applied work in panel data.

A review of the corporate investment literature shows that virtually all empirical work in the area considers panel data models with firm-fixed effects (Kaplan and Zingales 1997; Rauh 2006; Almeida and Campello 2007). From an estimation point of view, there are distinct advantages in exploiting repeated observations from individuals to identify the model (Blundell et al. 1992). In an investment model setting, exploiting firm effects contributes to estimation precision and allows for model consistency in the presence of unobserved idiosyncrasies that may be simultaneously correlated with investment and q. The baseline model in this literature has the form

$$I_{it}/K_{it} = \eta_i + \beta q^*_{it} + \alpha CF_{it}/K_{it} + u_{it}, \tag{57.27}$$

where I denotes investment, K capital stock, q* is marginal q, CF cash flow, η is the firm-specific effect, and u is the innovation term.

As mentioned earlier, if q* is measured with error, OLS estimates of β will be biased downwards. In addition, given that q and cash flow are likely to be positively correlated, the coefficient α is likely to be biased upwards in OLS estimations. In expectation, these biases should be reduced by the use of estimators like the ones discussed in the previous section.

57.4.1 Theoretical Expectations

In order to better evaluate the performance of the two alternative estimators, we develop some hypotheses about the effects of measurement-error correction on the estimated coefficients β and α from Eq. 57.27. Theory does not pin down the exact values that these coefficients should take. Nevertheless, one could argue that the two following conditions should be reasonable.

First, an estimator that addresses measurement error in q in a standard investment equation should return a higher estimate for β and a lower estimate for α when compared with standard OLS estimates. Recall that measurement error causes an attenuation bias on the estimate for the coefficient β. In addition, since q and cash flow are likely to be positively correlated, measurement error should cause an upward bias on the empirical estimate returned under the standard OLS estimation. Accordingly, if one denotes the OLS and the measurement-error consistent estimates, respectively, by (β^{OLS}, α^{OLS}) and (β^{MEC}, α^{MEC}), one should expect:

Condition 1. $\beta^{OLS} < \beta^{MEC}$ *and* $\alpha^{OLS} > \alpha^{MEC}$. Second, one would expect the coefficients for q and the cash flow to be nonnegative after treating the data for measurement error. The q-theory of investment predicts a positive correlation between investment and q (e.g., Hayashi 1982). If the theory holds and the estimator does a good job of adjusting for measurement error, then the cash flow coefficient should be zero ("neoclassical view"). However, the cash flow coefficient could be positive either because of the presence of financing frictions (as posited by Fazzari et al. 1988)[25] or

[25]However, financial constraints are not sufficient to generate a strictly positive cash flow coefficient because the effect of financial constraints is capitalized in stock prices and may thus be captured by variations in q (Chirinko 1993; Gomes 2001).

due to fact that cash flow picks up variation in investment opportunities even after we apply a correction for mismeasurement in q. Accordingly, one should observe:

Condition 2. $\beta^{MEC} \geq 0$ *and* $\alpha^{MEC} \geq 0$. Notice that these conditions are fairly weak. If a particular measurement-error consistent estimator does not deliver these basic results, one should have reasons to question the usefulness of that estimator in applied work.

57.4.2 Data Description

Our data collection process follows that of Almeida and Campello (2007). We consider a sample of manufacturing firms over the 1970–2005 period with data available from Compustat. Following those authors, we eliminate firm years displaying asset or sales growth exceeding 100 %, or for which the stock of fixed capital (the denominator of the investment and cash flow variables) is less than \$5 million (in 1976 dollars). Our raw sample consists of 31,278 observations from 3,084 individual firms. Summary statistics for investment, q, and cash flow are presented in Table 57.6. These statistics are similar to those reported by Almeida and Campello, among other papers. To save space we omit the discussion of these descriptive statistics.

57.4.3 Testing for the Presence of Fixed Effects and Heteroscedasticity

Before estimating our investment models, we conduct a series of tests for the presence of firm-fixed effects and heteroscedasticity in our data. As a general rule, these phenomena might arise naturally in panel data applications and should not be ignored. Importantly, whether they appear in the data can have concrete implications for the results generated by different estimators.

We first perform a couple of tests for the presence of firm-fixed effects. We allow for individual firm intercepts in Eq. 57.27 and test the null hypothesis that the coefficients associated with those firm effects are jointly equal to zero (Baltagi 2005). Table 57.7 shows that the F-statistic for this test is 4.4 (the associated p-value is 0.000). Next, we contrast the random effects OLS and the fixed effects OLS estimators to test again for the presence of fixed effects. The Hausman test statistic reported in Table 57.7 rejects the null hypothesis that the random effects model is appropriate with a test statistic of 8.2 (p-value of 0.017). In sum, standard tests strongly reject the hypothesis that fixed effects can be ignored.

We test for homoscedasticity using two different panel data-based methods. First, we compute the residuals from the least squares dummy variables estimator and regress the squared residuals on a function of the independent variables [see Frees (2004) for additional details]. We use two different combinations of independent regressors – (q_{it}, CF_{it}) and (q_{it}, q_{it}^2, CF_{it}, CF_{it}, CF_{it}^2) – and both of them robustly reject the null hypothesis of homoscedasticity. We report the results for the

Table 57.6 Descriptive statistics

Variable	Obs.	Mean	Std. dev.	Median	Skewness
Investment	22,556	0.2004	0.1311	0.17423	2.6871
q	22,556	1.4081	0.9331	1.1453	4.5378
Cash flow	22,556	0.3179	0.3252	0.27845	−2.2411

This table shows the basic descriptive statistics for q, cash flow, and investment. The data are taken from the annual Compustat industrial files over the 1970–2005 period. See text for details

Table 57.7 Diagnosis tests

Test	Test statistic	p-value
Pooling test	4.397	0.0000
Random effects vs. fixed effects	8.17	0.0169
Homoscedasticity 1	55.19	0.0000
Homoscedasticity 2	7,396.21	0.0000

This table reports results for specification tests. Hausman test for fixed effects models considers fixed effects models against the simple pooled OLS and the random effects model. A homoscedasticity test for the innovations is also reported. The data are taken from the annual Compustat industrial files over the 1970–2005 period. See text for details

first combination in Table 57.7, which yields a test statistic of 55.2 (p-value of 0.000). Our second approach for testing the null of homoscedasticity is the standard random effects Breusch-Pagan test. Table 57.7 shows that the Breusch-Pagan test yields a statistic of 7,396.2 (p-value of 0.000). Our tests hence show that the data strongly reject the hypothesis of error homoscedasticity.

57.4.4 Implementing the EW Identification Test

Our preliminary tests show that one should control for fixed effects when estimating investment models using real data. In the context of the EW estimator, it is thus appropriate to apply the within transformation before the estimation. However, in this section, we also present results for the data in level form to illustrate the point made in Sect. 57.3.2 that applying the within transformation compromises identification in the EW context. Prior papers adopting the EW estimator have ignored (or simply dismissed) the importance of fixed effects (e.g., Whited 2001, 2006).

We present the results for EW's identification test in Table 57.8. Using the data in level form, we reject the hypothesis of no identification in 12 out of 30 years (or 36 % rejection). For data that is transformed to accommodate fixed effects (within transformation), we find that in only 7 out of 33 (or 21 %) of the years between 1973 and 2005, one can reject the null hypothesis that the model is not identified at the usual 5 % level of significance. These results suggest that the power of the test is low and decreases further after applying the within transformation to the data. These results are consistent with Almeida and Campello's (2007) use of the EW estimator. Working with a 15-year Compustat panel, those authors report that they could only find a maximum of 3 years of data passing the EW identification test.

The results in Table 57.8 reinforce the notion that it is quite difficult to operationalize the EW estimator in real-world applications, particularly in situations in which the within transformation is appropriate due to the presence of fixed effects. We recognize that the EW identification test rejects the model for most of the data at hand. However, recall from Sect. 57.3.2 that the test itself is likely to be misleading ("over-rejecting" the data). In the next section, we take the EW estimator to the data (a standard Compustat sample extract) to illustrate the issues applied researchers face when using that estimator, contrasting it to an easy-to-implement alternative.

57.4.5 Estimation Results

We estimate Eq. 57.27 using the EW, OLS-IV, and AB-GMM estimators. For comparison purposes, we also estimate the investment equation using standard OLS and OLS with fixed effects (OLS-FE). The estimates for the standard OLS are likely to be biased, providing a benchmark to evaluate the performance of the other estimators. As discussed in Sect. 57.4.1, we expect estimators that improve upon the problem of mismeasurement to deliver results that satisfy Conditions 1 and 2 above.

As is standard in the empirical literature, we use an unbalanced panel in our estimations. Erickson and Whited (2000) propose a minimum distance estimator (MDE) to aggregate the cross-sectional estimates obtained for each sample year, but their proposed MDE is designed for balanced panel data. Following Riddick and Whited (2009), we use a Fama–MacBeth procedure to aggregate the yearly EW estimations.[26]

To implement our OLS-IV estimators, we first take differences of the model in Eq. 57.27. We then employ the estimator denoted by *OLS-IV A* from Sect. 57.2.2, using lagged levels of q and cash flow as instruments for (differenced) q_{it}. Our Monte Carlos suggest that identification in this context may require the use of longer lags of the model variables. Accordingly, we experiment with specifications that use progressively longer lags of q and cash flow to verify the robustness of our results.

Table 57.9 reports our findings. The OLS and OLS-FE estimates, reported in columns (1) and (2), respectively, disregard the presence of measurement error in q. The EW-GMM3, EW-GMM4, and EW-GMM5 estimates are reported in columns (3), (4), and (5). For the OLS-IV estimates reported in column (6), we use q_{t-2} as an instrument.[27] The AB-GMM estimator, reported in column (7), uses lags of q as instruments. Given our data structure, this implies using a total of 465 instruments. We account for firm-fixed effects by transforming the data.

[26]Fama–MacBeth estimates are computed as a simple standard errors for yearly estimates. An alternative approach could use the Hall–Horowitz bootstrap. For completeness, we present in the appendix the actual yearly EW estimates.

[27]In the next section, we examine the robustness of the results with respect to variation in the instrument set.

Table 57.8 The EW identification test using real data

Level				Within transformation			
			#Rejections null				#Rejections null
1973	t-statistic	1.961	0	1973	t-statistic	1.349	0
	p-value	0.375			p-value	0.509	
1974	t-statistic	5.052	0	1974	t-statistic	7.334	1
	p-value	0.08			p-value	0.026	
1975	t-statistic	1.335	0	1975	t-statistic	1.316	0
	p-value	0.513			p-value	0.518	
1976	t-statistic	7.161	1	1976	t-statistic	5.146	0
	p-value	0.028			p-value	0.076	
1977	t-statistic	1.968	0	1977	t-statistic	1.566	0
	p-value	0.374			p-value	0.457	
1978	t-statistic	9.884	1	1978	t-statistic	2.946	0
	p-value	0.007			p-value	0.229	
1979	t-statistic	9.065	1	1979	t-statistic	1.042	0
	p-value	0.011			p-value	0.594	
1980	t-statistic	9.769	1	1980	t-statistic	7.031	1
	p-value	0.008			p-value	0.03	
1981	t-statistic	10.174	1	1981	t-statistic	7.164	1
	p-value	0.006			p-value	0.028	
1982	t-statistic	3.304	0	1982	t-statistic	2.991	0
	p-value	0.192			p-value	0.224	
1983	t-statistic	5.724	0	1983	t-statistic	9.924	1
	p-value	0.057			p-value	0.007	
1984	t-statistic	15.645	1	1984	t-statistic	6.907	1
	p-value	0			p-value	0.032	

(continued)

Table 57.8 (continued)

Level				Within transformation			
			#Rejections null				#Rejections null
1985	t-statistic	16.084	1	1985	t-statistic	1.089	0
	p-value	0			p-value	0.58	
1986	t-statistic	4.827	0	1986	t-statistic	5.256	0
	p-value	0.089			p-value	0.072	
1987	t-statistic	19.432	1	1987	t-statistic	13.604	1
	p-value	0			p-value	0.001	
1988	t-statistic	5.152	0	1988	t-statistic	1.846	0
	p-value	0.076			p-value	0.397	
1989	t-statistic	0.295	0	1989	t-statistic	0.687	0
	p-value	0.863			p-value	0.709	
1990	t-statistic	0.923	0	1990	t-statistic	1.3	0
	p-value	0.63			p-value	0.522	
1991	t-statistic	3.281	0	1991	t-statistic	3.17	0
	p-value	0.194			p-value	0.205	
1992	t-statistic	2.31	0	1992	t-statistic	2.573	0
	p-value	0.315			p-value	0.276	
1993	t-statistic	1.517	0	1993	t-statistic	1.514	0
	p-value	0.468			p-value	0.469	
1994	t-statistic	2.873	0	1994	t-statistic	4.197	0
	p-value	0.238			p-value	0.123	
1995	t-statistic	0.969	0	1995	t-statistic	1.682	0
	p-value	0.616			p-value	0.431	

1996	t-statistic	17.845	1	1996	t-statistic	4.711	0
	p-value	0			p-value	0.095	
1997	t-statistic	0.14	0	1997	t-statistic	1.535	0
	p-value	0.933			p-value	0.464	
1998	t-statistic	0.623	0	1998	t-statistic	5.426	0
	p-value	0.732			p-value	0.066	
1999	t-statistic	0.354	0	1999	t-statistic	2.148	0
	p-value	0.838			p-value	0.342	
2000	t-statistic	13.44	1	2000	t-statistic	13.502	1
	p-value	0.001			p-value	0.001	
2001	t-statistic	3.159	0	2001	t-statistic	3.309	0
	p-value	0.206			p-value	0.191	
2002	t-statistic	13.616	1	2002	t-statistic	0.693	0
	p-value	0.001			p-value	0.707	
2003	t-statistic	12.904	1	2003	t-statistic	4.006	0
	p-value	0.002			p-value	0.135	
2004	t-statistic	5.212	0	2004	t-statistic	2.801	0
	p-value	0.074			p-value	0.246	
2005	t-statistic	2.365	0	2005	t-statistic	4.127	0
	p-value	0.306			p-value	0.127	
		Sum	12			Sum	7
		% of years	0.3636			% of years	0.2121

This table shows the test statistic and its *p*-value for the EW identification test, which tests the null hypothesis that the model is not identified. The tests are performed on a yearly basis. In the last columns, we collect the number of years in which the null hypothesis is rejected (sum) and compute the percentage of years in which the null is rejected. The data are taken from the annual Compustat industrial files over the 1970–2005 period. See text for details

Table 57.9 EW, GMM, and OLS-IV coefficients, real-world data

Variables	OLS	OLS-FE	EW-GMM3	EW-GMM4	EW-GMM5	OLS-IV	AB-GMM
q	0.0174***	0.0253***	0.0679	−0.3031	0.0230	0.0627***	0.0453***
	(0.002)	(0.003)	(0.045)	(0.302)	(0.079)	(0.007)	(0.006)
Cash flow	0.1310***	0.1210***	0.1299***	0.3841*	0.1554***	0.0434***	0.0460***
	(0.011)	(0.017)	(0.031)	(0.201)	(0.052)	(0.007)	(0.016)
Observations	22,556	22,556	22,556	22,556	22,556	17,348	19,748
F-stat *p*-value (first step)	–	–	–	–	–	0.000	–

This table shows the coefficients and standard deviations that we obtain when we use the OLS, EW, and the GMM estimators in Eq. 57.22. The table also displays the standard OLS-FE coefficients (after applying the differencing transformation to treat the fixed effects) in column (2) and OLS-IV in the last column. Robust standard errors in parentheses for OLS and GMM and clustered in firms for OLS-FE and OLS-IV. Each EW coefficient is an average of the yearly coefficients reported in Table 57.11 and the standard error for these coefficients is a Fama–MacBeth standard error. The table shows the EW coefficients for the data after applying the within transformation. The data are taken from the annual Compustat industrial files over the 1970–2005 period. See text for details. *, **, and *** represent statistical significance at the 10 %, 5 %, and 1 % levels, respectively

When using OLS and OLS-FE, we obtain the standard result in the literature that both q and cash flow attract positive coefficients [see columns (1) and (2)]. In the OLS-FE specification, for example, we obtain a q coefficient of 0.025 and a cash flow coefficient of 0.121. Columns (3), (4), and (5) show that the EW estimator does not deliver robust inferences about the correlations between investment, cash flow, and q. The q coefficient estimate varies significantly with the set of moment conditions used, even flipping signs. In addition, none of the q coefficients is statistically significant. The cash flow coefficient is highly inflated under EW, and in the case of the EW-GMM4 estimator, it is more than three times larger than the (supposedly biased) OLS coefficient. These results are inconsistent with Conditions 1 and 2 above. These findings agree with the Monte Carlo simulations of Sect. 57.3.3, which also point to a very poor performance of the EW estimator in cases in which fixed effects and heteroscedasticity are present.

By comparison, the OLS-IV delivers results that are consistent with Conditions 1 and 2. In particular, the q coefficient increases from 0.025 to 0.063, while the cash flow coefficient drops from 0.131 to 0.043. These results suggest that the proposed OLS-IV estimator does a fairly reasonable job at addressing the measurement-error problem. This conclusion is consistent with the Monte Carlo simulations reported above, which show that the OLS-IV procedure is robust to the presence of fixed effects and heteroscedasticity in simulated data. The AB-GMM results also generally satisfy Conditions 1 and 2. Notice, however, that the observed changes in the q and *cash flow* coefficients ("corrections" relative to the simple, biased OLS estimator) are less significant than those obtained under the OLS-IV estimation.

57.4.6 Robustness of the Empirical OLS-IV Estimator

It is worth demonstrating that the OLS-IV we consider is robust to variations in the set of instruments that is used for identification. While the OLS-IV delivered results that are consistent with our priors, note that we examined a just-identified model, for which tests of instrument quality are not available. As we have discussed previously, OLS-IV estimators should be used with care in this setting, since the underlying structure of the error in the latent variable is unknown. In particular, the Monte Carlo simulations suggest that it is important to show that the results remain when we use longer lags to identify the model.

We present the results from our robustness checks in Table 57.10. We start by adding one more lag of q (i.e., q_{t-3}) to the instrumental set. The associated estimates are in the first column of Table 57.10. One can observe that the slope coefficient associated with q increases even more with the new instrument (up to 0.090), while that of the cash flow variable declines further (down to 0.038). One problem with this estimation, however, is the associated J-statistic. If we consider a 5 % hurdle rule, the J-statistic of 4.92 implies that, with this particular instrumental set, we reject the null hypothesis that the identification restrictions are met (p-value of 3 %). As we have discussed, this could be expected if, for example, the measurement-error process has an MA structure.

Table 57.10 OLS-IV coefficients, robustness tests

Variables	(1)	(2)	(3)	(4)	(5)	(6)	(7)
q	0.0901***	0.0652***	0.0394***	0.0559***	0.0906***	0.0660***	0.0718***
	(0.014)	(0.014)	(0.015)	(0.012)	(0.033)	(0.024)	(0.026)
Cash flow	0.0383***	0.0455***	0.0434***	0.0449***	0.0421***	0.0450***	0.0444***
	(0.007)	(0.011)	(0.011)	(0.010)	(0.008)	(0.012)	(0.012)
Observations	15,264	11,890	12,000	13,448	12,000	10,524	10,524
F-stat p-value (first step)	0.000	0.000	0.000	0.000	0.001	0.000	0.000
J-stat	4.918	3.119	0.122	0.00698	0.497	8.955	5.271
J-stat p-value	0.0266	0.210	0.727	0.933	0.481	0.111	0.261

This table shows the results of varying the set of instruments that are used when applying the OLS-IV estimator to Eq. 57.22. In the first column we use the second and third lags of q as instruments for current (differenced) q, as in Table 57.6. In column (2), we use third, fourth, and fifth lags of q as instruments. In column (3), we use the fourth and fifth lags of q and the first lag of cash flow as instruments. In column (4), we use the third lag of q and fourth lag of cash flow as instruments. In column (5), we use the fourth and fifth lags of cash flow as instruments. In column (6), we use $\{q_{t-4}, q_{t-5}, q_{t-6}, CF_{t-3}, CF_{t-4}, CF_{t-5}\}$ as instruments. Finally, in column (7) $\{q_{t-5},q_{t-6},CF_{t-3},CF_{t-4},CF_{t-5}\}$ as instruments. The estimations correct the errors for heteroscedasticity and firm-clustering. The data are taken from the annual Compustat industrial files over the 1970–2005 period. See text for details. *, **, and *** represent statistical significance at the 10 %, 5 %, and 1 % levels, respectively

This suggests that the researcher should look for longer lagging schemes, lags that "erase" the MA memory of the error structure.

Our next set of estimations use longer lagging structures for our proposed instruments and even an instrumental set with only lags of cash flow, the exogenous regressors in the model. We use combinations of longer lags of q (such as the fourth and fifth lags) and longer lags of cash flow (fourth and fifth lags). This set of tests yields estimates that more clearly meet standard tests for instrument validity.[28] Specifically, the J-statistics now indicate we do not reject the hypothesis that the exclusion restrictions are met. The results reported in columns (2) through (7) of Table 57.10 also remain consistent with Conditions 1 and 2. In particular, the q coefficient varies from approximately 0.040 to 0.091, while the cash flow coefficient varies roughly from 0.044 to 0.046. These results are consistent with our simulations, which suggest that these longer lag structures should deliver relatively consistent, stable estimates of the coefficients for q and cash flow in standard investment regressions.

57.5 Concluding Remarks

OLS estimators have been used as a reference in empirical work in financial economics. Despite their popularity, those estimators perform poorly when dealing with the problem of errors in variables. This is a serious problem since in most empirical applications, one might raise concerns about issues such as data quality and measurement errors.

This chapter uses Monte Carlo simulations and real data to assess the performance of different estimators that deal with measurement error, including EW's higher-order moment estimator and alternative instrumental variable-type approaches. We show that in the presence of individual fixed effects, under heteroscedasticity, or in the absence of high degree of skewness in the data, the EW estimator returns biased coefficients for both mismeasured and perfectly measured regressors. The IV estimator requires assumptions about the autocorrelation structure of the measurement error, which we characterize and discuss in the chapter.

We also estimate empirical investment models using the two methods. Because real-world investment data contain firm-fixed effects and heteroscedasticity, the EW estimator delivers coefficients that are unstable across different specifications and not economically meaningful. In contrast, a simple OLS-IV estimator yields results that conform to theoretical expectations. We conclude that real-world investment data is likely to satisfy the assumptions that are required for identification of OLS-IV but that the presence of heteroscedasticity and fixed effects causes the EW estimator to return biased coefficients.

[28] All of the F-statistics associated with the first-stage regressions have p-values that are close to zero. These statistics (reported in Table 57.10) suggest that we do not incur a weak instrument problem when we use longer lags in our instrumental set.

Table 57.11 EW coefficients for real data (within transformation)

q coefficient				Cash flow coefficient		
Year	GMM3	GMM4	GMM5	GMM3	GMM4	GMM5
1973	−0.029	0.000	0.000	0.347	0.265	0.264
	(0.075)	(0.073)	(4.254)	(0.207)	(0.207)	(11.968)
1974	0.050	0.029	0.019	0.168	0.199	0.214
	(0.037)	(0.012)	(0.016)	(0.073)	(0.043)	(0.043)
1975	0.225	0.001	0.000	0.161	0.292	0.292
	(0.475)	(0.149)	(0.125)	(0.281)	(0.095)	(0.094)
1976	0.137	0.001	0.000	0.156	0.276	0.276
	(0.094)	(0.273)	(0.042)	(0.090)	(0.251)	(0.048)
1977	0.082	0.243	0.000	0.203	0.091	0.261
	(0.263)	(0.109)	(0.108)	(0.179)	(0.090)	(0.083)
1978	0.263	0.514	0.281	0.122	(0.067)	0.108
	(0.282)	(0.927)	(0.146)	(0.224)	(0.689)	(0.125)
1979	0.020	0.001	0.001	0.249	0.266	0.266
	(0.161)	(0.048)	(0.031)	(0.155)	(0.056)	(0.044)
1980	0.349	0.116	0.183	0.021	0.219	0.163
	(0.294)	(0.071)	(0.055)	(0.273)	(0.074)	(0.067)
1981	0.334	0.185	0.324	−0.145	0.061	−0.131
	(0.165)	(0.045)	(0.128)	(0.248)	(0.093)	(0.191)
1982	0.109	0.383	0.238	0.125	−0.206	−0.031
	(0.155)	(0.316)	(0.126)	(0.195)	(0.398)	(0.174)
1983	0.081	0.001	0.001	0.132	0.184	0.184
	(0.037)	(0.041)	(0.059)	(0.033)	(0.034)	(0.040)
1984	0.230	0.210	0.185	0.125	0.138	0.154
	(0.083)	(0.050)	(0.043)	(0.067)	(0.052)	(0.048)
1985	0.198	0.349	0.230	0.050	(0.018)	0.035
	(0.483)	(0.137)	(0.024)	(0.212)	(0.086)	(0.032)
1986	0.672	0.244	0.593	−0.179	0.070	−0.133
	(0.447)	(0.089)	(0.162)	(0.303)	(0.079)	(0.128)
1987	0.102	0.104	0.115	0.078	0.078	0.077
	(0.039)	(0.020)	(0.003)	(0.021)	(0.021)	(0.020)
1988	0.129	0.179	0.148	0.030	0.027	0.029
	(0.051)	(0.029)	(0.014)	(0.011)	(0.007)	(0.007)
1989	−0.365	−0.015	−0.111	0.285	0.162	0.196
	(1.797)	(0.082)	(0.196)	(0.642)	(0.063)	(0.078)
1990	−0.437	−0.419	−0.529	0.395	0.386	0.440
	(0.404)	(0.137)	(0.024)	(0.214)	(0.094)	(0.093)
1991	0.384	0.260	0.240	−0.098	0.007	0.023
	(0.225)	(0.105)	(0.038)	(0.199)	(0.099)	(0.055)
1992	0.105	0.102	0.040	0.086	0.088	0.148
	(0.016)	(0.008)	(0.016)	(0.034)	(0.033)	(0.037)
1993	0.274	0.322	0.452	−0.076	−0.118	−0.232
	(0.394)	(0.352)	(0.273)	(0.360)	(0.297)	(0.276)

(*continued*)

Table 57.11 (continued)

q coefficient				Cash flow coefficient		
Year	GMM3	GMM4	GMM5	GMM3	GMM4	GMM5
1994	−0.110	−4.436	−0.047	0.255	3.550	0.207
	(0.136)	(86.246)	(0.011)	(0.108)	(65.488)	(0.045)
1995	−0.574	−8.847	−2.266	0.537	5.898	1.633
	(1.862)	(145.827)	(5.565)	(1.275)	(94.154)	(3.749)
1996	0.220	0.167	0.196	0.101	0.106	0.103
	(0.068)	(0.022)	(0.013)	(0.036)	(0.033)	(0.030)
1997	0.089	0.177	0.158	0.059	0.020	0.028
	(0.082)	(0.042)	(0.021)	(0.042)	(0.041)	(0.034)
1998	−0.620	−0.245	−0.119	0.688	0.355	0.242
	(1.634)	(0.187)	(0.027)	(1.446)	(0.169)	(0.037)
1999	−0.031	−0.003	0.000	0.160	0.126	0.123
	(0.059)	(0.028)	(0.055)	(0.074)	(0.038)	(0.068)
2000	0.071	0.126	0.118	0.032	−0.029	−0.021
	(0.024)	(0.030)	(0.020)	(0.043)	(0.057)	(0.051)
2001	0.050	0.077	0.055	0.034	0.020	0.031
	(0.021)	(0.020)	(0.013)	(0.016)	(0.016)	(0.012)
2002	0.047	0.048	0.048	0.030	0.030	0.030
	(0.128)	(0.016)	(0.014)	(0.033)	(0.013)	(0.012)
2003	0.131	0.066	0.157	−0.013	0.025	−0.027
	(0.043)	(0.025)	(0.010)	(0.031)	(0.026)	(0.014)
2004	0.005	0.030	0.030	0.092	0.079	0.079
	(0.066)	(0.018)	(0.009)	(0.045)	(0.034)	(0.034)
2005	0.049	0.029	0.026	0.078	0.095	0.098
	(0.025)	(0.009)	(0.011)	(0.040)	(0.032)	(0.034)
Fama–MacBeth standard error	0.0679	−0.3031	0.0232	0.1299	0.3841	0.1554
	0.0455	0.3018	0.0787	0.0310	0.2027	0.05222

This table shows the coefficients and standard deviations that we obtain when we use the EW estimator in Eq. 57.22, estimated year by year. The table also shows the results for the EW estimator associated with GMM3, GMM4, and GMM5. The table also shows the EW coefficients for the data that is treated for fixed effects via the within transformation. The data are taken from the annual Compustat industrial files over the 1970–2005 period. See text for details

References

Agca, S., & Mozumdar, A. (2007). Investment-cash flow sensitivity: Myth or reality? Working paper, George Washington University.

Almeida, H., & Campello, M. (2007). Financial constraints, asset tangibility and corporate investment. *Review of Financial Studies, 20*, 1429–1460.

Altonji, J., & Segal, L. (1996). Small-sample bias in GMM estimation of covariance structures. *Journal of Business & Economic Statistics, 14*, 353–366.

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *Review of Economic Studies, 58*, 277–297.

Bakke, T., & Whited, T. (2009). What gives? A study of firms' reactions to cash shortfalls. Working paper, University of Rochester.

Baltagi, B. (2005). *Econometric analysis of panel data* (3rd ed.). Chichester: Wiley.

Baum, C., Schaffer, M., & Stillman, S. (2003). Instrumental variables and GMM: Estimation and testing. *Stata Journal, 3*, 1–31.

Bertrand, M., & Mullainathan, S. (2005). Bidding for oil and gas leases in the Gulf of Mexico: A test of the free cash flow model?. *mimeo*, University of Chicago and MIT.

Bertrand, M., & Schoar, A. (2003). Managing with style: The effect of managers on firm policies. *Quarterly Journal of Economics, 118*, 1169–1208.

Bertrand, M., Duflo, E., & Mullainathan, S. (2004). How much should we trust differences-in-differences estimates? *Quarterly Journal of Economics, 119*, 249–275.

Biorn, E. (2000). Panel data with measurement errors: Instrumental variables and GMM procedures combining levels and differences. *Econometric Reviews, 19*, 391–424.

Blanchard, O., Lopez-de-Silanes, F., & Shleifer, A. (1994). What do firms do with cash windfalls? *Journal of Financial Economics, 36*, 337–360.

Blundell, R., Bond, S., Devereux, M., & Schiantarelli, F. (1992). Investment and Tobin's Q: Evidence from company panel data. *Journal of Econometrics, 51*, 233–257.

Bond, S., & Cummins, J. (2000). The stock market and investment in the new economy. *Brookings Papers on Economic Activity, 13*, 61–124.

Chirinko, R. (1993). Business fixed investment spending: Modeling strategies, empirical results, and policy implications. *Journal of Economic Literature, 31*, 1875–1911.

Colak, G., & Whited, T. (2007). Spin-offs, divestitures, and conglomerate investment. *Review of Financial Studies, 20*, 557–595.

Cragg, J. (1997). Using higher moments to estimate the simple errors-in-variables model. *RAND Journal of Economics, 28*, 71–91.

Doran, H., & Schmidt, P. (2006). GMM estimators with improved finite sample properties using principal components of the weighting matrix, with an application to the dynamic panel data model. *Journal of Econometrics, 133*, 387–409.

Erickson, T., & Whited, T. (2000). Measurement error and the relationship between investment and Q. *Journal of Political Economy, 108*, 1027–1057.

Erickson, T., & Whited, T. (2002). Two-step GMM estimation of the errors-in-variables model using high-order moments. *Econometric Theory, 18*, 776–799.

Fazzari, S., Hubbard, R. G., & Petersen, B. (1988). Financing constraints and corporate investment. *Brooking Papers on Economic Activity, 1*, 141–195.

Frees, E. (2004). *Longitudinal and panel data analysis and applications in the social sciences*. New York: Cambridge University Press.

Gan, J. (2007). Financial constraints and corporate investment: Evidence from an exogenous shock to collateral. *Journal of Financial Economics, 85*, 709–734.

Gomes, J. (2001). Financing investment. *American Economic Review, 91*, 1263–1285.

Griliches, Z., & Hausman, J. A. (1986). Errors in variables in panel data. *Journal of Econometrics, 31*, 93–118.

Hadlock, C. (1998). Ownership, liquidity, and investment. *RAND Journal of Economics, 29*, 487–508.

Hayashi, F. (1982). Tobin's marginal q and average q: A neoclassical interpretation. *Econometrica, 50*, 213–224.

Hayashi, F. (2000). *Econometrics*. Princeton: Princeton University Press.

Hennessy, C. (2004). Tobin's Q, debt overhang, and investment. *Journal of Finance, 59*, 1717–1742.

Holtz-Eakin, D., Newey, W., & Rosen, H. S. (1988). Estimating vector autoregressions with panel data. *Econometrica, 56*, 1371–1395.

Hoshi, T., Kashyap, A., & Scharfstein, D. (1991). Corporate structure, liquidity, and investment: Evidence from Japanese industrial groups. *Quarterly Journal of Economics, 106*, 33–60.

Hubbard, R. G. (1998). Capital market imperfections and investment. *Journal of Economic Literature, 36*, 193–227.

Kaplan, S., & Zingales, L. (1997). Do financing constraints explain why investment is correlated with cash flow? *Quarterly Journal of Economics, 112*, 169–215.

Lamont, O. (1997). Cash flow and investment: Evidence from internal capital markets. *Journal of Finance, 52*, 83–110.

Lyandres, E. (2007). External financing costs, investment timing, and investment-cash flow sensitivity. *Journal of Corporate Finance, 13*, 959–980.

Malmendier, U., & Tate, G. (2005). CEO overconfidence and corporate investment. *Journal of Finance, 60*, 2661–2700.

Petersen, M. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies, 22*, 435–480.

Poterba, J. (1988). Financing constraints and corporate investment: Comment. *Brookings Papers on Economic Activity, 1*, 200–204.

Rauh, J. (2006). Investment and financing constraints: Evidence from the funding of corporate pension plans. *Journal of Finance, 61*, 33–71.

Riddick, L., & Whited, T. (2009). The corporate propensity to save. *Journal of Finance, 64*, 1729–1766.

Shin, H., & Stulz, R. (1998). Are internal capital markets efficient? *Quarterly Journal of Economics, 113*, 531–552.

Stein, J. (2003). Agency information and corporate investment. In G. Constantinides, M. Harris, & R. Stulz (Eds.), *Handbook of the economics of finance*. Amsterdam: Elsevier/North-Holland.

Wansbeek, T. J. (2001). GMM estimation in panel data models with measurement error. *Journal of Econometrics, 104*, 259–268.

Whited, T. (2001). Is it inefficient investment that causes the diversification discount? *Journal of Finance, 56*, 1667–1691.

Whited, T. (2006). External finance constraints and the intertemporal pattern of intermittent investment. *Journal of Financial Economics, 81*, 467–502.

Xiao, Z., Shao, J., & Palta, M. (2008). A unified theory for GMM estimation in panel data models with measurement error. Working paper, University of Wisconsin Madison.

58 Realized Distributions of Dynamic Conditional Correlation and Volatility Thresholds in the Crude Oil, Gold, and Dollar/Pound Currency Markets

Tung-Li Shih, Hai-Chin Yu, Der-Tzon Hsieh, and Chia-Ju Lee

Contents

Abstract

This chapter proposes a modeling framework for the study of co-movements in price changes among crude oil, gold, and dollar/pound currencies that are conditional on volatility regimes. Methodologically, we extend the dynamic

T.-L. Shih
Department of Hospitality Management, Ming Dao University, Changhua Peetow, Taiwan
e-mail: tungli@mdu.edu.tw

H.-C. Yu (✉)
Department of International Business, Chung Yuan University, Chungli, Taiwan
e-mail: haichin@cycu.edu.tw; haichinyu@hotmail.com

D.-T. Hsieh
Department of Economics, National Taiwan University, Taipei, Taiwan
e-mail: dthsieh@ccms.ntu.edu.tw

C.-J. Lee
College of Business, Chung Yuan University, Chungli, Taiwan
e-mail: g9604601@cycu.edu.tw

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_58,

conditional correlation (DCC) multivariate GARCH model to examine the volatility and correlation dynamics depending on the variances of price returns involving a threshold structure. The results indicate that the periods of market turbulence are associated with an increase in co-movements in commodity (gold and oil) prices. By contrast, high market volatility is associated with a decrease in co-movements between gold and the dollar/pound or oil and the dollar/pound. The results imply that gold may act as a safe haven against major currencies when investors face market turmoil. By looking at different subperiods based on the estimated thresholds, we find that the investors' behavior changes in different subperiods. Our model presents a useful tool for market participants to engage in better portfolio allocation and risk management.

Keywords

Dynamic conditional correlation • Volatility threshold • Realized distribution • Currency market • Gold • Oil

58.1 Introduction

Commodity markets in recent years have experienced dramatic growth in trading volume as well as widespread price volatility. With few exceptions, most of the commodities have experienced an impressive bull run and have generally outperformed traditional investments. For example, the prices of commodities such as crude oil have risen dramatically, and the crude oil price almost reached a new high of US$200 per barrel in 2011. In the meantime, the price of gold hit a new high of US$1,700 in 2011. These price surprises have influenced not only the commodity markets but also the currency markets and the international parity of foreign exchange. By the fall of 2007, the increasing speculation in commodity markets was associated with the devaluation of the US dollar.

Among these commodities, gold appears to have exhibited a more stable price trend than crude oil. From the beginning of the financial crisis in 1997 up until 2011, the price of gold has risen by almost 42 %. For many years, gold has been viewed as a safe haven from market turbulence. However, very few empirical studies have examined the role of gold as a safe-haven asset and even fewer have examined gold's safe-haven role with respect to major currency exchange rates, especially those of the two major currencies – the US dollar and the British pound.

The reason why we choose exchange rates as a comparative baseline is that, for commodities that are traded continuously in organized markets, a change in a major currency exchange rate will result in an instant adjustment in the prices of commodities in at least one currency and perhaps in both currencies if both countries are "large." For instance, when the dollar depreciates against the pound, the dollar prices of commodities tend to rise (and pound prices fall) even though the fundamentals of the markets remain unchanged.

This widely expanded and complex volatility in commodity prices increases the importance of modeling real volatility and correlation, because a good estimate helps facilitate portfolio optimization, risk management, and hedging activities. Although some of the literature assumes volatility and correlation to be constant in the past years, it is widely recognized that they indeed vary over time. This recognition has spurred a vibrant body of work regarding the dynamic properties of market volatility. To date, very little is known about the volatility dynamics between the commodity and currency markets, for instance, in the case of gold and its possible correlations with oil and major currencies. This chapter intends to address this gap.

The main purpose of this study is to examine the dynamic relationships among gold, oil, and the dollar/pound to further understand the hedging ability of gold relative to another commodity or currency. That is to say, if gold acts as a financial safe haven against the dollar (or oil), it allows for systematic feedback between changes in the price of gold, oil, and the dollar/pound exchange rate. Specifically, this chapter asks, does gold act as a safe haven against the dollar/pound, as a hedge, or as neither? Are gold and oil highly correlated with each other? Movements in the price of gold, oil, and the dollar/pound are analyzed using a model of dynamic conditional correlation covering 20 years of daily data.

Studies related to this issue are few. Capie et al. (2005) point out that gold acts as an effective hedge against the US dollar by estimating elasticity relative to changes in the exchange rate. However, their approach involves the use of a single-equation model in which the exchange rate is assumed to be unaffected by the time of the dependent variable, the price of gold. Our chapter improves their work by employing a dynamic model of conditional correlations in which all variables are treated symmetrically. Besides, although Baur and Lucey (2010) find evidence in support of gold providing a haven from losses incurred in the bond and stock markets, they neglect the interactions with the currency market and, like Capie et al. (2005), do not consider feedback in their model of returns. Nikos (2006) uses correlation analysis to estimate the correlation of returns between gold and the dollar and shows that the correlation between the dollar and gold is −0.19 and −0.51 for two different periods. These findings imply that gold is a contemporaneous safe haven in extreme currency market conditions. Steinitz (2006) utilizes the same method to estimate the correlations of weekly returns between gold and Brent oil for two periods of 1 year and 5 years, respectively, and shows that the correlations between gold and Brent oil are 0.310 and 0.117, respectively. Boyer and Fillion (2007) report on the financial determination of Canadian oil and gas stock returns and conclude that a weakening of the Canadian dollar against the US dollar has a negative impact on stock returns.

If correlations and volatilities vary over time, the hedge ratio should be adjusted to account for the new information. Other work, such as Baur and McDermott (2010), similarly neglects feedback in its regression model. Further studies investigate the concept of a safe-haven asset without reference to gold. For example, Ranaldo and Soderlind (2009) and Kaul and Sapp (2006) examine safe-haven currencies, while Upper (2000) examines German government bonds as safe-haven instruments. Andersen et al. (2007) show that exchange rate volatility outstrips bond volatility in the US, British, and German markets. Thus, currency risk is worth exploring and being hedged.

As a general rule, commodities are priced in US dollars. Since the US currency has weakened that a bull run of commodity prices appeared, the question arises as to which the increases in commodity prices have been a product of the depreciation in the US dollar. Furthermore, it would be interesting to examine how to provide a hedge against the dollar that varies across different commodities. It also needs to be asked which investment instruments are more suitable for diversification purposes to protect against changes in the US currency.

This chapter investigates the following issues. First, how do the time-varying correlations and associated distributions appear in the crude oil, gold, and dollar/pound markets? Second, what is the shape of each separate distribution of various volatility levels among the crude oil, gold, and dollar/pound markets? Third, by employing the volatility threshold DCC model put forward by Kasch and Caporin (2012), is the high volatility (exceeding a specified threshold) of the assets associated with an increasing degree of correlation?

We find that the volatility thresholds of oil and gold correspond to two major events – *the First Gulf War in 1990 and the 911 event in 2001*. We also find that the increase in commodity (crude oil and gold) prices was a reflection of the falling US dollar, especially after the 911 event. The evidence shows that the DCC between crude oil and gold was 0.1168, while those for the gold/dollar/pound and oil/dollar/pound markets were −0.2826 and −0.0369, respectively, with the latter being significantly higher than in the other subperiods.

This remainder of this chapter is organized as follows. Section 58.2 provides a review of the literature. Section 58.3 describes the data and summary statistics for crude oil, gold, and the dollar/pound exchange rate. Section 58.4 presents the dynamic conditional correlation model and reports the results of its volatility threshold. And also provides the results for subperiods separated by the thresholds found. Finally, Sect. 58.5 discusses the results and concludes.

58.2 Literature Review

Engle et al. (1994) investigate how the returns and volatilities of stock indices between Tokyo and New York are correlated and find that, except for a lagged return spillover from New York to Tokyo after the crash, there was no significant lagged spillover in returns or in volatilities. Ng (2000) examines the size and the impact of volatility spillover from Japan and the USA to six Pacific Basin equity markets. Using four different specifications of correlation by constructing volatility spillover models, he distinguishes the volatility between local idiosyncratic shock, regional shock from Japan, and global shock from the USA and finds significant spillover effects from regional to Pacific Basin economies. Andersen et al. (2001a) find strong evidence that volatilities and correlations move together in a manner broadly consistent with the latent factor structure. Andersen et al. (2001b) found that volatility movements are highly correlated across the deutsche mark and yen against the US dollar. Furthermore, the correlation between the two exchange rates increases with volatility. Engle (2002) finds that the breakdown of the correlations between the deutsche mark and the pound and lira in August 1992 is very apparent. In addition, after the euro is launched, the estimated currency correlation essentially moves to 1.

Recently, Doong et al. (2005) examined the dynamic relationship and pricing between stocks and exchange rates for six Asian emerging markets. They found that the currency depreciation is accompanied by a fall in stock prices. The conditional variance-covariance process of changes in stock prices and exchange rates is time varying. Lanza et al. (2006) estimate the dynamic conditional correlations in the daily returns for West Texas Intermediate (WTI) oil forward and future prices from January 3, 1985 to January 16, 2004, and find that the dynamic conditional correlations vary dramatically. Chiang et al. (2009) investigate the probability distribution properties, autocorrelations, dynamic conditional correlations, and scaling analysis of Dow-Jones and NASDAQ Intraday returns from August 1, 1997 to December 31, 2003. They find the correlations to be positive and to mostly fluctuate in the range of 0.6–0.8. Furthermore, the variance of the correlation coefficients has been declining and appears to be stable during the post-2001 period. Pérez-Rodríguez (2006) applies a multivariate DCC-GARCH technique to examine the structure of the short-run dynamics of volatility returns on the euro, yen, and British pound against the US dollar over the period from 1999 to 2004 and finds strong dynamic relationships between currencies. Tastan (2006) applies multivariate GARCH to capture the time-varying variance-covariance matrix for stock market returns (Dow-Jones Industrial Average Index and S&P500 Index) and changes in exchange rates (euro/dollar exchange rates). He also plots news impact surfaces for variances, covariances, and correlation coefficients to sort out the effects of shocks. Chiang et al. (2007a) apply a dynamic conditional correlation model to nine Asian daily stock-return series from 1990 to 2003 and find evidence of a contagion effect and herding behavior. Chiang et al. (2007b) examine A-share and B-share market segmentation conditions by employing a dynamic multivariate GARCH model and show that stock returns in both A- and B-shares are positively correlated with the daily change in trading volume or abnormal volume.

58.3 Data

Our data consist of the daily prices of crude oil and gold, and the US dollar/British pound exchange rate, and are obtained from the *AREMOS* database over the period from January 1, 1986 to December 31, 2007 for a total of 5,165 observations. The West Texas Intermediate crude oil price is chosen to represent the oil spot market, and the price of 99.5 % fine gold, the London afternoon fixing,

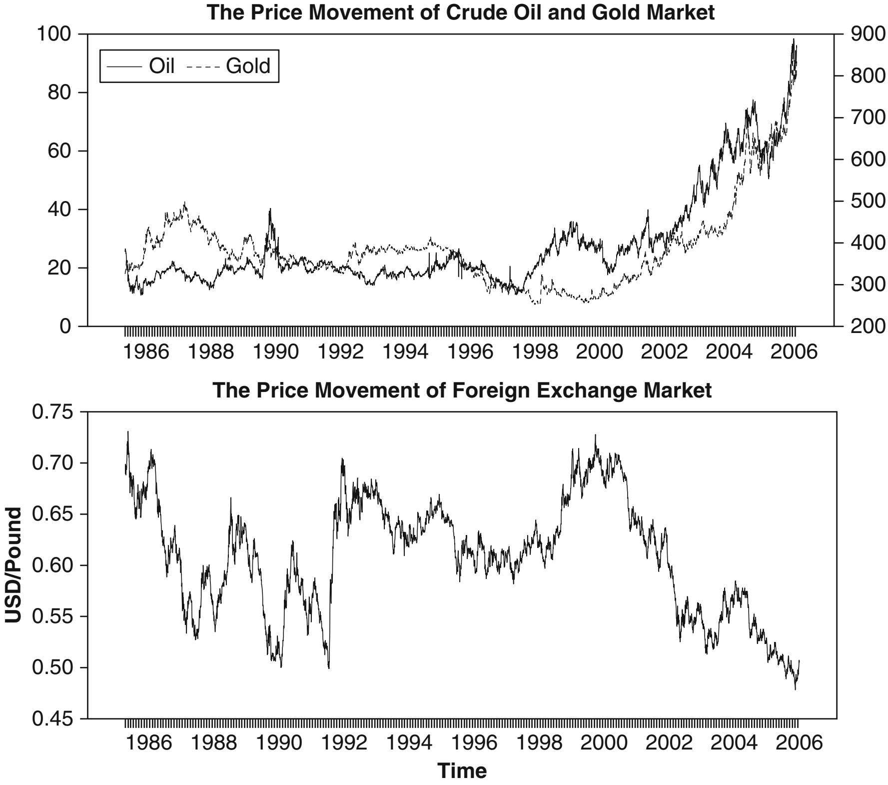


Fig. 58.1 The price movement for the sampled markets from January 1, 1986 to December 31, 2007, for a total of 5,405 observations

is chosen to represent the gold spot market. The daily dollar/pound exchange rate, which represents the major currencies, is selected to estimate the volatility of the FX market.

In our sample period, the crude oil price was testing the $100 per barrel threshold by November 2007. Meanwhile, the price of gold was relatively stable varying between $415 and $440 per ounce from January to September 2005. However, in the fourth quarter of 2005, the gold price jumped dramatically and hit $500 per ounce. In April 2006, the gold price broke through the $640 level. 2007 was a strong year, with the price steadily rising from $640 on January 2 with a closing London fixed price of over $836 on December 31, 2007. Since then, prices have continued to increase to reach new record highs of over $1,700 in 2011.

Figure 58.1 displays the price movements for oil, gold, and the dollar/pound over the sample period. As shown in Fig. 58.1, gold traded between a low of $252 (August 1999) and a high of $836 (December 31, 2007) per ounce at the fixing, while oil traded between a low of $10 (in late 1998, in the wake of the Asian

Table 58.1 Summary statistics of the daily returns among crude oil, gold, and dollar/pound[a] (January 1, 1986 to December 31, 2007)

	Crude oil	Gold	Dollar/pound
Mean	0.024	0.017	−0.006
	(0.536)	(0.159)	(0.4769)
Max	0.437	0.070	0.0379
Min	−0.404	−0.063	−0.0329
Standard dev.	0.029	0.009	0.006
Skewness[b]	−0.012	−0.031	0.164**
	(0.729)	(0.354)	(0.0000)
Kurtosis[b]	37.915**	5.968**	2.439**
	(0.0000)	(0.0000)	(0.0000)
Jarque-Bera[c]	323,692.99**	8,019.20**	1,363.59**
	(0.0000)	(0.0000)	(0.0000)

[a]The table summarizes the daily returns of estimates for the West Texas Intermediate crude oil, gold, and dollar/pound markets. The sample covers the period from January 1, 1986 through December 31, 2007 for a total of 5,405 observations
[b]The three markets are far away from the skewness and kurtosis of 0 and 3, respectively, implying that the three markets are not normally distributed
[c]Jarque-Bera is the Jarque-Bera test statistic, distributed χ^2_2
**Denotes significance at the 0.05 level

Financial Crisis and the United Nations' oil-for-food program) and a high of $99.3 (November 2007) per barrel. These large variations in the price of both gold and oil indicate that the DCC and realized distribution are better approaches for detecting the trading pattern of investors.

Table 58.1 reports the statistics of daily returns for crude oil, gold, and the dollar/pound exchange rate. The daily returns are calculated as the first differences of the natural log of the prices times 100. The results show that the crude oil has the highest return, followed by gold and the dollar/pound.

58.4 Dynamic Conditional Correlation

58.4.1 Dynamic Conditional Correlation Between Gold, Oil, and the Dollar/Pound

It is generally recognized that financial markets are highly integrated in terms of price movements, since prices soaring in one market can spill over to another market instantly. One simple method to explore the relationship between the two markets is to calculate the correlation coefficient. We then specify a multivariate model, which is capable of computing the dynamic conditional correlation (DCC) that is capable of capturing ongoing market elements and shocks. The DCC model is specified as Eq. 58.1.

Table 58.2 The correlation among crude oil, gold, and FX of dollar/pound (January 1, 1986 to December 31, 2007)

	Oil	Gold	FX
Oil	1		
Gold	0.7488	1	
FX	−0.5592	−0.6260	1

$$\rho_{ij,t} = \frac{E_{t-1}\left(r_{i,t} r_{j,t}\right)}{\sqrt{E_{t-1}\left(r_{i,t}^2\right) E_{t-1}\left(r_{j,t}^2\right)}} \tag{58.1}$$

where the conditional correlation $\rho_{ij,t}$ is based on information known in the previous period E_{t-1} and i, j represent the three markets 1, 2, and 3. Based on the laws of probability, all correlations defined in this way must lie within the interval [−1, 1]. This is different from the constant correlation we have usually used and assumed throughout a given period. To clarify the relationship between the conditional correlations and conditional variances, it is convenient to express the returns as the conditional standard deviation times the standardized disturbance as suggested by Engle (2002) in Eq. 58.2 below:

$$h_{i,t} = E_{t-1}\left(r_{i,t}^2\right), \quad r_{i,t} = \sqrt{h_{i,t}\varepsilon_{i,t}}, \qquad i = 1, 2, 3 \tag{58.2}$$

Since the correlation coefficients among crude oil, gold, and dollar/pound FX markets provide useful measures of the long-term relationship between each pair of markets, Table 58.2 presents a simple correlation matrix in which the calculation is based on the constant coefficient given by Eq. 58.1. Some preliminary information is obtained below. First, the crude oil and gold are highly correlated with a coefficient of 0.7488, a result that is in line with Steinitz (2006). Secondly, both gold and crude oil are highly negatively related to the dollar/pound with coefficients of −0.6260 and −0.5592, respectively, which is consistent with the report of Nikos (2006).

As the autoregressive conditional heteroskedasticity (ARCH) model has become the most useful model in investigating the conditional volatility since Engle (1982), we then follow this model in our analysis. The ARCH model adopts the effect of past residuals that helps explain the phenomenon of volatility clustering. Bollerslev (1986) proposed the generalized autoregressive conditional heteroskedasticity (GARCH) model, which has created a new field in the research on volatility and is widely used in financial and economic time series. Some of his research attempts to discuss the effects of more than one variable simultaneously. For instance, Bollerslev (1990) proposed the constant conditional correlation (CCC) model which makes a strong assumption, namely, that the correlation among the variables remains constant in order to simplify the estimation. Engle (2002) later proposed a dynamic conditional correlation (DCC) model, which allows the correlation to be time varying and, by involving fewer complicated calculations, is capable of dealing with numerous variables.

In this chapter, we follow the Engle (2002) approach, which has clear computational advantages over multivariate GARCH models in that the number of parameters to be estimated remains constant and loosens the assumptions of the multivariate conditional correlations, in order to develop the dynamic conditional correlation (DCC) model. The DCC model can be viewed as a generalization of the Bollerslev (1990) constant conditional correlation (CCC) estimator. It differs only in that it allows the correlation to be time varying, which parameterizes the conditional correlations directly. The estimation takes place in two steps, in that a series of univariate GARCH estimates are first obtained followed by the correlation coefficients. The characteristics of the DCC model are that the multivariate conditional correlations are dynamic and not constant and confirm that the real conditional correlations of financial assets in general and the time-varying covariance matrices can be estimated. This model involves a less complicated calculation without losing too much generality and is able to deal with numerous variables.

Following Engle (2002) and Chiang et al. (2009), the mean equation is assumed to be represented by Eq. 58.1, where the multivariate conditional variance is given by

$$H_{\tau,t} = D_{\tau,t} V_{\tau,t} D_{\tau,t}, \tag{58.3}$$

where τ is a time interval, which can be a day, an hour, or 1 min. Here τ is a daily interval. $V_{\tau,t}$ is a symmetric conditional correlation matrix of ε_t and $D_{\tau,t}$ is a (2×2) matrix with the conditional variances $h_{\tau,ii,t}$ for two stock returns (where i = gold, oil, or the dollar/pound exchange rate) on the diagonal. That is, $D_{\tau,t} = diag\left[\sqrt{\sigma^2_{\tau,ii,t}}\right]_{(2,2)}$. Equation 58.3 suggests that the dynamic properties of the covariance matrix $H_{\tau,t}$ are determined by $D_{\tau,t}$ and $V_{\tau,t}$ for a given τ, a time interval that can be 1 min, 1 day, or 1 week and so on. The DCC model proposed by Engle (2002) involves a two-stage estimation of the conditional covariance matrix H_t in Eq. 58.3. In the first stage, univariate volatility models are fitted for each of the returns, and estimates of $\sqrt{\sigma^2_{\tau,ii,t}}$ $(i = 1, 2, \text{ and } 3)$ are obtained by using Eq. 58.4. In the second stage, return residuals are transformed by their estimated standard deviations from the first stage. That is $\eta_{\tau,i,t} = \varepsilon_{\tau,i,t}/\sqrt{\sigma^2_{\tau,ii,t}}$, where $\eta_{\tau,i,t}$ is used to estimate the parameters of the conditional correlation. The evolution of the correlation in the DCC model is given by Eq. 58.5:

$$\sigma^2_{\tau,ii,t} = c_{\tau,i} + \alpha_{\tau,i}\varepsilon^2_{\tau,i,t-1} + \beta_{\tau,i}\sigma^2_{\tau,ii,t-1}, \quad i = 1, 2 \tag{58.4}$$

$$Q_{\tau,t} = \left(1 - \alpha_{\tau,i} - \beta_{\tau,i}\right)\overline{Q}_\tau + \alpha_{\tau,i}\, \eta_{\tau,i,t-1}\eta'_{\tau,i,t-1} + \beta_{\tau,i} Q_{\tau,t-1}, \tag{58.5}$$

where $Q_{\tau,t} = (q_{\tau,ij,t})$ is the 2 × 2 time-varying covariance matrix of $\eta_{\tau,i,t}, \overline{Q}_\tau = E\left[\eta_{\tau,i,t}\eta'_{\tau,i,t}\right]$ is the 2 × 2 unconditional variance matrix of $\eta_{\tau,i,t}$, and $a_{\tau,i}$ and $\beta_{\tau,i}$ are non-negative scalar parameters satisfying $(\alpha_{\tau,i} + \beta_{\tau,i}) < 1$. Since Q_t does not generally have ones on its diagonal, we scale it to obtain a proper correlation matrix $V_{\tau,t}$ Thus,

$$V_{\tau,t} = \left(diag\left(Q_{\tau,t}\right)\right)^{-1/2} Q_{\tau,t}\left(diag\left(Q_{\tau,t}\right)\right)^{-1/2}, \tag{58.6}$$

where $\left(diag\left(Q_{\tau,t}\right)\right)^{-1/2} = diag\left(1/\sqrt{q_{\tau,11,t}},\quad 1/\sqrt{q_{\tau,22,t}}\right)$.

Here $V_{\tau,t}$ in Eq. 58.6 is a correlation matrix with ones on the diagonal and off-diagonal elements of less than one in absolute value terms, as long as $Q_{\tau,t}$ is positive definite. A typical element of $V_{\tau,t}$ takes the form:

$$\rho_{\tau,12,t} = q_{\tau,12,t} / \sqrt{q_{\tau,11,t} q_{\tau,22,t}} \tag{58.7}$$

The dynamic correlation coefficient, $\rho_{\tau,12,t}$, can be obtained by using the element of $Q_{\tau,t}$ in Eq. 58.5, which is given by Eq. 58.8 below:

$$q_{\tau,ij,t} = \left(1 - a_{\tau,i} - b_{\tau,i}\right)\overline{\rho}_{\tau,ij} + a_{\tau,i}\, \eta_{\tau,i,t-1}\eta'_{\tau,j,t-1} + b_{\tau,i} q_{\tau,ij,t-1}, \tag{58.8}$$

The mean reversion requires that $(a_{\tau,i} + b_{\tau,i}) < 1$. In general terms, the essence of this concept is the assumption that both an asset's high and low prices are temporary and that the asset's price will tend to move toward the average price over time. Besides, the estimates of the dynamic correlation coefficients, $\rho_{ij,t}$, between each pair of the three markets have been specified as in Eq. 58.1.

58.4.2 Empirical Results of Dynamic Conditional Correlation

In this section, we present the estimation results of the models outlined above. The estimation results are presented in Table 58.1, which provides the dynamic correlations of returns across crude oil, gold, and the dollar/pound foreign exchange rate with each other. The estimated a and β for three markets are listed in Tables 58.3 and 58.4. The likelihood ratio does not support the rejection of the null hypothesis of the scalar dynamic conditional correlation. It can be seen that the sum of the estimated coefficients in the variance equations $(\alpha + \beta)$ is close to 1 for all of the cases, implying that the volatility appears to be highly persistent. As for the Ljung-Box Q-statistic of the serial correlation of the residuals, the results show that that the serial correlations in the error series are regarded as adequate.

Calvert et al. (2006) observed that through the dynamic conditional correlation distribution, we can more fully understand the real impacts in international markets.

Table 58.3 DCC estimates: three markets[a] (January 1, 1986 to December 31, 2007)

DCC			
α	0.0202**	β	0.9651**
	(52.3020)		(1,050.395)

[a]The t-statistic is given in parentheses
**Denotes significance at the 0.05 level

Table 58.4 Estimation results from the DCC-GARCH model[a]

	Mean equation Constant	Variance equation α	β	Persistence	Ljung-Box Q-statistic
Oil	0.0004**	0.1399**	0.8306**	0.9705	185.7203**
	(1.6549)	(21.6431)	(143.5303)		
Gold	−1.54E-06**	0.0664**	0.9272**	0.9936	61.8887**
	(−0.1658)	(15.0320)	(202.223)		
FX	−0.0001**	0.0417**	0.9479**	0.9896	32.6628
	(−1.6045)	(9.2628)	(168.722)		

[a]The persistence level of the variance is calculated as the summation of the coefficients in the variance equations ($\alpha + \beta$). The z-statistic is given in parentheses. The Ljung-Box Q-statistic tests the serial correlation of the residuals
**Denotes significance at the 0.05 level

It can also help with portfolio selection and risk management. In Fig. 58.2, which reports the results of the dynamic conditional correlation, the estimated correlation coefficients are time varying, reflecting some sort of portfolio shift between each two items.

The correlation between crude oil and gold was estimated using the DCC integrated method, and the results, shown in Fig. 58.2a, are quite interesting. The correlations are found to be generally positive around 0.2 except for mid-1990 which turns out to be highly correlated with a correlation of around 0.6605. The possible interpretation for the high correlation is due to the Iraqi invasion of Kuwait and the Gulf War. The crude oil price jumped from about $15 to over $33 per barrel during that time, so that investors channeled their money into the gold market because of their fear of inflation. This fact accords with the “flight to quality” concept, which represents the action of investors moving their capital away from riskier or more volatile assets to the ones considered to be safer and less volatile. The correlation between gold and the dollar/pound exchange rate is shown in Fig. 58.2b for the integrated DCC in the last 20 years. Whereas for most of the period the correlations were between −0.1 and −0.3, there were two notable drops, where the stock market crashed in October 1987 and in late 2002, and we also find two peaks, one in the middle of 1990 and the other in late 1998 where the gold price dropped to $252 per ounce. Fig. 58.2c shows the correlation between crude oil and

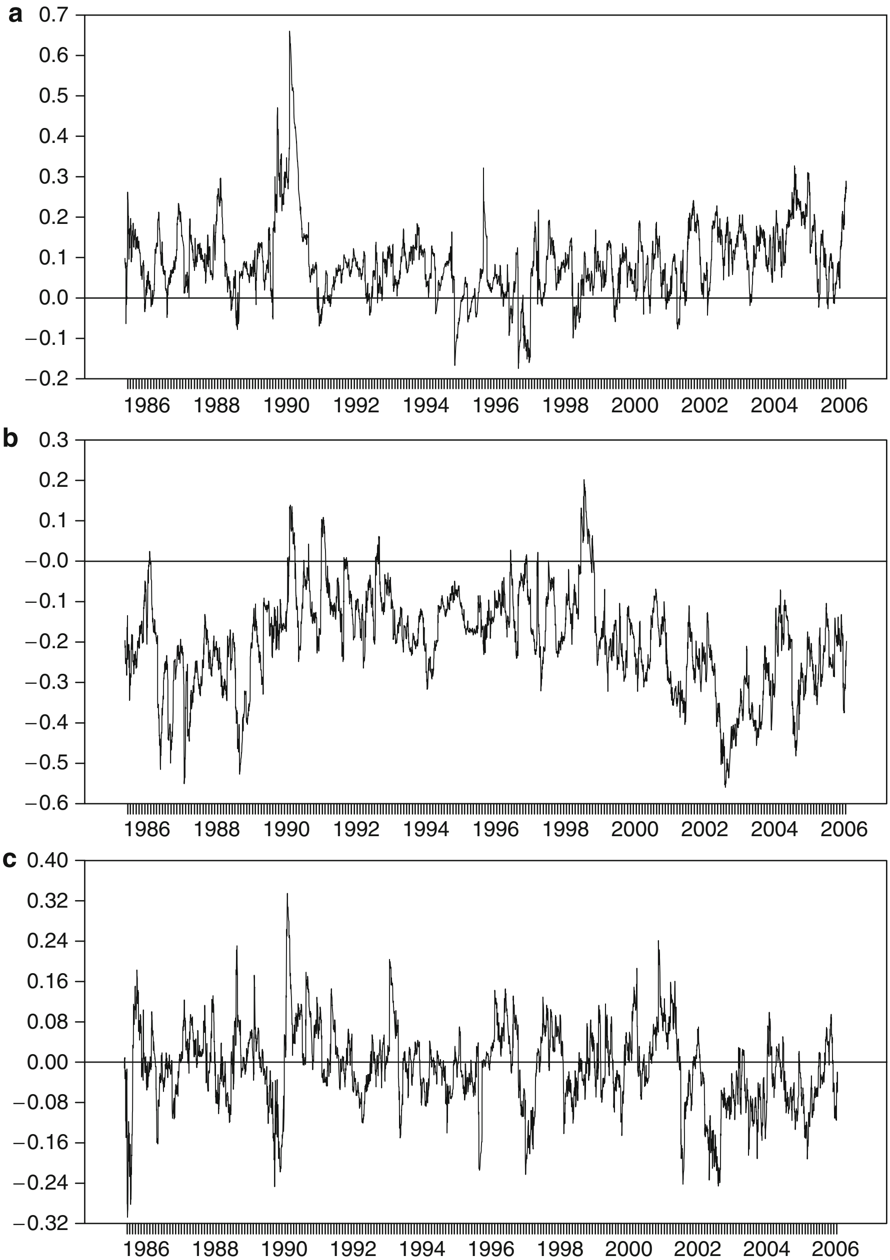


Fig. 58.2 The time series of dynamic conditional correlations (DCC) among each pair of three markets: (**a**) daily DCC between crude oil and gold, (**b**) daily DCC between gold and dollar/pound, and (**c**) daily DCC between crude oil and dollar/pound

the dollar/pound that was estimated using the DCC integrated method. Except in the mid-1990s when they are highly correlated with a coefficient of 0.32, the correlation between crude oil and the dollar/pound is generally negative (with a coefficient of −0.08 at the beginning of 1986 and a coefficient of −0.16 at the beginning of 2003, respectively).

Key issues relevant in financial economic applications include, for example, whether and how volatility and correlation move together. It is widely recognized among both finance academics and practitioners that they vary importantly over time (Andersen et al. 2001a, b; Engle 2002; Kasch and Caporin 2012). Such questions are difficult to answer using conventional volatility models, and so we wish to use the dynamic conditional correlation model to explain the phenomenon. From Fig. 58.3, the bivariate scatter plots of volatilities and correlations, it is hard to tell if there is a strong positive association between each of the two markets sampled. The correlation between two financial assets will affect the diversification of the portfolio. If two financial assets are highly negatively correlated, the effect of diversification will be significant, meaning that the portfolio can balance the returns. This is the so-called idea of not putting all one's eggs in the same basket.

According to the empirical data, the dynamic conditional correlation for the overall gold and dollar/pound (at −0.1986) or the overall oil and dollar/pound (at −0.0116) can moderately diversify investment risks and can thereby increase the rate of return. Investors can add commodities and their related derivatives to portfolios, in an effort to diversify away from traditional investments and assets. These results are in line with Capie et al. (2005) who found a negative relationship between the gold price and the sterling/dollar and yen/dollar foreign exchange rates and Nikos (2006) who found that the correlation between the dollar and gold is significantly negative. The conclusion we can draw from the results is that gold is by far the most relevant commodity in hedging against the US dollar. Capie et al. (2005) observed that gold has served as a hedge against fluctuations in the foreign exchange value of the dollar. Secondly, gold has become particularly relevant during times of US dollar weakness. In addition to that, the dynamic conditional correlation for the overall crude oil and gold markets is 0.0889, and a similar correlation was documented for the Brent crude oil and gold markets by Steinitz (2006).

To characterize the distributions of dynamic conditional correlation among the sampled markets, the summary statistics of the probability distributions for DCC are shown in Table 58.5, and the associated distributions of DCC for the sampled markets are shown in Fig. 58.4. We can find that the average DCC between the crude oil and gold markets is 0.0889 with a standard deviation of 0.0916. The distribution of the daily DCC between the crude oil and gold markets reflects a slightly right-skewed (at 1.2021) and leptokurtic distribution (at 4.6799), implying that a positive DCC occurs more often than a negative DCC between the crude oil and gold markets. Furthermore, the average DCC between the gold and

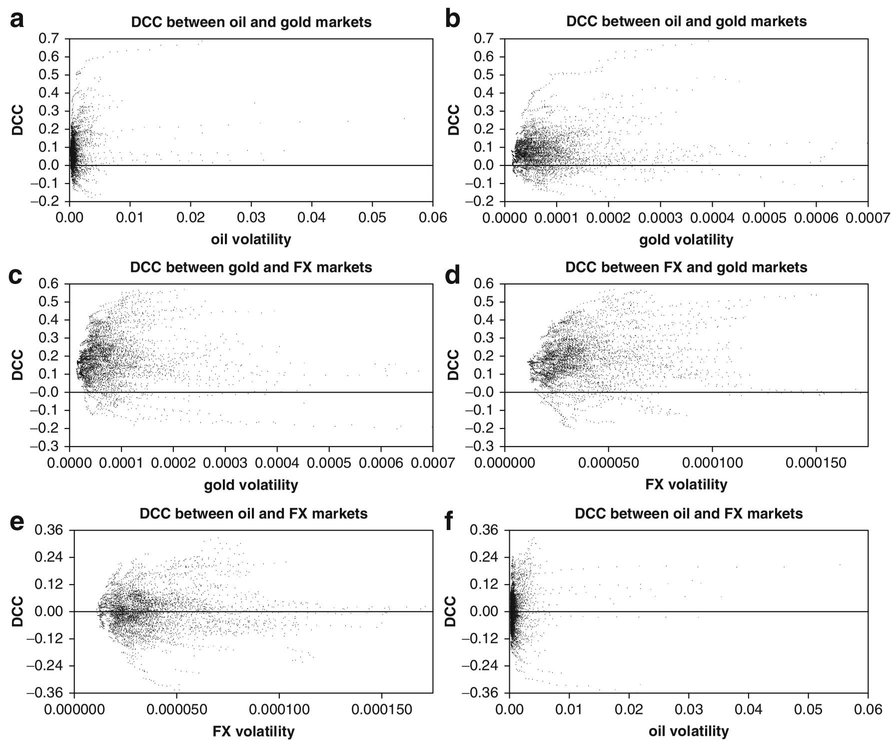


Fig. 58.3 Bivariate scatter plots of volatilities and correlations. The scatter plots of the daily DCC between (**a**) crude oil and gold and (**b**) crude oil and dollar/pound against the crude oil volatility; the daily DCC between (**c**) gold and crude oil and (**d**) gold and dollar/pound against the gold volatility; and the daily DCC between (**e**) dollar/pound and crude oil and (**f**) dollar/pound and gold against the dollar/pound volatility have been shown

Table 58.5 Summary statistics of probability distributions of DCC for each pair of oil, gold, and FX

Panel A The DCC distributions for the sampled markets from January 1, 1986 to December 31, 2007						
	Mean	Standard dev.	Max	Min	Skewness	Kurtosis
DCC between oil and gold	0.0889	0.0916	0.6605	−0.1744	1.2021	4.6799
DCC between gold and FX	−0.1986	0.1197	0.2019	−0.5596	−0.1529	0.3304
DCC between FX and oil	−0.0116	0.0799	0.3349	−0.3076	−0.0201	0.8678
Panel B The DCC distributions for the sampled markets from January 1, 1986 through July 31, 1990						
	Mean	Standard dev.	Max	Min	Skewness	Kurtosis
DCC between oil and gold	0.0910	0.0669	0.3003	−0.0783	0.3752	0.1478
DCC between gold and FX	−0.2567	0.1052	0.0246	−0.5506	−0.2751	0.0666
DCC between FX and oil	−0.0030	0.0734	0.2308	−0.3076	−0.5582	1.9547
Panel C The DCC distributions for the sampled markets from August 1, 1990 through August 31, 2001						
	Mean	Standard dev.	Max	Min	Skewness	Kurtosis
DCC between oil and gold	0.0720	0.1040	0.6605	−0.1744	1.8367	6.5149
DCC between gold and FX	−0.1258	0.0857	0.2019	−0.3222	0.7307	0.8036
DCC between FX and oil	−0.0004	0.0772	0.3349	−0.2469	0.1648	1.3984
Panel D The DCC distributions for the sampled markets from September 1, 2001 through December 31, 2007						
	Mean	Standard dev.	Max	Min	Skewness	Kurtosis
DCC between oil and gold	0.1168	0.0758	0.3270	−0.0772	−0.0521	−0.4706
DCC between gold and FX	−0.2826	0.1003	−0.0688	−0.5596	−0.2534	−0.4195
DCC between FX and oil	−0.0369	0.0832	0.2413	−0.2459	0.1549	−0.0225

dollar/pound markets is −0.1986 with a standard deviation of 0.1197. The distribution of the daily DCC between the gold and dollar/British pound markets reflects a slightly left-skewed (at −0.1529) and platykurtic distribution (at 0.3304), implying that a negative DCC occurs more often than a positive DCC between gold and the dollar/pound. Moreover, the average DCC between the dollar/pound and crude oil is −0.0116 with a standard deviation of 0.0799. The distribution of daily DCCs between the dollar/pound and crude oil markets reflects a slightly left-skewed (at −0.0201) and platykurtic distribution (at 0.8678), implying that negative DCCs occur more often than positive DCCs between the dollar/pound and crude oil markets.

According to the empirical results, we rank the sequence of the volatility in order to analyze the volatility effect in the correlation. We show the panel of volatility and dynamic conditional correlation in Tables 58.6, 58.7, and 58.8. We can clearly realize from the sample mean, from the daily DCC between the crude oil and gold markets, that higher volatility can accompany the larger DCC. However, the higher volatility can also accompany the smaller DCC between the gold and dollar/pound markets and the crude oil and dollar/pound markets. This is because the correlations between these markets are negative.

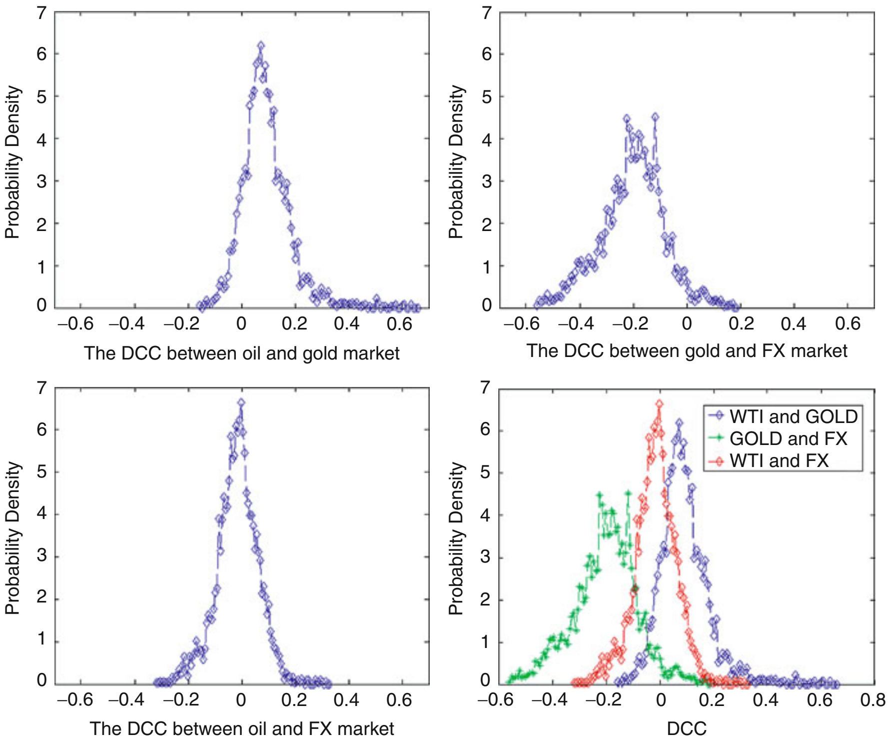


Fig. 58.4 The distributions of dynamic conditional correlations among gold, oil, and FX from 1986 to 2007

To further quantify this volatility effect in correlation, we classify the volatility into two categories, low volatility days and high volatility days,[1] and according to the results, we rank the sequence of the volatility. The group of low volatility days means that the volatility is less than the 10th percentile value and the group of high volatility days means that the volatility is greater than the 90th percentile value. The results are shown in Fig. 58.5a–c that reports the DCC distributions for low volatility days and high volatility days.

It is found that some special characteristics of DCC exist among the oil, gold, and FX markets. First, distributions for low volatility days are obviously different from those for high volatility days. Those for low volatility days approximate leptokurtic distributions, whereas those for high volatility days approximate

[1]Following Andersen et al. (2001b), the authors classify the days into two groups: low volatility days and high volatility days. The empirical results show that the distribution of correlations shifts rightward when volatility increases.

Table 58.6 DCC distributions of crude oil and gold markets

Panel A DCC between crude oil and gold against the crude oil market volatility				
	Crude oil and gold against the crude oil market volatility			
Volatility	Mean	Standard dev.	Skewness	Kurtosis
0–10 %	0.0779	0.0379	0.6205	1.8582
10–20 %	0.0785	0.0589	0.7672	0.6458
20–30 %	0.0887	0.0680	0.3950	−0.1434
30–40 %	0.0924	0.0754	0.6638	0.9140
40–50 %	0.0904	0.0769	0.4299	0.8993
50–60 %	0.0900	0.0825	0.1764	0.5715
60–70 %	0.0839	0.0820	0.3908	0.7808
70–80 %	0.0735	0.0872	0.4688	1.4680
80–90 %	0.0908	0.1273	1.2185	2.3300
90–100 %	0.1212	0.1530	0.8691	1.2595
Panel B DCC between crude oil and gold against the gold market volatility				
	Crude oil and gold against the gold market volatility			
Volatility	Mean	Standard dev.	Skewness	Kurtosis
0–10 %	0.0537	0.0516	−0.1667	−0.6070
10–20 %	0.0817	0.0567	0.0538	1.4446
20–30 %	0.0790	0.0704	0.4538	1.2475
30–40 %	0.0908	0.0734	0.7867	1.8984
40–50 %	0.0842	0.0879	0.7319	2.4974
50–60 %	0.0723	0.0918	0.2724	1.5955
60–70 %	0.0838	0.0899	0.6821	1.9649
70–80 %	0.0868	0.0880	0.9889	3.9028
80–90 %	0.1002	0.0996	0.9919	3.0296
90–100 %	0.1579	0.1386	0.9773	1.9451

platykurtic distributions. Secondly, the average DCCs of the high volatility days are greater than the average DCCs of the low volatility days for the DCCs between crude oil and gold, implying that the correlation between gold and oil increases with volatility. Furthermore, the distribution of DCCs shifts rightward when volatility increases. Similar results are found for equity returns as reported by Sonlinik et al. (1996) and in realized exchange rate returns by Andersen et al. (2001b). Thirdly, the average DCCs for the high volatility days are smaller than the average DCCs for lower volatility days across gold and the dollar/pound, and oil and the dollar/pound. This implies that the correlation between gold (oil) and foreign exchange rates decreases with volatility, and the distribution of DCCs shifts leftward when volatility increases. Finally, the standard deviations of the distributions for high volatility days are obviously greater than the standard deviations of the distributions for low volatility days.

Table 58.7 DCC distributions of gold and dollar/pound markets

Panel A DCC between gold and FX against the gold market volatility				
	Gold and FX against the gold market volatility			
Volatility	Mean	Standard dev.	Skewness	Kurtosis
0–10 %	−0.1507	0.0452	−0.2642	−0.2864
10–20 %	−0.1468	0.0811	−0.1898	0.1166
20–30 %	−0.2048	0.1069	0.0397	−0.2823
30–40 %	−0.2209	0.1039	0.3412	0.0578
40–50 %	−0.2105	0.1184	0.2681	−0.3950
50–60 %	−0.2070	0.1131	0.0924	−0.2283
60–70 %	−0.1984	0.1239	−0.0848	−0.2003
70–80 %	−0.2188	0.1411	−0.2455	−0.1963
80–90 %	−0.2231	0.1425	−0.2051	0.0326
90–100 %	−0.2064	0.1498	0.3059	0.0513
Panel B DCC between gold and FX against the FX market volatility				
	Gold and FX against the FX market volatility			
Volatility	Mean	Standard dev.	Skewness	Kurtosis
0–10 %	−0.1419	0.0621	0.1926	0.8518
10–20 %	−0.1618	0.0861	0.1325	−0.1298
20–30 %	−0.1839	0.0943	0.0866	−0.0156
30–40 %	−0.1968	0.1036	0.5927	0.7147
40–50 %	−0.2012	0.1163	0.4529	1.0610
50–60 %	−0.2314	0.1150	0.3684	0.9286
60–70 %	−0.2189	0.1239	0.2181	0.1255
70–80 %	−0.2201	0.1217	0.1057	−0.1704
80–90 %	−0.2170	0.1510	−0.2916	−0.4897
90–100 %	−0.2170	0.1558	−0.1548	−0.6496

58.4.3 Volatility Threshold Dynamic Conditional Correlation

To further check if different subperiods have various patterns, we utilize the volatility threshold model addressed by Kasch and Caporin (2012)[2] to examine whether increasing volatility (exceeding a specified threshold) is associated with an increasing correlation. The volatility threshold DCC model is specified as Eq. 58.9:

$$q_{ij,t} = \left(1 - \alpha^2 - \beta^2\right)\overline{q_{ij}} - \gamma_i\gamma_j\overline{v_{ij}} + \alpha^2\varepsilon_{i,t-1}\varepsilon_{j,t-1} + \beta^2 q_{ij,t-1} + \gamma_i\gamma_j v_{ij,t} \tag{58.9}$$

[2]Kasch and Caporin (2012) extended the multivariate GARCH dynamic conditional correlation of Engle to analyze the relationship between the volatilities and correlations. The empirical results indicated that high volatility levels significantly affect the correlations of the developed markets, while high volatility does not seem to have a direct impact on the correlations of the transition blue chip indices with the rest of the markets. It is easy to see that the volatility and correlation move together.

Table 58.8 DCC distributions of crude oil and dollar/pound markets

Panel A DCC between crude oil and FX against the FX market volatility				
	Crude oil and FX against the FX market volatility			
Volatility	Mean	Standard dev.	Skewness	Kurtosis
0–10 %	−0.0147	0.0530	−0.2458	0.3703
10–20 %	−0.0142	0.0651	−0.4224	−0.1734
20–30 %	−0.0140	0.0695	0.0757	0.2648
30–40 %	−0.0028	0.0804	0.0465	−0.2710
40–50 %	−0.0108	0.0832	0.0264	0.1784
50–60 %	−0.0206	0.0720	0.1618	0.0328
60–70 %	−0.0081	0.0789	0.5080	1.0521
70–80 %	−0.0081	0.0773	0.4689	1.6993
80–90 %	−0.0161	0.1042	0.1100	0.5576
90–100 %	−0.0068	0.1019	−0.7085	0.2311
Panel B DCC between crude oil and FX against the crude oil market volatility				
	Crude oil and gold against the gold market volatility			
Volatility	Mean	Standard dev.	Skewness	Kurtosis
0–10 %	−0.0001	0.0453	0.5049	0.7991
10–20 %	0.0045	0.0609	−0.1145	0.1341
20–30 %	−0.0016	0.0675	−0.2397	−0.2455
30–40 %	−0.0087	0.0715	−0.2668	0.1219
40–50 %	−0.0164	0.0736	−0.2498	0.1421
50–60 %	−0.0228	0.0743	−0.2262	0.0489
60–70 %	−0.0174	0.0732	0.0187	−0.2162
70–80 %	−0.0167	0.0811	−0.0200	0.0117
80–90 %	−0.0028	0.0934	0.0397	0.0698
90–100 %	−0.0310	0.1234	0.4996	0.1905

where v_t is a dummy variables matrix defined as

$$v_{ij,t} = \begin{cases} 1 \ \text{if} \ \ h_{i,t} > fh_i(k) \ \ \text{or} \ \ h_{j,t} > fh_j(k) \\ 0 \ \text{otherwise} \end{cases} \tag{58.10}$$

where $fh_i(k)$ is the kth fractional of the volatility series $h_{i,t}$.

When thresholds are found, the whole period will be divided into various sub-periods based on these thresholds. This separation helps detect any changes in investor behavior after crucial events. It is then known whether a time horizon is a key factor influencing the patterns of return and volatility.

Table 58.9 presents the estimation results of the volatility threshold DCC models. The estimation was based on various volatility threshold levels at 50 %, 75 %, 90 %, and 95 %. The results in Table 58.9 show that the correlation between the oil and gold prices is significantly affected by the volatility of oil at the 50 %, 75 %, and 90 %

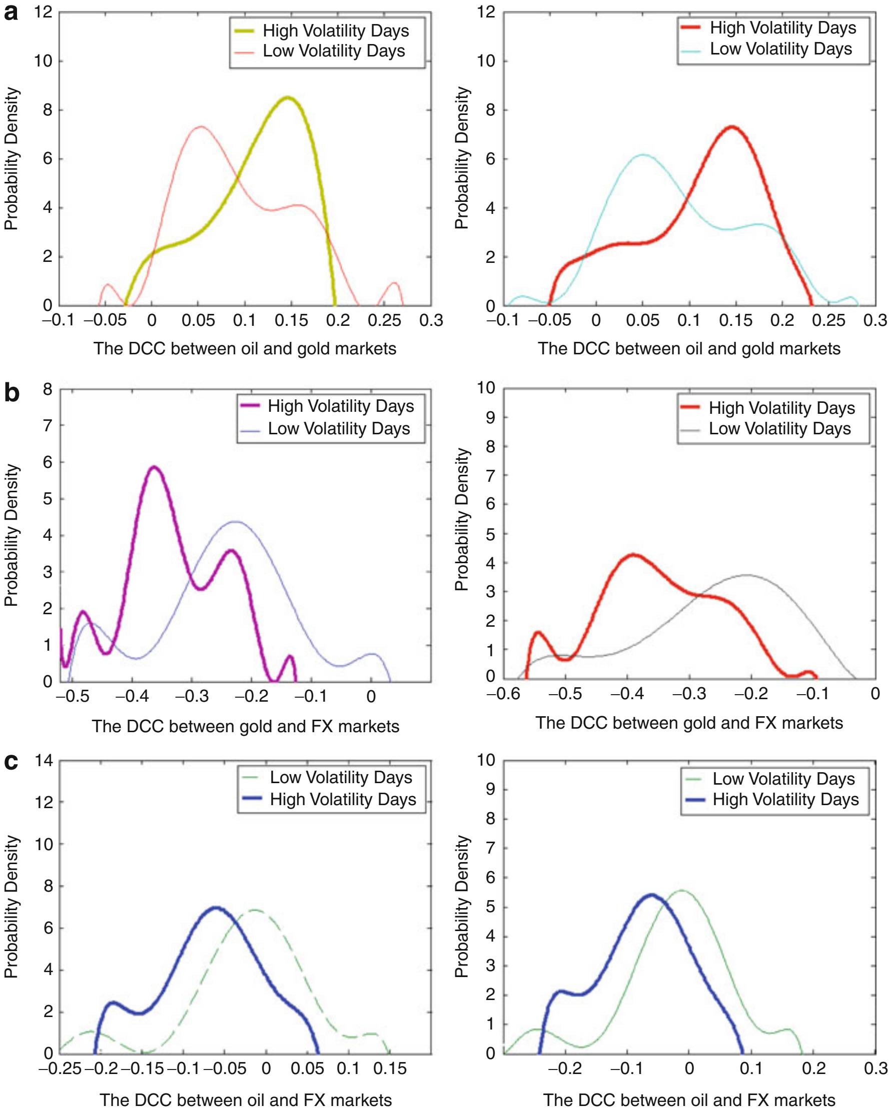


Fig. 58.5 (**a**) The distributions of DCC between crude oil and gold on low and high volatility days. (**b**) The distributions of DCC between gold and dollar/pound on low and high volatility days. (**c**) The distributions of DCC between oil and dollar/pound on low and high volatility days

(with the exception of 95 %) thresholds. Interestingly, these estimated thresholds were quite consistent with the real events of *the First Gulf War in 1990 and the 911 attack in 2001*. We then separate the period into three subperiods to further examine whether the investors' behaviors change after the events.

Table 58.9 The volatility threshold dynamic conditional correlation[a]

Panel A Crude oil and gold				
	50 %	75 %	90 %	95 %
α^2	0.0171	0.0146	0.0187	0.0229
	(1.956)	(1.870)	(2.072)	(2.424)
β^2	0.9546	0.9548	0.9362	0.9233
	(36.377)	(41.162)	(27.21)	(23.705)
$\gamma_{WTI}\gamma_{GOLD}$	0.0075	0.0116	0.0206	0.0253
	(2.045)	(2.300)	(2.376)	(1.636)
Panel B Gold and dollar/pound				
	50 %	75 %	90 %	95 %
α^2	0.0161	0.0150	0.0150	0.0149
	(3.016)	(3.052)	(2.938)	(2.894)
β^2	0.9784	0.9808	0.9804	0.9803
	(108.063)	(124.94)	(120.37)	(119.94)
$\gamma_{GOLD}\gamma_{FX}$	−0.00032	0.0012	0.0023	0.0046
	(−0.190)	(0.805)	(0.932)	(1.079)
Panel C Crude oil and dollar/pound				
	50 %	75 %	90 %	95 %
α^2	0.0131	0.0121	0.0109	0.0126
	(2.294)	(1.825)	(1.633)	(2.013)
β^2	0.9519	0.9580	0.9639	0.9518
	(33.260)	(28.286)	(33.633)	(28.095)
$\gamma_{WTI}\gamma_{FX}$	0.00006	0.0015	0.0034	0.0099
	(0.016)	(0.418)	(0.785)	(0.991)

[a]This table presents the quasi-maximum likelihood estimates of volatility threshold of dynamic conditional correlation. The t-statistics are given in parentheses

58.4.4 Does Investors' Behavior Change over Subperiods?

The three subperiods based on our estimated thresholds are before the first Gulf War (January 1, 1986 to July 31, 1990), after the first Gulf War up to the 911 attack (August 1, 1990 to August 31, 2001), and after the 911 attack (September 1, 2001 to December 1, 2007). We then examine the dynamic co-movement in each pair of markets in various subperiods (Rigobon and Sack 2005; Guidi et al. 2007[3]).

Our sampled period covers economic hardship and soaring energy prices. Soaring energy prices make gold an attractive hedging asset against inflation in that a positive correlation with oil is expected over time, especially after the 911 event. The evidence in Table 58.10 shows that oil and gold are highly

[3]Guidi et al. (2007) examined the impact of relevant US decisions on oil spot price movements from January 1986 to December 2005. They identified the following conflict periods: the Iran-Iraq conflict, January 1985 until July 1988; Iraq's invasion of Kuwait, August 1990 until February 1991; and the US-led forces' invasion of Iraq, March 2003 until December 2005.

Table 58.10 Simple correlation matrix of oil, gold, and dollar/pound markets among the three subperiods

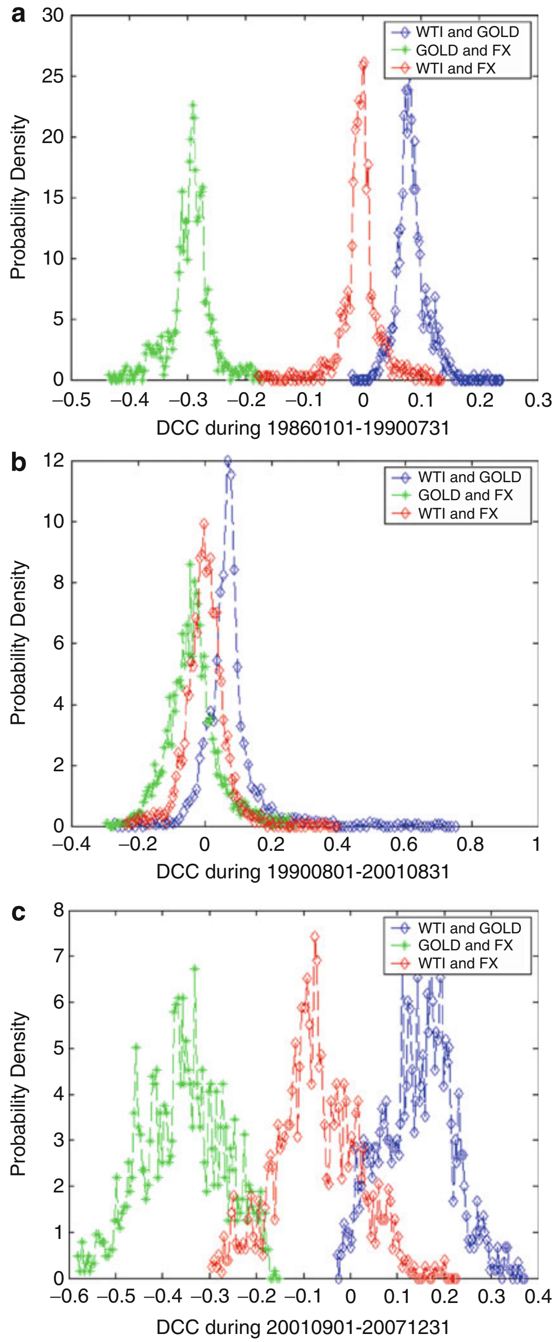
	Oil	Gold	FX
Panel A The correlation coefficients between crude oil, gold, and dollar/pound markets from January 1, 1986 through July 31, 1990			
	Oil	Gold	FX
Oil	1		
Gold	0.1648	1	
FX	−0.1517	−0.5228	1
Panel B The correlation coefficients between crude oil, gold, and dollar/pound markets from August 1, 1990 through August 31, 2001			
	Oil	Gold	FX
Oil	1		
Gold	−0.2465	1	
FX	0.0628	−0.2096	1
Panel C The correlation coefficients between crude oil, gold, and dollar/pound markets from September 1, 2001 through December 31, 2007			
	Oil	Gold	FX
Oil	1		
Gold	0.9264	1	
FX	−0.8124	−0.8142	1

correlated with a high coefficient of 0.9264 between September 1, 2001 and December 31, 2007 (after the 911 event). In the meantime, the correlation between gold and the dollar/pound is −0.8142, and between oil and the dollar/pound is −0.8124, showing the fears of a depreciation in the US dollar push commodity prices up significantly.

The results of Panels B, C, and D in Table 58.5 along with Figs. 58.6 and 58.7 show that the DCCs between oil and gold are all positive in the three subperiods (with coefficients of 0.0910, 0.0720, and 0.1168, respectively). Obviously, higher oil prices spark inflationary concerns and make gold a value reserve for wealth. By contrast, the correlation between oil and the dollar/pound has been negative, and a DCC of −0.0369 is shown after the 911 attack. In this period, the oil price increased from a low of US$19 per barrel in late January 2002 to US$96 per barrel in late December 2007. Meanwhile, the dollar conversely tumbled 29 % from US$0.7 to a US$0.5 per pound.

Historical data also show that the prices of oil and gold are rising over time. The increase in the prices of gold was a reflection of the falling US dollar. The evidence confirms that the DCCs between gold and the dollar/pound in the three subperiods are all negative (with average DCC of −0.2567, −0.1258, and −0.2826 in the first, second, and third periods, respectively). Generally speaking, during the subperiods of market crises, increasingly high correlations in commodity prices were observed, with oil and gold moving in the same direction and the dollar/pound moving in opposite directions.

Fig. 58.6 The distributions of dynamic conditional correlations among each pair of the three markets over three subperiods: (**a**) January 1, 1986 to July 31, 1990, (**b**) August 1, 1990 to August 31, 2001, and (**c**) September 1, 2001 to December 31, 2007



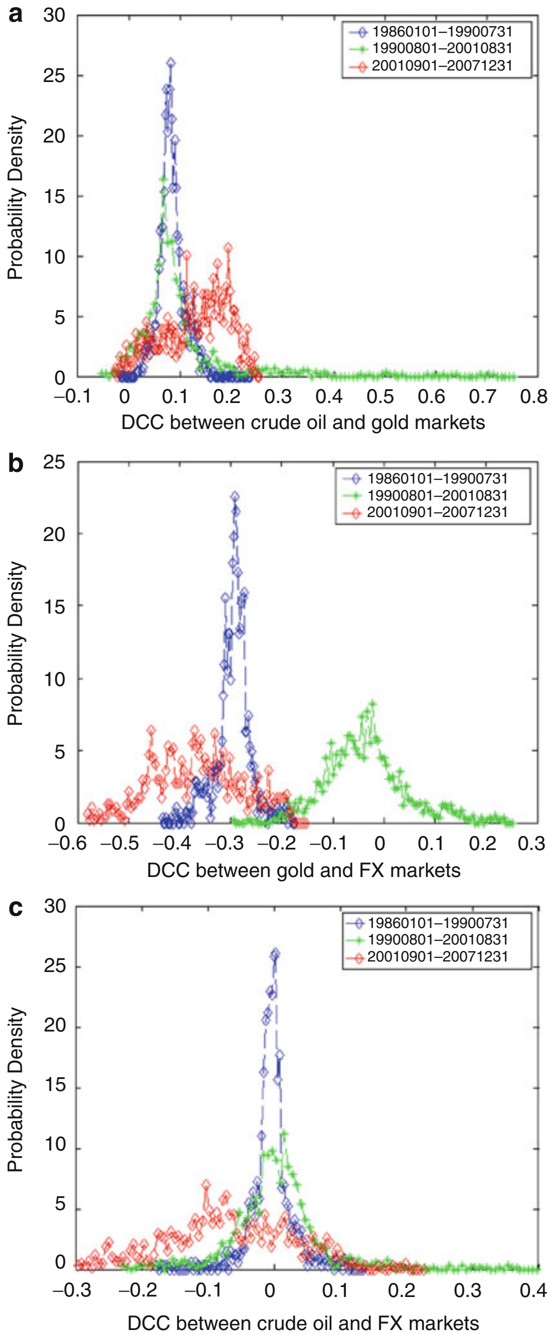


Fig. 58.7 The distributions of dynamic conditional correlations for (**a**) crude oil and gold, (**b**) gold and dollar/pound, (**c**) crude oil and dollar/pound over three subperiods

58.5 Conclusions and Implications

Using the dynamic conditional correlation model, we have estimated the cross correlation and volatility among crude oil, gold, and dollar/pound currencies from 1986 to 2007. After exploring the time-varying correlations and realized distributions, several regularities have been found that help illustrate the characteristics of the crude oil, gold, and currency markets.

First, the correlation coefficients between each pair of the three assets are found to be time varying instead of constant. As such, besides considering the mean and standard deviation of the underlying assets, investors need to follow the co-movement in the relevant assets in order to make better portfolio hedging and risk management decisions across these assets. The results of the dynamic correlation coefficients between the gold and dollar/pound show that gold is by far the most relevant commodity in terms of serving as a hedge against the US dollar. These results are in line with the reports suggested by Nikos (2006) and Capie et al. (2005). Our findings are helpful in terms of arriving at a more optimal allocation of assets based on their multivariate returns and associated risks.

Besides, the distributions of low volatility days are found to approximate leptokurtic distributions in the gold, oil, and dollar/pound markets, whereas the high volatility days approximate platykurtic distributions. Furthermore, the DCC between oil and gold is increasing with volatility, indicating that the distribution of DCC shifts rightward when volatility increases. By contrast, the DCCs between gold and the dollar/pound and crude oil and the dollar/pound are decreasing with the volatility. Our findings in terms of oil and gold are consistent with the reports of Sonlinik et al. (1996) and Andersen et al. (2001b) who use different approaches.

Moreover, by estimating the volatility threshold dynamic conditional correlation model addressed by Kasch and Caporin (2012), we find that high volatility values (exceeding some specified thresholds) are associated with an increase in correlation values in various subperiods. Remarkably, investors' behaviors are seen to have changed in different subperiods. During periods of market turmoil, such as the First Gulf War in 1990 and the 911 terror attack in 2001, an increase in correlation between the prices of oil and gold, as well as a decrease in correlation between the oil (gold) and dollar/pound currencies, is observed. These behaviors make gold an attractive asset against major currencies for value-preserving purposes. For market participants from long-term hedging perspective, our results provide useful information on asset allocation across commodity and currency markets during market turmoil.

Acknowledgments This chapter was presented at the Seventh International Business Research Conference in Sydney, Australia, and the Taiwan Finance Association Annual Meeting in Taichung, Taiwan. The authors are grateful for the helpful and suggestive comments from Ken Johnson, C. L. Chiu, Ming-Chi Lee, and other participants. The authors also appreciate the financial grants for attending the conference from the National Science Council.

References

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Ebens, H. (2001a). The distributions of realized stock return volatility. *Journal of Financial Economics, 61*, 43–76.

Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2001b). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association, 96*(453), 42–55.

Andersen, T., Bollerslev, T., Diebold, F., & Vega, C. (2007). Real time price discovery in global stock, bond and foreign exchange markets. *Journal of International Economics, 73*, 251–277.

Baur, D. G., & Lucey, B. M. (2010). Is gold a hedge or a safe haven? An analysis of stocks, bonds and gold. *Financial Review, 45*(2), 217–229.

Baur, D. G., & McDermott, T. K. (2010). Is gold a safe haven? International evidence. *Journal of Banking and Finance, 34*(8), 1886–1898.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics, 31*, 307–327.

Bollerslev, T. (1990). Modeling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH. *Review of Economics and Statistics, 72*(3), 498–505.

Boyer, M. M., & Fillion, D. (2007). Common and fundamental factors in stock returns of Canadian oil and gas companies. *Energy Economics, 29*, 428–453.

Calvert, L. E., Fischer, A. J., & Thompson, S. B. (2006). Volatility comovement: A multi-frequency approach. *Journal of Econometrics, 131*(1–2), 179–215.

Capie, F., Mills, T. C., & Wood, G. (2005). Gold as a hedge against the dollar. *Journal of International Financial Markets, Institutions and Money, 15*(4), 343–352.

Chiang, T. C., Jeon, B. N., & Li, H. M. (2007a). Dynamic correlation analysis of financial contagion: Evidence from Asian markets. *Journal of International Money and Finance, 26*(7), 1206–1228.

Chiang, T. C., Tan, L., & Li, H. M. (2007b). Empirical analysis of dynamic correlations of stock returns: Evidence from Chinese A-share and B-share markets. *Quantitative Finance, 7*(6), 651–667.

Chiang, T. C., Yu, H. C., & Wu, M. C. (2009). Statistical properties, dynamic conditional correlation, scaling analysis of high-frequency intraday stock returns: Evidence from Dow-Jones and nasdaq indices. *Physica A, 388*(8), 1555–1570.

Doong, S. C., Yang, S. Y., & Wang, A. T. (2005). The dynamic relationship and pricing of stocks and exchange rates: Empirical evidence from Asian emerging markets. *Journal of the American Academy of Business, 7*(1), 118–123.

Engle, R. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation. *Econometrica, 50*, 987–1008.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate GARCH models. *Journal of Business & Economic Statistics, 20*(3), 339–350.

Engle, R., Ito, T., & Lin, W. L. (1994). Do bulls and bears move across borders? International transmission of stock returns and volatility. *Review of Financial Studies, 7*, 507–538.

Guidi, M., Russell, A., & Tarbert, H. (2007). The efficiency of international oil markets in incorporating US announcements during conflict and non-conflict periods. *Petroleum Accounting and Financial Management Journal, 26*(2), 67–86.

Kasch, M., & Caporin, M. (2012). Volatility threshold dynamic conditional correlations: An international analysis. SSRN: http://ssrn.com/abstract=968233

Kaul, A., & Sapp, S. (2006). Y2K fears and safe haven trading of the US dollar. *Journal of International Money and Finance, 25*, 760–779.

Lanza, A., Manera, M., & McAleer, M. (2006). Modeling dynamic conditional correlations in WTI oil forward and futures returns. *Finance Research Letters, 3*, 114–132.

Ng, A. (2000). Volatility spillover effects from Japan and the US to the Pacific-Basin. *Journal of International Money and Finance, 19*(2), 207–233.

Nikos, K. (2006). Commodity prices and the influence of the US dollar. World Gold Council.

Pérez-Rodríguez, J. (2006). The Euro and other major currencies floating against the U.S. dollar. *Atlantic Economic Journal, 34*, 367–384.

Ranaldo, A., & Soderlind, P. (2009). Safe haven currencies. *Review of Finance, 14*(3), 385–407.

Rigobon, R., & Sack, B. (2005). The effect of war on US financial markets. *Journal of Banking and Finance, 29*(7), 1769–1789.

Sonlinik, B., Boucrelle, C., & Le Fur, Y. (1996). International market correlation and volatility. *Financial Analysts Journal, 52*(5), 17–34.

Steinitz, J. (2006). Gold: Volatility, correlation and teacups. World Gold Council.

Tastan, H. (2006). Estimating time-varying conditional correlations between stock and foreign exchange markets. *Physica A, 360*, 445–458.

Upper, C. (2000). How safe was the "safe-haven"? Financial market liquidity during the 1998 turbulences. Discussion paper 1/00, Deutsche Bundesbank.

Pre-IT Policy, Post-IT Policy, and the Real Sphere in Turkey

59

Ahmed Hachicha and Cheng-Few Lee

Contents

A. Hachicha
Department of Economic Development, Faculty of Economics and Management of Sfax, University of Sfax, Sfax, Tunisia
e-mail: hachicha.ahmed@fsegs.rnu.tn

C.-F. Lee (✉)
Department of Finance and Economics, Rutgers Business School, Rutgers, The State University of New Jersey, Piscataway, NJ, USA

Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
e-mail: cflee@business.rutgers.edu; lee@business.rutgers.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_59,

Abstract

We estimate two SVECM (structural vector error correction) models for the Turkish economy based on imposing short-run and long-run restrictions that account for examining the behavior of the real sphere in the pre-IT policy (before inflation-targeting adoption) and post-IT policy (after inflation-targeting adoption).

Responses reveal that an expansionary interest policy shock leads to a decrease in price level, a fall in output, an appreciation in the exchange rate, and an improvement in the share prices in the very short run for the most of pre-IT period.

Central Bank of the Republic of Turkey (CBT) stabilizes output fluctuations in the short run while maintaining a very medium-run inflation target since January 2006. One of the most important results of this study is that the impact of a monetary policy shock on the real sphere is insignificant during the post-IT policy.

Keywords

SVECM models • Turkish economy • Pre-IT policy • Post-IT policy • Real sphere

59.1 Introduction

The effectiveness of inflation targeting on the real sphere has recently been the subject of a vast and ever-growing literature. Being well defined in theory, we can say that there is a lack especially in the identification of the repercussion of inflation-targeting framework on macroeconomic variables. This is one of the most interesting subjects that merit to be watched nowadays. This chapter will be a new contribution to the empirical literature.

We are going to present the eligible method that allows us to discover the relation between the inflation targeting and the real sphere in Turkey.

Due to the successful experience in some neighboring countries with the adopting of inflation-targeting regime, the Turkish economy was encouraged to adopt such a monetary policy to overcome one of the deepest crises of Turkish economy in 2001 especially when the Central Bank of the Republic of Turkey was obliged to observe floating its currency.

In this study, we analyze the following research questions: “How can the real sphere react to a pre-inflation-targeting regime?” “Does inflation targeting enhance output growth?” “Are we going to find similar results regarding the effectiveness of pre-IT policy and IT policy?”

To deal with our objective, we investigate a structural vector error correction model (SVECM) analysis with long-run and short-run restrictions for IT policy and pre-IT policy to extract conclusions through examining the responses of macroeconomic data, respectively, to a monetary policy shock. Monetary transmission mechanisms based on a structural vector error correction model were studied by King et al. (1991), Ehrmann (1998), Lütkepohl et al. (1998), Ramaswamy and Sloek (1998), Cecchetti (1995), Debondt (2000), Clements et al. (2001), Kakes and Sturm (2002), Nadenichek (2006), Hachicha and Chaabane (2007), Ivrendi and Guloglu (2010), Lucke (2010), and Bhuiyan (2012).

The structure of this chapter is as follows: Sect. 59.2 provides the empirical methodology and outlines the SVECM technique. Section 59.3 defines the data. Section 59.4 presents long-run and short-run matrices estimation of the SVECM technique before and after IT adoption. Section 59.5 highlights a comparative analysis for the empirical results studying the impact of an interest rate shock on the output, inflation rate, share prices, and exchange rates in the pre-IT and post-IT policy. Section 59.6 makes some concluding remarks.

59.2 Empirical Methodology

Searching for an answer to my problematic, we resort to a structural vector error correction model (SVECM) technique with contemporaneous and long-run restrictions developed by Breitung et al. (2004). The main advantage of adopting a structural vector error correction model instead of a structural vector autoregressive model is that it gives us the opportunity to use cointegration restrictions which implement constraints on the long-run effects of the permanent shocks (Lutkepohl 2005). In what follows, we explore the SVEC model and forecasting technique.

59.2.1 SVEC Model and Forecasting Technique

$$\Delta y_t = \alpha\beta^{*\prime}\begin{pmatrix} y_{t-1} \\ D^{\infty}_{t-1} \end{pmatrix} + \Gamma_1\Delta y_{t-1} + \ldots + \Gamma_p\Delta y_{t-p} + CD_t + u_t \tag{59.1}$$

where $y_t = (y_{1t}, \ldots y_{kt})\,'$ is a vector of K endogenous variables, D^{∞}_{t-1} contains all deterministic terms included in the cointegration relations, and D_t contains all remaining deterministic variables (constant, seasonal dummy). The residual vector u_t is assumed to be a K-dimensional process and unobservable zero means white noise process with positive definite covariance matrix $E(u\,u\,') = \Sigma_u$.

The parameter matrices α and β have dimensions (K × r) and they have to have rank r. They specify the long-run part of the model with β containing the cointegrating relations and α representing the loading coefficients. The column dimension of η is also r and its row dimension corresponds to the dimension of D^{∞}_{t-1}. The notation $\beta^* = \begin{pmatrix} \beta \\ \eta \end{pmatrix}$ will be used in the following and the row dimension of β^* will be denoted by K^*. Hence, β^* is a (K × r) matrix. The cointegrating rank has to be in the range $1 \leq r \leq k - 1$.

SVEC (structural vector correction) is a model that can identify the shocks to be traced in an impulse response analysis by imposing restrictions on the matrix of long-run effects of shocks and the matrix B of contemporaneous effects of the shocks.

59.2.2 Matrix B Definition

According to the theorem of Johansen (1995), the VEC model has the following moving average representation:

$$y_t = \Xi \sum_{i=1}^{t} u_i + \Xi^*(L)u_t + y_0^* \tag{59.2}$$

where $y_t = (y_{1t}, \ldots y_{kt})'$ is a vector of K observable series and y_0^* contains the initial values. The long-run effects of shocks are represented by the first term in Eq. 59.2, $\Xi \sum_{i=1}^{t} u_i$, which captures the common stochastic trends from time (1) till time (t). The matrix B is defined such that $u_t = B\varepsilon_t$, and assuming that it is in reduced form, the matrix of long-run effects of the u_t residuals is

$$\Xi = \beta\perp \left(\alpha\perp' \left(I_k - \sum_{i=1}^{p-1} \Gamma_i \right) \beta\perp \right)^{-1} \alpha'\perp \tag{59.3}$$

Hence the long-run effects of ε shocks are given by ΞB. $rk(\Xi) = K - r$ and it follows that ΞB has rank $K - r$. Thus the matrix ΞB can have at most r columns of zero.

On that account, there can be at most r shock with transitory effect andis at least $(k - r)$ shocks have permanent effects. Due to the reduced rank of the matrix, each column of zeros stands for only $(k - r)$ independent restrictions. $(K - r)(K - r - 1)/2$ additional restrictions are needed to exactly identify the permanent shocks and $r(r - 1)/2$ additional contemporaneous restrictions identify the transitory shocks.

59.2.3 The Confidence Interval

The impulse responses are computed from the estimated VAR coefficients, and the Hall percentile interval is chosen to build confidence intervals (CI) that reflect the estimation's unpredictability:

$$CI = \left[\phi_1 - t^*_{(1-\gamma/2)}, \phi_2 - t^*_{(\gamma/2)} \right] \tag{59.4}$$

According to Hall (1992), $t^*_{\gamma/2}$ and t^* $(1 - \gamma/2)$ are the $\gamma/2$ and the $(1 - \gamma/2)$ quantiles of the distribution of $CI = \langle \phi_1 - \phi_2 \rangle$, respectively.

Table 59.1 Forecasting 1st undifferenced series (pre-IT period)

Reference: Lütkepohl (1993), IMTSA, 2ed, ch. 5.2.6, ch. 10.5					
CI coverage: 0.95					
Forecast horizon: 1 period					
Using standard confidence intervals					
y(N) in levels used in the forecast					
Time	MMR_log-level (N)	ExRate_log-level (N)	Share_prices_log-level (N)	IP_log-level (N)	CPI_log-level (N)
2005 M12	0.0000	0.0000	0.0000	0.0000	0.0000
ExRate_log					
Time	Forecast	Lower CI	Upper CI	+/−	
2006 M1	0.3098	0.1963	0.4234	0.1135	
Share_prices_log					
Time	Forecast	Lower CI	Upper CI	+/−	
2006 M1	4.8843	4.6852	5.0834	0.1991	
IP_log					
Time	Forecast	Lower CI	Upper CI	+/−	
2006 M1	4.7326	4.5883	4.8769	0.1443	
CPI_log					
Time	Forecast	Lower CI	Upper CI	+/−	
2006 M1	4.6779	4.6504	4.7055	0.0275	

59.2.4 Forecasting with SVEC Processes

According to Pfaff (2008), forecasting is based on h-step at time T:

$$y_{T+h/T} = A_1 y_{T+h-1/T} + \ldots + A_p y_{T+h-p/T} + B_0 x_{T+h} + \ldots + B_q x_{T+h-q} + CD_{T+h} \tag{59.5}$$

The forecasts are computed recursively for h $\geq$ 1 of an empirical VECM (p) process according to

$$y_{T+h/T} = A_1 y_T + \ldots + A_p y_{T+h-p} + B_0 x_{T+h} + \ldots + B_q x_{T+1-q} + CD_{T+h} \tag{59.6}$$

The forecasting errors are

$$y_{T+h} - y_{T+h/T} = u_{T+h} + \Phi_1 u_{T+h-1} + \ldots + \phi_{h-1} u_{T+1}, \tag{59.7}$$

With $\Phi_0 = I_k$ and Φ_s can be computed recursively according to

$$\Phi_s = \sum_{i=1}^{s} \Phi_{s-j} A_j, \quad s = 1, 2 \ldots \tag{59.8}$$

According to Lütkepohl (1991), $\Phi_0 = I_k$ and $A_j = 0$ for $j \rangle p$. Appendix 1 explores

Table 59.2 Forecasting 1st undifferenced series (post-IT period)

Reference: Lütkepohl (1993), IMTSA, 2ed, ch. 5.2.6, ch. 10.5					
CI coverage: 0.95					
Forecast horizon: 1 period					
Using standard confidence intervals					
y(N) in levels used in the forecast					
Time	MMR_log-level (N)	ExRate_log-level (N)	Share_prices_log-level (N)	IP_log-level (N)	CPI_log-level (N)
2005 M12	0.0000	0.0000	0.0000	0.0000	0.0000
MMR_log					
Time	Forecast	Lower CI	Upper CI	+/−	
2011 M12	1.6238	1.2610	1.9866	0.3628	
ExRate_log					
Time	Forecast	Lower CI	Upper CI	+/−	
2011 M12	0.6154	0.5210	0.7099	0.0944	
Share_prices_log					
Time	Forecast	Lower CI	Upper CI	+/−	
2011 M12	5.2127	5.0617	5.3637	0.1510	
IP_log					
Time	Forecast	Lower CI	Upper CI	+/−	
2011 M12	4.8102	4.6942	4.9262	0.1160	
Consumerprices_log					
Time	Forecast	Lower CI	Upper CI	+/−	
2011 M12	5.1506	5.1336	5.1676	0.0170	

the forecast error covariance matrix, presents forecasting series estimations (Tables 59.1 and 59.2), and plots figures (Figs. 59.1 and 59.2).

59.3 Data Definition

The dataset consists of monthly observations from January 2000 up to November 2011 divided into two periods, i.e., the inflation-targeting framework was adopted in Turkey in January 2006. The first period named "**Pre-IT**" treats the case before the adoption of inflation-targeting policy and starts from 2000 M1.

The second period called "**Post-IT**" takes into consideration the adoption of inflation-targeting policy and starts from 2006 M1. This division highlights the importance of the topic treated in our paper and does not affect our estimation results. We choose deliberately monthly frequency to maximize the number of observations to get a robust estimation of each period.

The empirical models are estimated separately for each period. The interest rate (R) is measured by the log of money market rate (MMR); the price level (P) is measured by the log of consumer prices (2005=100); the real output (Y) is measured by the log of industrial production index (2005=100);

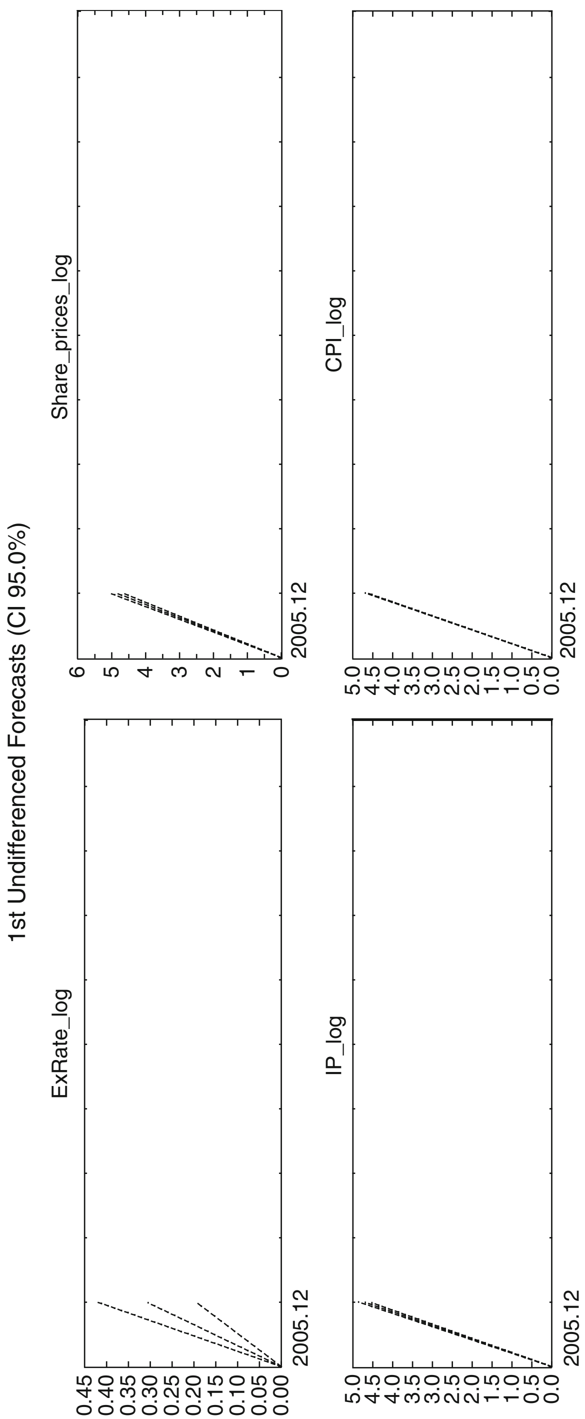


Fig. 59.1 Time series forecasts (2000 M1–2005 M12)

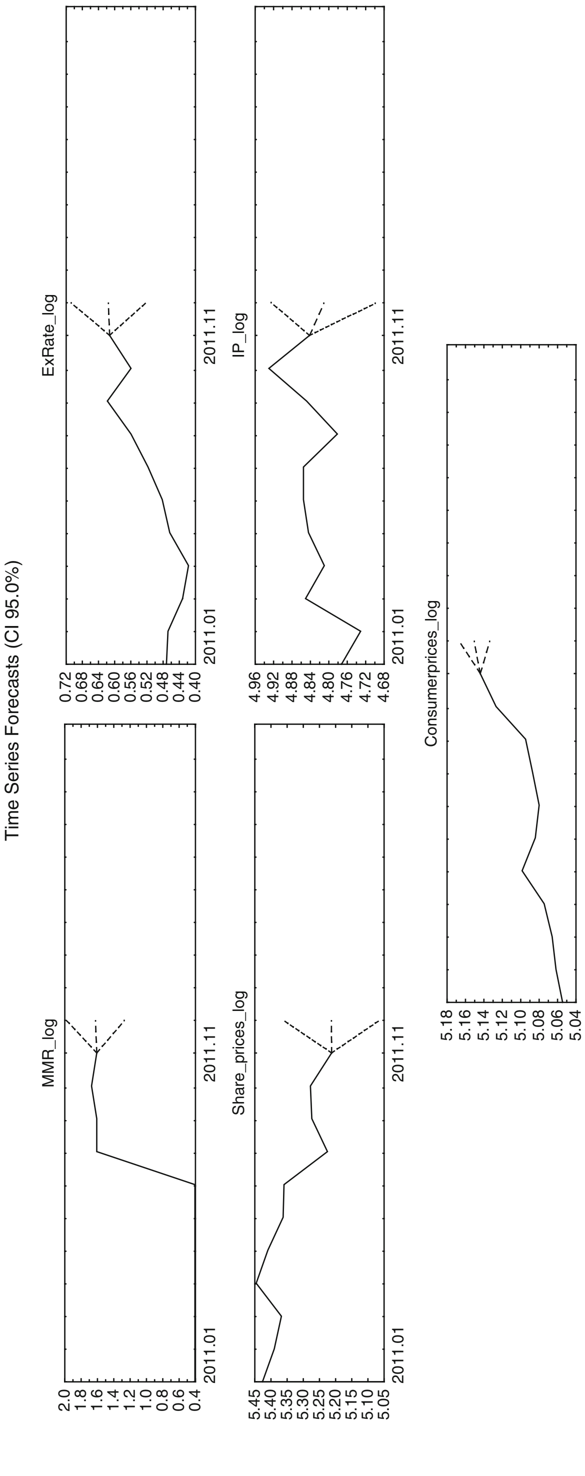


Fig. 59.2 Time series forecasts (2006 M1–2011 M11)

the real exchange rate (REXR) is measured by the log of (EXR). Due to lack of adequate data for the Tobin's Q, the share prices index (share prices) is taken as a proxy for wealth channel, and it is measured by the log of share prices.

The main source of data is IMF's International Financial Statistics and the Central Bank of the Republic of Turkey.

59.4 Empirical Analysis

We start our empirical analysis by investigating the univariate time series properties of the variables. Secondly, we use the AIC information criteria to determine the lag length of the VECM process. It suggests a lag length of $P = 11$ when maximum lag length is $pmax = 11$. This lag length is also confirmed with the same information criteria for $p = 12$ (Lütkepohl 1991). Then, we test the number of cointegration relation for each system before and after IT adoption, separately. We use Johansen's (1988, 1995) approach to test for the existence of a cointegrating relationship among the variables.

The maximum eigenvalue (λmax) and the trace tests for each model suggest one or two cointegration relations among five variables. In Appendix 2, we show results of Johansen cointegration test for the two systems. Specifically in Tables 59.3 and 59.4, we report cointegration results of the estimated parameters for the whole SVECM systems.

The structural shocks being identified, the VECM model is transformed into a VMA model (moving average) which makes it possible to compute the dynamics of the various endogenous variables following a structural shock due to

Table 59.3 Cointegration test (pre-IT)

Sample range: [2000 M3, 2005 M12], $T = 70$					
Johansen trace test for: MMR_log ExRate_log Share_prices_log IP_log Consumerprices_log					
Included lags (levels): 2					
Dimension of the process: 5					
Trend and intercept included					
Response surface computed:					
r_0	LR	pval	90 %	95 %	99 %
0	133.83	0.0000	84.27	88.55	96.97
1	86.49	0.0001	60.00	63.66	70.91
2	50.90	0.0055	39.73	42.77	48.87
3	24.33	0.0754	23.32	25.73	30.67
4	10.55	0.1051	10.68	12.45	16.22
Optimal endogenous lags from information criteria					
Sample range: [2000 M11, 2005 M12], $T = 62$					
Optimal number of lags (searched up to 10 lags of levels)					
Akaike info criterion: 10					
Final prediction error: 10					
Hannan-Quinn criterion: 10					
Schwarz criterion: 1					

Table 59.4 Cointegration test (post-IT)

Sample range: [2006 M3, 2011 M11], $T = 69$					
Johansen trace test for: MMR_log ExRate_log Share_prices_log IP_log Consumerprices_log					
r_0	LR	pval	90 %	95 %	99 %
0	79.82	0.0284	72.74	76.81	84.84
1	46.76	0.1928	50.50	53.94	60.81
2	27.58	0.2642	32.25	35.07	40.78
3	15.03	0.2293	17.98	20.16	24.69
4	4.83	0.3131	7.60	9.14	12.53
Optimal endogenous lags from information criteria					
Sample range: [2006 M11, 2011 M11], $T = 61$					
Optimal number of lags (searched up to 10 lags of levels)					
Akaike info criterion: 10					
Final prediction error: 10					
Hannan-Quinn criterion: 10					
Schwarz criterion: 1					

the Johansen procedure which supposes the ignorance of restrictions on beta. Then, we introduce restrictions in matrices of the long and short run. In addition, we know from the common trends literature that in a five-dimensional system with two cointegration relations determined previously, only three shocks can have permanent effect. We impose over identifying restrictions on the cointegrating vectors using the ML method proposed by Johansen (1995).

The identified cointegration relations can be used to set up a full VECM, where no further restrictions are imposed to form an estimate for Σ_u. Moreover, long-run and contemporaneous identifying restrictions derived from the theory are used to form estimates for matrix B or A. We know from the previous results that we need K (K − 1)/2 = 5(5 − 1)/2 = 10 additional linearly independent restrictions coming from economic theory to exactly identify the structural shocks.

$$\mathrm{A} = \mathrm{C}(1)\mathrm{B} = \begin{pmatrix} * & * & * & * & * \\ 0 & 0 & * & * & * \\ * & * & * & * & * \\ * & * & * & 0 & * \\ * & 0 & * & * & * \end{pmatrix}$$

While

$$\mathrm{B} = \begin{pmatrix} * & 0 & 0 & 0 & 0 \\ * & * & 0 & * & 0 \\ * & * & * & 0 & 0 \\ * & * & * & * & 0 \\ * & * & * & * & 0 \end{pmatrix}$$

Then, we estimate the standard VECM with identifying restrictions explained above for the first non-inflation-targeting period and reestimate a second model for the second inflation-targeting period. The zeros represent the restricted elements and the asterisks denote unrestricted elements.

The bootstrap estimation allows us to determine unknown values of the short and long-run matrix for two periods.

59.4.1 Pre-IT (Before IT Adoption)

$$\text{Matrix A} = \mathrm{C}(1)\mathrm{B}$$
$$= \begin{pmatrix} 0.134 & 0.0125 & -0.1012 & -0.0668 & -0.00587 \\ 0 & 0 & 0.004 & 0.0649 & 0.0427 \\ -0.0644 & 0.0609 & 0.0622 & -0.0068 & 0.0068 \\ -0.0362 & -0.0602 & 0.0289 & 0 & 0.0325 \\ 0.0139 & 0 & 0.0112 & 0.0064 & 0.0325 \end{pmatrix}$$

While

$$\text{Matrix B} = \begin{pmatrix} 0.3601 & 0 & 0 & 0 & 0 \\ -0.0089 & 0.0005 & 0 & 0.0622 & 0 \\ -0.0415 & 0.0596 & 0.0725 & 0 & 0 \\ -0.0156 & -0.0613 & 0.0382 & 0.0061 & 0 \\ 0.0018 & 0.0007 & 0.0057 & 0.0028 & 0 \end{pmatrix}$$

59.4.2 Post-IT (After IT Adoption)

$$\text{Matrix A} = \mathrm{C}(1)\mathrm{B} = \begin{pmatrix} 0.1815 & 0.029 & 0.0428 & 0.0113 & 0.0123 \\ 0 & 0 & -0.003 & -0.0479 & -0.0324 \\ -0.0134 & -0.0784 & 0.02134 & -0.0011 & -0.0214 \\ 0.0012 & -0.0211 & 0.0012 & 0 & 0.0213 \\ 0.0011 & 0 & 0.0007 & 0.001 & 0.0012 \end{pmatrix}$$

While

$$\text{Matrix B} = \begin{pmatrix} 0.1851 & 0 & 0 & 0 & 0 \\ -0.0003 & 0.002 & 0 & -0.0471 & 0 \\ -0.0138 & -0.0756 & 0.0041 & 0 & 0 \\ -0.0035 & 0.0168 & 0.0571 & 0.0148 & 0 \\ 0.001 & 0.0003 & 0.0011 & 0.0011 & 0 \end{pmatrix}$$

To investigate the impulse response analysis, we compute impulse responses from the full SVECM, and we try to benefit from the important number of series

Table 59.5 Selected impulse responses: "impulse variable −> response variable" related to Appendix 3. Selected confidence interval (CI): (a) 95 % Hall percentile CI ($B = 100$ $h = 20$)

Time point estimate	MMR_log −>ExRate_log	MMR_log −>Share_prices_log	MMR_log −>IP_log	MMR_log −>CPI_log
Point estimate	−0.0089	−0.0415	−0.0156	0.0018
CI a)	[−0.0228, 0.0195]	[−0.0722, −0.0147]	[−0.0325, 0.0063]	[−0.0025, 0.0065]
1 Point estimate	−0.0035	−0.0555	−0.0281	0.0092
CI a)	[−0.0082, 0.0153]	[−0.0911, −0.0184]	[−0.0511, 0.0019]	[0.0027, 0.0153]
2 Point estimate	−0.0014	−0.0609	−0.0330	0.0120
CI a)	[−0.0030, 0.0107]	[−0.1013, −0.0169]	[−0.0603, 0.0026]	[0.0053, 0.0199]
3 Point estimate	−0.0005	−0.0631	−0.0349	0.0132
CI a)	[−0.0011, 0.0071]	[−0.1054, −0.0156]	[−0.0640, 0.0041]	[0.0060, 0.0216]
4 Point estimate	−0.0002	−0.0639	−0.0357	0.0136
CI a)	[−0.0004, 0.0046]	[−0.1070, −0.0134]	[−0.0654, 0.0053]	[0.0055, 0.0223]
5 Point estimate	−0.0001	−0.0642	−0.0360	0.0138
CI a)	[−0.0002, 0.0030]	[−0.1077, −0.0132]	[−0.0660, 0.0061]	[0.0049, 0.0226]
6 Point estimate	−0.0000	−0.0643	−0.0361	0.0138
CI a)	[−0.0001, 0.0019]	[−0.1079, −0.0131]	[−0.0662, 0.0065]	[0.0047, 0.0227]
7 Point estimate	−0.0000	−0.0644	−0.0361	0.0139
CI a)	[−0.0000, 0.0012]	[−0.1080, −0.0130]	[−0.0663, 0.0068]	[0.0046, 0.0228]
8 Point estimate	−0.0000	−0.0644	−0.0361	0.0139
CI a)	[−0.0000, 0.0008]	[−0.1081, −0.0130]	[−0.0663, 0.0069]	[0.0046, 0.0228]
9 Point estimate	−0.0000	−0.0644	−0.0362	0.0139
CI a)	[−0.0000, 0.0005]	[−0.1081, −0.0129]	[−0.0663, 0.0070]	[0.0045, 0.0228]
10 Point estimate	−0.0000	−0.0644	−0.0362	0.0139
CI a)	[−0.0000, 0.0003]	[−0.1081, −0.0129]	[−0.0663, 0.0071]	[0.0045, 0.0228]
11 Point estimate	−0.0000	−0.0644	−0.0362	0.0139
CI a)	[−0.0000, 0.0002]	[−0.1081, −0.0129]	[−0.0663, 0.0071]	[0.0045, 0.0228]
12 Point estimate	−0.0000	−0.0644	−0.0362	0.0139
CI a)	[−0.0000, 0.0001]	[−0.1081, −0.0129]	[−0.0663, 0.0071]	[0.0045, 0.0228]
13 Point estimate	−0.0000	−0.0644	−0.0362	0.0139
CI a)	[−0.0000, 0.0001]	[−0.1081, −0.0129]	[−0.0663, 0.0071]	[0.0045, 0.0228]

(*continued*)

Table 59.5 (continued)

Time point estimate				
	MMR_log −>ExRate_log	MMR_log −>Share_prices_log	MMR_log −>IP_log	MMR_log −>CPI_log
14 Point estimate	−0.0000	−0.0644	−0.0362	0.0139
CI a)	[−0.0000, 0.0001]	[−0.1081,−0.0129]	[−0.0663, 0.0071]	[0.0045, 0.0228]
15 Point estimate	−0.0000	−0.0644	−0.0362	0.0139
CI a)	[−0.0000, 0.0000]	[−0.1081,−0.0129]	[−0.0663, 0.0071]	[0.0045, 0.0228]
16 Point estimate	−0.0000	−0.0644	−0.0362	0.0139
CI a)	[−0.0000, 0.0000]	[−0.1081,−0.0129]	[−0.0663, 0.0071]	[0.0045, 0.0228]
17 Point estimate	−0.0000	−0.0644	−0.0362	0.0139
CI a)	[−0.0000, 0.0000]	[−0.1081,−0.0129]	[−0.0663, 0.0071]	[0.0045, 0.0228]
18 Point estimate	−0.0000	−0.0644	−0.0362	0.0139
CI a)	[−0.0000, 0.0000]	[−0.1081,−0.0129]	[−0.0663, 0.0071]	[0.0045, 0.0228]
19 Point estimate	−0.0000	−0.0644	−0.0362	0.0139
CI a)	[−0.0000, 0.0000]	[−0.1081,−0.0129]	[−0.0663, 0.0071]	[0.0045, 0.0228]
20 Point estimate	−0.0000	−0.0644	−0.0362	0.0139
CI a)	[−0.0000, 0.0000]	[−0.1081,−0.0129]	[−0.0663, 0.0071]	[0.0045, 0.0228]

taken in this chapter and presented for a long period by estimating series two times; the first treats the case before the adoption of inflation targeting by Turkish monetary policy authorities. The second allows us to examine real sphere during the adoption of inflation-targeting policy.

59.5 Interest Rate Shock and Real Sphere

Appendixes 3 and 4 show the responses of the real exchange rate, the shares prices, the industrial production, and the consumer prices to a monetary policy shock during the pre-IT and the post-IT policy, respectively. The confidence bounds were bootstrapped, since this gives more accurate confidence coverage compared to the asymptotic ones. Results are reported in Tables 59.6 and 59.7 in Appendix 5.

It is worth noting that for each figure presented in Appendixes 3 and 4, the horizontal axis of graphs shows the number of periods after a monetary policy shock has been initialized. The vertical axis measures the response of the relevant variables.

Table 59.6 Selected impulse responses: "impulse variable –> response variable" related to Appendix 4

Selected confidence interval (CI):

a) 95 % Hall percentile CI (B=100 h=20)

	MMR_log –>ExRate_log	MMR_log –>Share_prices_log	MMR_log –>IP_log	MMR_log –>CPI_log
Point estimate	−0.0003	−0.0138	−0.0035	0.0010
CI a)	[−0.0037, 0.0029]	[−0.0327,−0.0009]	[−0.0236, 0.0140]	[−0.0005, 0.0027]
1 Point estimate	−0.0001	−0.0135	−0.0006	0.0010
CI a)	[−0.0016, 0.0013]	[−0.0334,−0.0004]	[−0.0081, 0.0039]	[−0.0006, 0.0028]
2 Point estimate	−0.0000	−0.0135	0.0005	0.0011
CI a)	[−0.0007, 0.0006]	[−0.0343,−0.0001]	[−0.0032, 0.0034]	[−0.0006, 0.0028]
3 Point estimate	−0.0000	−0.0134	0.0009	0.0011
CI a)	[−0.0003, 0.0003]	[−0.0347,−0.0000]	[−0.0031, 0.0047]	[−0.0006, 0.0028]
4 Point estimate	−0.0000	−0.0134	0.0011	0.0011
CI a)	[−0.0001, 0.0001]	[−0.0349, 0.0000]	[−0.0034, 0.0053]	[−0.0007, 0.0028]
5 Point estimate	−0.0000	−0.0134	0.0012	0.0011
CI a)	[−0.0001, 0.0001]	[−0.0349, 0.0000]	[−0.0038, 0.0056]	[−0.0007, 0.0028]
6 Point estimate	−0.0000	−0.0134	0.0012	0.0011
CI a)	[−0.0000, 0.0000]	[−0.0350, 0.0000]	[−0.0039, 0.0057]	[−0.0007, 0.0028]
7 Point estimate	−0.0000	−0.0134	0.0012	0.0011
CI a)	[−0.0000, 0.0000]	[−0.0350, 0.0000]	[−0.0040, 0.0058]	[−0.0007, 0.0028]
8 Point estimate	−0.0000	−0.0134	0.0012	0.0011
CI a)	[−0.0000, 0.0000]	[−0.0350, 0.0001]	[−0.0040, 0.0058]	[−0.0007, 0.0028]
9 Point estimate	−0.0000	−0.0134	0.0012	0.0011
CI a)	[−0.0000, 0.0000]	[−0.0350, 0.0001]	[−0.0040, 0.0058]	[−0.0007, 0.0028]
10 Point estimate	−0.0000	−0.0134	0.0012	0.0011
CI a)	[−0.0000, 0.0000]	[−0.0350, 0.0001]	[−0.0040, 0.0058]	[−0.0007, 0.0028]
11 Point estimate	−0.0000	−0.0134	0.0012	0.0011
CI a)	[−0.0000, 0.0000]	[−0.0350, 0.0001]	[−0.0040, 0.0058]	[−0.0007, 0.0028]
12 Point estimate	−0.0000	−0.0134	0.0012	0.0011
CI a)	[−0.0000, 0.0000]	[−0.0350, 0.0001]	[−0.0040, 0.0058]	[−0.0007, 0.0028]

(*continued*)

Table 59.6 (continued)

Selected confidence interval (CI):				
a) 95 % Hall percentile CI (B=100 h=20)				
	MMR_log −>ExRate_log	MMR_log −>Share_prices_log	MMR_log −>IP_log	MMR_log −>CPI_log
13 Point estimate	−0.0000	−0.0134	0.0012	0.0011
CI a)	[−0.0000, 0.0000]	[−0.0350, 0.0001]	[−0.0040, 0.0058]	[−0.0007, 0.0028]
14 Point estimate	−0.0000	−0.0134	0.0012	0.0011
CI a)	[−0.0000, 0.0000]	[−0.0350, 0.0001]	[−0.0040, 0.0058]	[−0.0007, 0.0028]
15 Point estimate	−0.0000	−0.0134	0.0012	0.0011
CI a)	[−0.0000, 0.0000]	[−0.0350, 0.0001]	[−0.0040, 0.0058]	[−0.0007, 0.0028]
16 Point estimate	−0.0000	−0.0134	0.0012	0.0011
CI a)	[−0.0000, 0.0000]	[−0.0350, 0.0001]	[−0.0040, 0.0058]	[−0.0007, 0.0028]
17 Point estimate	−0.0000	−0.0134	0.0012	0.0011
CI a)	[−0.0000, 0.0000]	[−0.0350, 0.0001]	[−0.0040, 0.0058]	[−0.0007, 0.0028]
18 Point estimate	−0.0000	−0.0134	0.0012	0.0011
CI a)	[−0.0000, 0.0000]	[−0.0350, 0.0001]	[−0.0040, 0.0058]	[−0.0007, 0.0028]
19 Point estimate	−0.0000	−0.0134	0.0012	0.0011
CI a)	[−0.0000, 0.0000]	[−0.0350, 0.0001]	[−0.0040, 0.0058]	[−0.0007, 0.0028]
20 Point estimate	−0.0000	−0.0134	0.0012	0.0011
CI a)	[−0.0000, 0.0000]	[−0.0350, 0.0001]	[−0.0040, 0.0058]	[−0.0007, 0.0028]

The simulation analysis covers 20 periods. The solid lines for each graph denote impulse responses. The dotted lines are approximately 95 error bands (with 95 % confidence intervals) that are derived from a bootstrap routine with 100 replications. Bootstrap confidence bands are computed by percentile method advanced by Hall (1992) and Lütkepohl et al. (2001).

59.5.1 Pre-IT (Before IT Adoption)

Figure 59.3 presented in Appendix 3 reveals the responses of our series after an unexpected increase of the short-term interest rate which normally leads to an exchange rate appreciation (Mishkin 2001). This cannot be seen in the response of real exchange rate due to nonsignificant response in the short and long run.

Table 59.7 Kenneth S, Rogoff and Carmen M. Reinhart (2010)

Default, Restructuring, Banking Crises, Growth Collapses, and IMF Programs: Turkey, 1800–2010 (calculations since independence – 1923 – reported)

External default/ restructuring	Duration (in years)	Domestic default/ restructuring	Banking crisis (first year)	Hyper-inflation dates	Share of years in external default	Share of years in inflation crisis	5 worst output collapses year (decline)[a]
1876–1881	6	n.a.	1931	n.a.	19.5	35.6	1927 (9.1)
1915–1928	14		1982				1932 (6.0)
1931–1932	2		*1991*				1994 (5.5)
1940–1943	4		2000				2001 (5.7)
1959	1						2009 (*5.6*)
1965	1						
1978–1979	2						
1982	1						
2000–2001 "near"	2						
Number of episodes:							
8		0	4	0			
Memorandum item on IMF programs, 1952–2009							
Dates of programs							Total number 18
1961–1970, 1978–1980, 1983–1984, 1994, 1999, 2002							

[a]Excludes World Wars I and II
Summary of private forecasts for 2009 in italics

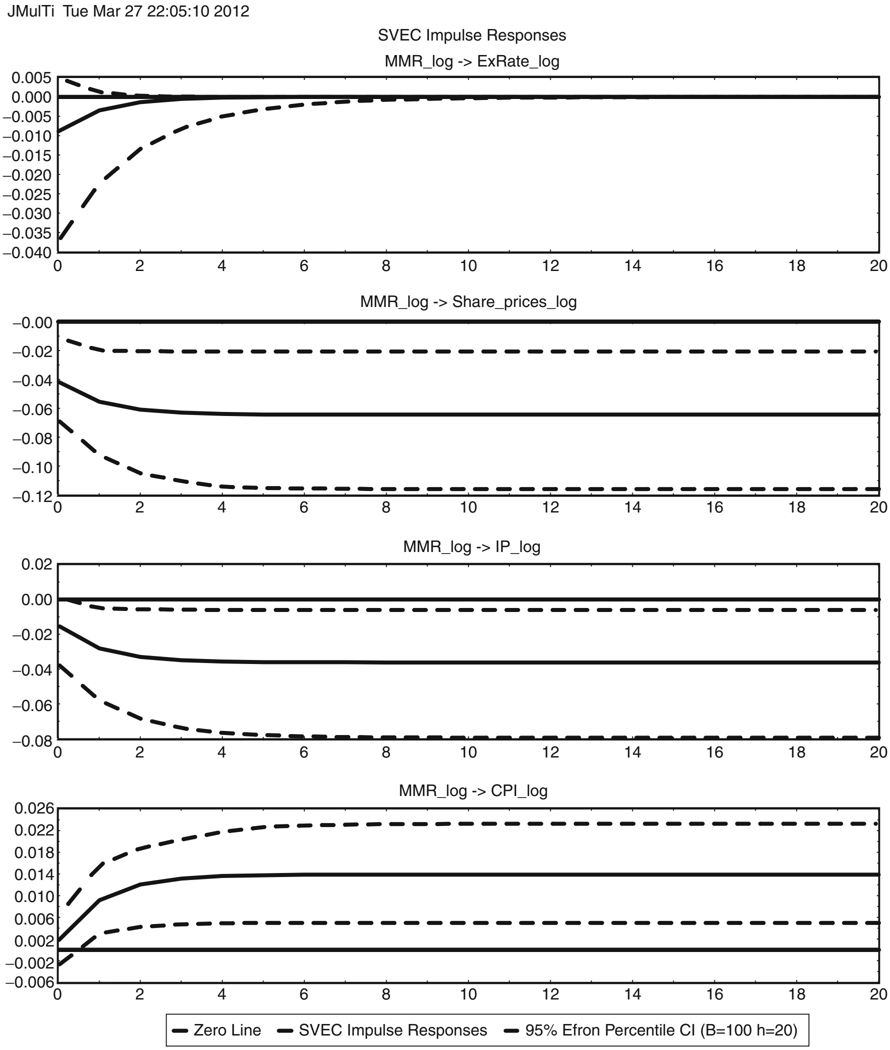


Fig. 59.3 Responses of the main macro variables to a monetary policy shock before IT

Despite the trial by Turkish authorities to implement a new stabilization and structural adjustment program, the banking sector was in turbulence in May 2001. According to Table 59.7, this result could be explained by the predominance of many preponderant banking crisis dates especially in 1991 and 2000 which were accompanied by a considerable output collapse in 2001 (Rogoff and Reinhart 2010).

59.5.2 Post-IT (After IT Adoption)

Figure 59.4 presented in Appendix 4 reveals the responses of our series after an unexpected increase of the short-term interest rate. Seeking to reduce the size of the money supply, normally, a contractionary monetary policy shock in the very

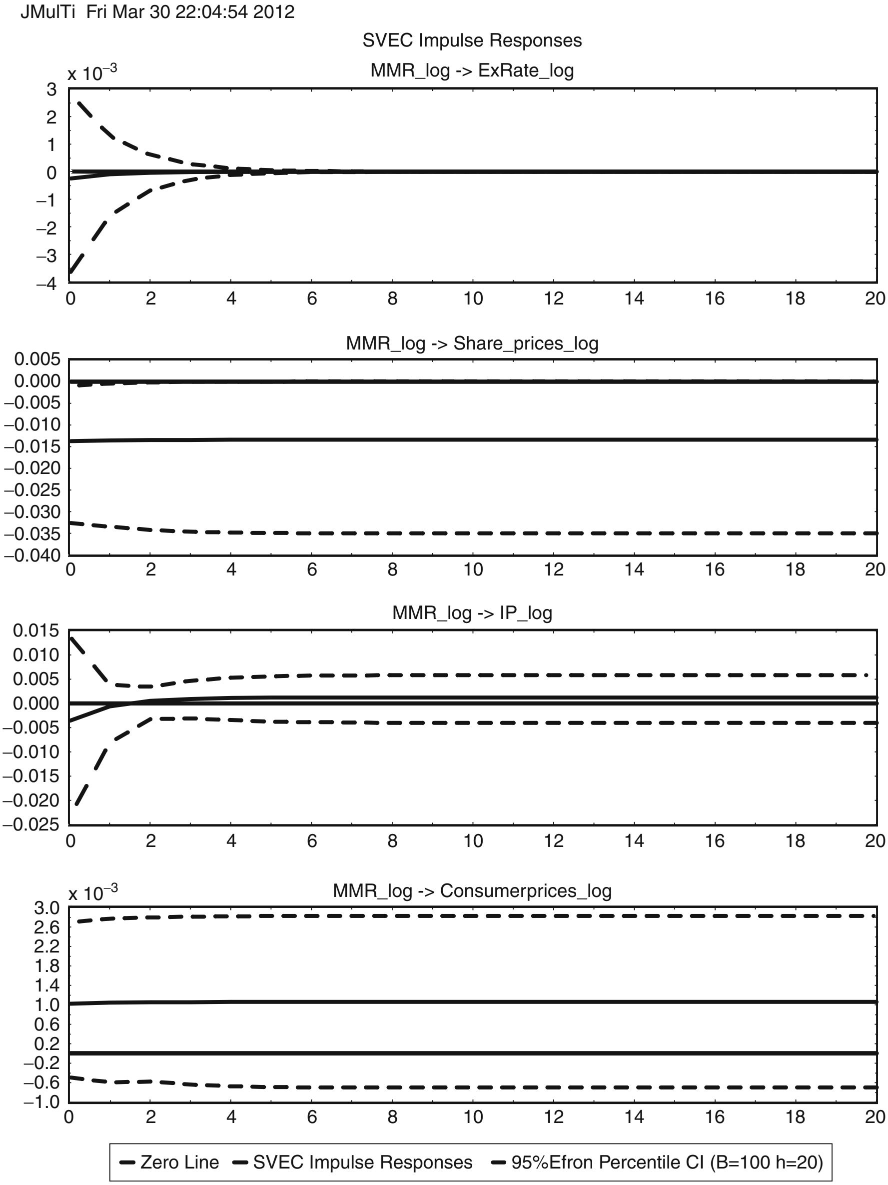


Fig. 59.4 Responses of the main macro variables to a monetary policy shock after IT adoption

short run depresses exchange rate which in turn reduces share prices, and it also appreciates consumer prices which in turn reduces the output. The other important result obtained from this study is that a contractionary policy has a nonsignificant effect on price level in only non-inflation-targeting period. This may be interpreted as a particular case of monetary policy effectiveness in inflation-targeting period. But, this does not mean that monetary policy in inflation-targeting period is more effective than monetary policy in non-inflation-targeting period. This result contradicts the findings of Akyurek and Kutan (2008) who advance that the performances of Turkey after inflation-targeting regime are better than its pre-inflation-targeting regime, at least in terms of containing inflation.

However, the findings reveal that monetary policy affects the price level, share prices, real exchange rate, and output level in the very short run before inflation-targeting policy. These results are similar to those advanced by Fair (2007) and Ball and Sheridan (2005) who show no evidence that inflation targeting improves a country's performance.

To conclude, our results are similar to those of the IMF World Economic Outlook (2005) which offers evidence of either "no increase or a decrease in the volatility of output" due to inflation-targeting policy.

59.6 Conclusion

In this chapter, we based our analysis on two SVEC models with long- and short-run restrictions to detect the impact and dynamic effects of a contractionary monetary policy shock on the output, inflation rate, share prices, and exchange rate. The overall responses of the macroeconomic variables in our SVEC models are consistent with most common theoretical expectations that we discussed in this chapter: an expansionary interest policy shock leads to a decrease in price level, a fall in output, an appreciation in the exchange rate, an improvement in the share prices in the very short run for the most of non-inflation-targeting period. For this pre-IT period, we did not find any evidence of empirical anomalies advanced in empirical literature, i.e., the price puzzle and the exchange rate puzzle.

Our approach in this chapter overcomes such empirical anomalies only for the pre-IT policy. The exchange rate puzzle and the price puzzle were observed for the whole inflation-targeting period.

One of the most important results of this study is that the impact effect of a monetary policy shock on the real sphere is negative and generally statistically significant only in pre-IT period.

Appendix 1: Joint Forecast Error Covariance Matrix, Pre-IT, and Post-IT Forecasting Results

The forecast errors have zero mean and, hence, the forecasts are unbiased. The joint forecast error covariance matrix for all forecasts up to horizon h is

$$Cov\begin{pmatrix} y_{T+1}-y_{T+1/T} \\ \ddot{} \\ . \\ \\ y_{T+h}-y_{T+h/T} \end{pmatrix} = \begin{pmatrix} I & 0 & . & . & 0 \\ \Phi_1 & I & . & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ \Phi_{h-1} & Q_{h-2} & . & . & I \end{pmatrix} \left(\sum\nolimits_u \otimes I_h\right) \begin{pmatrix} I & 0 & . & . & 0 \\ \Phi_1 & I & . & . & 0 \\ . & . & . & . & . \\ . & . & . & . & . \\ \Phi_{h-1} & \Phi_{h-2} & . & . & I \end{pmatrix}' \tag{59.9}$$

where $y_{T+1/T} = y_{T+j}$ for $j \leq 0$.

Assuming normally distributed disturbances, these results can be used for setting up forecast intervals for any linear combination of these forecasts.

See Tables 59.1 and 59.2, Figs. 59.1 and 59.2.

Appendix 2: Pre-IT and Post-IT Cointegration Tests

See Tables 59.3 and 59.4.

Appendix 3: Pre-IT Macro Variables Responses

See Fig. 59.3.

Appendix 4: Post-IT Macro Variables Responses

See Fig. 59.4.

Appendix 5: Impulse Responses and Confidence Intervals

See Tables 59.5 and 59.6.

References

Akyurek, C., & Kutan, M. A. (2008). Inflation targeting, policy rates and exchange rate volatility: Evidence from Turkey. *Comparative Economic Studies, 50*, 460–493.

Ball, L., & Sheridan, N. (2005). Does inflation targeting matter? In B. Bernanke & M. Woodford (Eds.), *The inflation-targeting debate*. Chicago: University of Chicago Press.

Bhuiyan, R. (2012). Inflationary expectations and monetary policy: Evidence from Bangladesh. *Empirical Economics*, Online First™, 26 April 2012.

Breitung, J., Bruggemann, R., & Lutkepohl, H. (2004). Structural vector autoregressive modelling and impulse responses. In H. Lutkepohl & M. Kratzig (Eds.), *Applied Time Series Econometrics*.

Cecchetti, S. (1995). Distinguishing theories of monetary transmission mechanism. *Economic Review of Federal Reserve of Saint-Louis Review*, 77 May-June, 83–97.

Clements, B. J., Kontolemis, G. Z., & Levy, J. V. (2001). Monetary policy under EMU: Differences in the transmission mechanisms? IMF Working paper, 01/102.

DeBondt, G. (2000). Credit channels and consumption in Europe: Empirical evidence. BIS Working paper, *69*.

Ehrmann, M. (1998). Will EMU generate asymmetry? Comparing monetary policy transmissions across European countries. *European University Institute*, Working paper ECO 98/28.

Fair, R. C. (2007). Evaluating inflation targeting using a macroeconometric model. *Economics*: *The Open-Access*, *Open-Assessment E-Journal*, *1*.

Hachicha, A., & Chaabane, A. (2007). Monetary policy transmission mechanism in Tunisia. *Euro-Mediterranean Economics and Finance Review, 1*, 104–26.

Hall, P. (1992). *The bootstrap and Edgeworth expansion*. New York: Springer.

IMF, (2005). Does inflation targeting work in emerging markets? In *World economic outlook* (Chapter 4, pp. 161–178).

Ivrendi, M., & Guloglu, B. (2010). Monetary shocks, exchange rates and trade balances: Evidence from inflation targeting countries. *Economic Modelling, 27*(5), 1144–1155.

Johansen, S. (1988). Statistical analysis of cointegrating vectors. *Journal of Economics and Dynamic Control, 12*, 231–254.

Johansen, S. (1995). *Likelihood-based inference in cointegrated vector autoregressive models*. Oxford: Oxford University Press.

Kakes, J. I., & Sturm, J. E. (2002). Monetary policy and bank lending: Evidence from German banking groups. *Journal of Banking and Finance, 26*, 2077–2092.

King, R., Plosser, C., Stock, J., & Watson, M. (1991). Stochastic trends and economic fluctuations. *The American Economic Review, 81*(4), 819–840.

Lucke, B. (2010). Identification and overidentification in SVECMs. *Economics Letters, 108*(3), 318–321.

Lütkepohl, H. (1993). *Introduction to multiple time series analysis*. Berlin: Springer.

Lutkepohl, H. (2005). Structural vector autoregressive analysis for cointegrated variables. EUI Working paper ECO.

Lütkepohl, H., Benkwitz, A., & Wolters, J. (1998). A money demand system for German M3. *Empirical Economics, 23*, 382–384.

Lütkepohl, H., Benkwitz, A., & Wolters, J. (2001). Comparison of bootstrap confidence intervals for impulse responses of German monetary systems. *Macroeconomic Dynamics, 5*, 81–100.

Mishkin, F. S. (2001). The transmission mechanism and the role of asset prices in monetary policy. NBER Working paper 8617, December.

Nadenichek, J. (2006). The J-curve effect: An examination using a structural vector error correction model. *International Journal of Applied Economics, 3*(2), 34–47.

Pfaff, B. (2008). VAR, SVAR and SVEC models: Implementation within R package vars. *Journal of Statistical Software*, *27*(4). http://www.jstatsoft.org/

Ramaswamy, R., & Sloek, T. (1998). The real effects of monetary policy in the European Union: What are the differences? *IMF Staff Papers, 45*, 347–395.

Rogoff, K., & Reinhart, C. M. (2010). Growth in a time of debt. NBER Working paper 15795.

Determination of Capital Structure: A LISREL Model Approach

60

Cheng-Few Lee and Tzu Tai

Contents

Abstract

Most previous studies investigate theoretical variables which affect the capital structure of a firm; however, these latent variables are unobservable and generally estimated by accounting items with measurement errors. The use of these observed accounting variables as theoretical explanatory latent variables will cause error-in-variable problems during the analysis of the factors of capital structure.

C.-F. Lee (✉)
Department of Finance and Economics, Rutgers Business School, Rutgers, The State University of New Jersey, Piscataway, NJ, USA

Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
e-mail: lee@business.rutgers.edu; cflee@business.rutgers.edu

T. Tai
Department of Finance and Economics, Rutgers, The State University of New Jersey, Piscataway, NJ, USA
e-mail: tzutai@pegasus.rutgers.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_60,

Since Titman and Wessels (*Journal of Finance 43*, 1–19, 1988) first utilize LISREL system to analyze the determinants of capital structure choice based on a structural equation modeling (SEM) framework, Chang et al. (*The Quarterly Review of Economic and Finance 49*, 197–213, 2009) and Yang et al. (*The Quarterly Review of Economics and Finance 50*, 222–233, 2010) extend the empirical work on capital structure research and obtain more convincing results by using multiple indicators and multiple causes (MIMIC) model and structural equation modeling (SEM) with confirmatory factor analysis (CFA) approach, respectively.

In this chapter, we employ structural equation modeling (SEM) in LISREL system to solve the measurement errors problems in the analysis of the determinants of capital structure and find the important factors consistent with capital structure theory by using date from 2002 to 2010. The purpose of this chapter is to investigate whether the influences of accounting factors on capital structure change and whether the important factors are consistent with the previous literature.

Keywords

Capital structure • Structural equation modeling (SEM) • Multiple indicators and multiple causes (MIMIC) model • LISREL system • Simultaneous equations • Latent variable • Determinants of capital structure • Error in variable problem

60.1 Introduction

In previous research in capital structure, many models are derived based on theoretical variables; however, these variables are often unobservable in the real world. Therefore, many studies use the accounting items from the financial statements as proxies to substitute for the theoretically derived variables. In the regression analysis, the estimated parameters from accounting items as proxies for unobservable theoretical attributes would cause some problems. First, there are measurement errors between the observable proxies and latent variables. According to the previous theoretical literature in corporate finance, a theoretical variable can be formed with either one or several observed variables as a proxy. But there is no clear rule to allocate the unique weights of observable variables as the perfect proxy of a latent variable. Second, because of unobservable attributes to capital structure choice, researchers can choose different accounting items to measure the same attribute in accordance with the various capital structure theory and the their bias economic interpretation. The use of these observed variables as theoretical explanatory latent variables in both cases will cause error-in-variable problems. Joreskog (1977) Joreskog and Sorbom (1981, 1989) and Jorekog and Goldberger (1975) first develop the structure equation modeling (hereafter called SEM) to analyze the relationship between the observed variables as the indicators and the latent variables as the attributes of the capital structure choice.

Since Titman and Wessels (1988) (hereafter called TW) first utilize LISREL system to analyze the determinants of capital structure choice based on a structural equation modeling (SEM) framework, Chang et al. (2009) and Yang et al. (2010) extend the empirical work on capital structure choice and obtain more convincing results. These papers employ structural equation modeling (SEM) in LISREL system to solve the measurement errors problems in the analysis of the determinants of capital structure and to find the important factors consistent with capital structure theories. Although TW initially apply SEM to analyze the factors of capital structure choice, their results are insignificant and poor to explain capital structure theories. Maddala and Nimalendran (1996) point out the problematic model specification as the reason for TW's poor finding and propose a multiple indicators and multiple causes (hereafter called MIMIC) model to improve the results. Chang et al. (2009) reproduce TW's research on determinants of capital structure choice but use MIMIC model to compare the results with TW's. They state that the results show the significant effects on capital structure in a simultaneous cause-effect framework rather than in SEM framework. Later, Yang et al. (2010) incorporate the stock returns with the research on capital structure choice and utilize structural equation modeling (SEM) with confirmatory factor analysis (CFA) approach to solve the simultaneous equations with latent determinants of capital structure. They assert that a firm's capital structure and its stock return are correlated and should be decided simultaneously. Their results are mainly same as TW's finding; moreover, they also find that the stock returns as a main factors of capital structure choice.

In this chapter, we compare the results of the determinants of capital structure from the period 2002–2010 with the results in previous chapter by using LISREL system. The purpose of this chapter is to investigate whether the influences of accounting factors on capital structure are of difference from TW's results and whether the important factors are consistent with the theories in previous literature. During the financial crisis, the influences of accounting factors on the firm's capital structure may have some difference due to the extremely decline of the equity market in the economic recession. Also, the method of reducing measurement error via the average of 3-year data may be invalid because the samples will have different time series pattern after significant event such as current financial crisis. Therefore, this chapter aims at whether the parameters used in TW paper are still significant or not during the financial crisis.

This chapter is organized as follows. In Sect. 60.2, we review the accounting items as proxies for latent variables in TW paper and describe the sample data used in LISREL system. Then, in Sect. 60.3, we introduce SEM to investigate the determinants of capital structure choice and illustrate the SEM approach for TW work in LISREL program. The results of empirical work and the analysis of the comparison with TW's finding are shown in Sect. 60.4. Finally, Sect. 60.5 represents the conclusions of this study.

Table 60.1 Attributes and indicators

Attributes	Indicators
Asset structure	Intangible asset/total assets (INT_TA)
	Inventory plus gross plant and equipment/total assets (IGP_TA)
Non-debt tax shield	Investment tax credits/total asset (ITC_TA)
	Depreciation/total asset (D_TA)
	Non-debt tax shields/total asset (NDT_TA)
Growth	Capital expenditures/total asset (CE_TA)
	The growth of total asset (GTA)
	Research and development/sales (RD_S)
Uniqueness	Research and development/sales (RD_S)
	Selling expense/sales (SE_S)
Industry classification	SIC code (IDUM)
Size	Natural logarithm of sales (LnS)
Volatility	The standard deviation of the percentage change in operating income (SIGOI)
Profitability	Operating income/sales (OI_S)
	Operating income/total assets (OI_TA)
Capital structure (dependent variables)	Long-term debt/market value of equity (LT_MVE)
	Short-term debt/market value of equity (ST_MVE)
	Convertible debt/market value of equity (C_MVE)
	Long-term debt/book value of equity (LT_BVE)
	Short-term debt/book value of equity (ST_BVE)
	Convertible debt/book value of equity (C_BVE)

60.2 Determinants of Capital Structure and Data

Before we utilize SEM approach to analyze the determinants of capital structure, the observable indicators are first briefly described in this section and then the data used in this chapter is subsequently introduced.

60.2.1 Determinants of Capital Structure

TW provide eight characteristics to determine the capital structure: asset structure, non-debt tax shields, growth, uniqueness, industry classification, size, volatility, and profitability. These attributes are unobservable; therefore, some useful and observable accounting items are classified into these eight characteristics in accordance with the previous literature on capital structure. The attributes as latent variables, their indicators as independent variables, and the indicators of capital structure as dependent variables are shown in Table 60.1. The parentheses in indicators are the notations used in LISREL system. Moreover, TW adopt the long-term debt, the short-term debt, and the convertible debt over either market value of equity or book value of equity as the indicators of capital structure as shown in the bottom of Table 60.1.

Based on the trade-off theory and agency theory, firms with larger tangible and collateral assets may have less bankruptcy, asymmetry information, and agency costs. Myers and Majluf (1984) indicate that companies with larger collateral assets attempt to issue more secured debt to reduce the cost arising from information asymmetry. Moreover, Jensen and Meckling (1976) and Myers (1977) state that there are agency costs related to underinvestment problem in the leveraged firm. Therefore, the collateral assets are positive correlated to debt ratios. According to TW paper, the ratio of intangible assets to total assets (INT_TA) and the ratio of inventory plus gross plant and equipment to total assets (IGP_TA) are viewed as the indicators to evaluate the asset structure attribute.

DeAngelo and Masulis (1980) extend Miller's (1977) model to analyze the effect of non-debt tax shields increasing the costs of debt for firms. Bowen et al. (1982) find their empirical work on the influence of non-debt tax shields on capital structure consistent with DeAngelo and Masulis's (1980) optimal debt model. Following Fama and French (2002) and TW paper, the indicators of non-debt tax shields are investment tax credits over total asset (ITC_TA), depreciation over total asset (D_TA), and non-debt tax shields over total asset (NDT_TA) which NDT is defined as in TW paper with the corporate tax rate 34 %.

According to TW paper, we use capital expenditures over total asset (CE_TA), the growth of total asset (GTA), and research and development over sales (RD_S) as the indicators of growth attribute. TW argue the negative relationship between growth opportunities and debt because growth opportunities only add firm's value but cannot collateralize or generate taxable income. Furthermore, the indicators of uniqueness include development over sales (RD_S) and selling expense over sales (SE_S). Titman (1984) indicates that uniqueness negatively correlates to debt because the firms with high-level uniqueness will cause customers, suppliers, and workers to suffer relatively high costs of finding alternative products, buyers, and jobs when firms liquidate. SIC code (IDUM) as proxy of industry classification attribute followed Titman's (1984) and TW's suggestions that firms manufacturing machines and equipment have high liquidation cost and thus more likely to issue less debt.

The indicator of size attribute is measured by natural logarithm of sales (LnS). The financing cost of firms may relate to firm size since small firms have higher cost of nonbank debt financing (see Bevan and Danbolt 2002). Therefore size is supposed to be positive associated with debt level. Besides, volatility attribute is estimated by the standard deviation of the percentage change in operating income (SIGOI). The large variance in earnings means higher possibility of financial distress; therefore, to avoid bankruptcy to happen, firms with larger volatility of earnings will have less debt. Finally, the pecking order theory developed in Myers (1977) paper indicates that firms prefer to use internal finance rather than external finance when raising capital. The profitable firms are likely to have less debt and

Table 60.2 The Compustat code of observable data

Accounting	Code	Accounting	Code
Total asset	AT	Net income	NI
Intangible asset	INTAN	R&D expense	RDIP
Inventory	INVT	Sales	SALE
Gross plant and equipment	PPEGT	Selling expense	XSGA
Investment tax credits	ITCB	SIC code	SIC
Depreciation	DPACT	Short-term debt	DLC
Income tax	TXT	Long-term debt	DLTT
Operating income	EBIT	Convertible debt	DCVT
Interest payment	XINT	Book value of equity	SEQ
Capital expenditures	CAPX	Market value of equity	MKVALT

profitability and hence are negatively related to debt level. Following TW paper, the indicators of profitability are operating income over sales (OI_S) and operating income over total assets (OI_TA).

60.2.2 Data

The sample period from 2002 to 2010 is the same as the duration of variables used in TW paper. The sources of data are from Compustat and CRSP in WRDS. The codes of the accounting items used to calculate the observed variables in Compustat are shown in Table 60.2.

The process of dealing with data is same as TW. The firms with incomplete record on variables and with negative values of total asset and operating income are deleted from the samples. After combination of all data from Compustat and CRSP, one indicator of non-debt tax shields is also excluded because almost all samples are zero or insignificant. For indicator of industry classification, we use dummy variable which equal to one for firms with SIC codes between 3400 and 4000 and equal to zero otherwise.

According to the problem of measurement errors, TW suggested that the sampling period should be divided into three subperiods. In each subperiod, the variables are the average of 3-year data due to random year-to-year fluctuations in variables. The dependent variables and independent variables used to measure uniqueness, non-debt tax shields, asset structure, and the industry classification are measured during the period 2005–2007. The indicators of size and profitability are measured during the period 2002–2004. Two independent variables, capital expenditures/total asset (CE_TA) and the growth of total asset (GTA), are measured during 2008–2010. Finally, the standard deviation of the percentage change in operating income (SIGOI) is estimated during the whole sample period 2002–2010 in order to obtain as the same measure as in TW paper.

60.3 Methodology and LISREL System

In this section, we first introduce the SEM approach and present an example of path diagram to show the structure of structural model and measurement model in SEM framework. Then, the specified structure in SEM approach is given in accordance with the constraints in TW paper and the code is illustrated in Appendix.

60.3.1 SEM Approach

The SEM incorporates three equations as follows:

$$\textit{Structural model: } \eta = \beta\eta + \Phi\xi + \varsigma \tag{60.1}$$

$$\textit{Measurement model for } y\text{: } y = \Lambda_{\mathrm{y}}\eta + \upsilon \tag{60.2}$$

$$\textit{Measurement model for } x\text{: } x = \Lambda\xi + \delta \tag{60.3}$$

where x is the matrix of observed independent variables as the indicators of attributes, y is the matrix of observed dependent variables as the indicators of capital structure, ξ is the matrix of latent variables as attributes, and η is the latent variables that link determinants of capital structure (a linear function of attributes) to capital structure(y).

Figure 60.1 shows an example of the path diagram of SEM approach where the observed independent variables $\mathrm{x} = (\mathrm{x}_1, \mathrm{x}_2, \mathrm{x}_3)'$ are located in rectangular, the observed dependent variables $\mathrm{y} = (\mathrm{y}_1, \mathrm{y}_2)'$ are set in hexagons, variables $\eta = (\eta_1, \eta_2)'$, $\xi = (\xi_1, \xi_2)'$ in ovals denote the latent variables, and the corresponding sets of disturbance are $\varsigma = (\varsigma_1, \varsigma_2)'$, $\upsilon = (\upsilon_1, \upsilon_2)'$, and $\delta = (\delta_1, \delta_2, \delta_3)'$.

The structural model can be specified as the system of equations which combines (60.1) and (60.2), and then we can obtain the structural model in TW paper as follows:

$$y = \Gamma\xi + \varepsilon \tag{60.4}$$

In this chapter, the accounting items can be viewed as the observable independent variables (x) which are the causes of attributes as the latent variables (ξ), and the debt-equity ratios represented the indicators of capital structure are the observable dependent variables (y).

The fitting function for maximum likelihood estimation method for SEM approach is the following:

$$F = \log|\Sigma| + tr\left(S\Sigma^{-1}\right) - \log|S| - (p+q) \tag{60.5}$$

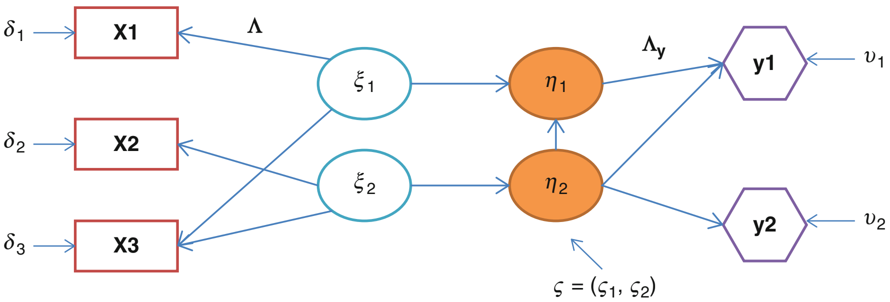


Fig. 60.1 Path diagram of SEM approach
In this path diagram, the SEM formulas (60.1), (60.2), and (60.3) are specified as follows: $\beta = \begin{bmatrix} 0 & \beta_1 \\ 0 & 0 \end{bmatrix}, \Phi = \begin{bmatrix} \Phi_1 & 0 \\ 0 & \Phi_2 \end{bmatrix}, \Lambda_y = \begin{bmatrix} \Lambda_{y1} & \Lambda_{y2} \\ 0 & \Lambda_{y3} \end{bmatrix}, \Lambda = \begin{bmatrix} \Lambda_1 & 0 \\ 0 & \Lambda_2 \\ \Lambda_3 & \Lambda_4 \end{bmatrix}$ where Λ_{y1}, Λ_{y2}, Λ_{y3}, Λ_1, Λ_2, Λ_3, and Λ_4 denote unknown factor loadings; β_1, Φ_1, and Φ_2 denote unknown regression weights; υ_1, υ_2, δ_1, δ_2, and δ_3 denote measurement errors; and ς_1, and ς_2 denote error terms

where S is the observed covariance matrix, Σ is the model-implied covariance matrix, p is the number of independent variables (x), and q is the number of dependent variables (y).

60.3.2 Illustration of SEM Approach in LISREL System

In general, SEM consists of two parts, the measurement model and structural model. The measurement model analyzes the presumed relations between the latent variables viewed as the attributes and observable variables viewed as the indicators. For example, in TW paper, capital expenditures over total assets (CE_TA) and research and development over sales (RD_S) are the indicators of the growth attributes (Growth). In the measurement model, each indicator is assumed to have measurement error associated with it. On the other hand, the structure model presents the relationship between unobserved variables and outcome. For instance, the relationship between attributes and the capital structure is represented by the structure model. Moreover, the relationship between the capital structure and its indicators estimated by debt-equity ratios is modeled by the measurement model.

TW also specific settings include zero measurement error of variables, the standard deviation of the percentage change in operating income (SIGOI) and SIC code (IDUM), measurement errors uncorrelated with each other indicator, with the latent variables, and with the errors in the structural equations. The attributes of volatility and industry classification equal to their indicators, respectively. Based on eight attributes as latent variables, thirteen indicators for determinants of capital structure choice, and six indicators of capital structure in TW paper, the SEM measurement model formula (60.3) and structural model formula (60.4) are specified as follows:

$$x = \begin{bmatrix} GTA \\ CE\_TA \\ RD\_S \\ SE\_S \\ D\_TA \\ NDT\_TA \\ INT\_TA \\ IGP\_TA \\ LnS \\ OI\_TA \\ OI\_S \\ SIGOI \\ IDUM \end{bmatrix}, \xi = \begin{bmatrix} Growth \\ Uniqueness \\ Non\_Debt\_Tax\_Shields \\ Asset\_Structure \\ Size \\ Profitability \\ Volatility \\ Industry\_Dummy \end{bmatrix},$$

$$\Lambda = \begin{bmatrix} \Lambda_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \Lambda_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \Lambda_3 & \Lambda_4 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \Lambda_5 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \Lambda_6 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \Lambda_7 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \Lambda_8 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \Lambda_9 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \Lambda_{10} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \Lambda_{11} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \Lambda_{12} & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \end{bmatrix}, \delta = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \delta_3 \\ \delta_4 \\ \delta_5 \\ \delta_6 \\ \delta_7 \\ \delta_8 \\ \delta_9 \\ \delta_{10} \\ \delta_{11} \\ 0 \\ 0 \end{bmatrix},$$

$$y = \begin{bmatrix} LT\_MVE \\ ST\_MVE \\ C\_MVE \\ LT\_BVE \\ ST\_BVE \\ C\_BVE \end{bmatrix}, \Gamma = \begin{bmatrix} \Gamma_{1,1} & \cdots & \Gamma_{1,8} \\ \vdots & \ddots & \vdots \\ \Gamma_{6,1} & \cdots & \Gamma_{6,8} \end{bmatrix}, \varepsilon = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{bmatrix}$$

where the variables for x, y, and ξ are defined as in Table 60.1. The codes of SEM in LISREL system are illustrated in Appendix.

60.4 Empirical Results and Analysis

In our empirical research, the estimates of the parameters of measurement model are presented in Tables 60.3 and 60.4. The regression coefficients of

Table 60.3 Measurement model: factor loadings of indicators for independent variables (x)

Attributes (the latent variables)									
Variable(x)	Growth	Uniqueness	Non-debt tax shields	Asset structure	Size	Profitability	Volatility	Industry dummy	σ_δ^2
GTA	**1.42** (4.25)								0.22
CE_TA	0.71 (1.84)								1.09
RD_S	0.0087 (0.46)	**0.064** (3.16)							0.0017
SE_S		−0.33 (−0.55)							14.91
D_TA			0.02 (0.47)						0.029
NDT_TA			0.094 (0.76)						0.033
INT_TA				**0.047** (5.15)					−0.00063
IGP_TA				0.0061 (0.93)					0.005
LnS					0.0079 (0.45)				0.00035
OI_TA						**0.041** (2.96)			0.018
OI_S						**−0.12** (−1.9)			0.027
SIGOI							1.000		0.000
IDUM								1.000	0.000

The *bold numbers* are significant at 5 % level where the t-statistics are in parentheses

Table 60.4 The estimated covariance matrix of attributes

	Growth	Uniqueness	Non-debt tax shields	Asset structure	Size	Profitability	Volatility	Industry dummy
Growth	1							
Uniqueness	−0.06 (−0.35)	1						
Non-debt tax shields	**0.2** (2.01)	**1.42** (2.75)	1					
Asset structure	0.1 (0.74)	0.02 (0.15)	−0.12 (−1.78)	1				
Size	−**0.17** (−3.36)	**0.11** (2.07)	**0.54** (5.47)	−**0.15** (−6.24)	1			
Profitability	−0.08 (−1.52)	**0.19** (4.37)	**0.39** (3.11)	−**0.65** (−5.99)	**0.42** (6.69)	1		
Volatility	0.03 (1.24)	0.01 (0.62)	0.04 (1.74)	−**0.06** (−4.42)	0 (0.02)	**0.11** (4.79)	**0.04** (10.37)	
Industry dummy	0.02 (1.19)	0.02 (0.97)	**0.07** (2.28)	−**0.06** (−4.05)	0.03 (0.30)	**0.12** (5.46)	**0.04** (7.04)	**0.04** (7.88)

The *bold numbers* are significant at 5 % level where the t-statistics are in parentheses

Table 60.5 Estimates of structural coefficients

	Attributes							
Debt measure	Growth	Uniqueness	Non-debt tax shields	Asset structure	Size	Profitability	Volatility	Industry dummy
LT_MVE	0.53	2.32	0.46	0.3	0.3	0.19	0.3	1.19
ST_MVE	0.055	0.0065	0.0034	−0.071	−0.071	0.026	−0.071	0.028
C_MVE	0.094	0.022	0.0073	−0.023	−0.023	−0.011	−0.023	0.039
LT_BVE	−0.00016	0.034	−0.0017	**−0.37**	0.026	−0.0092	0.026	0.026
ST_BVE	0.12	**−0.16**	0.17	−0.086	−0.086	0.0085	0.086	0.015
C_BVE	0.14	0.0047	0.044	−0.13	−0.13	−0.05	−0.13	−0.045

The *bold numbers* are significant at 5 % level matrix Λ in (60.3) are illustrated in Table 60.3, and most coefficients of independent variables except operating income over sales (OI_S) and intangible asset over total assets (INT_TA) are the same direction as TW paper. However, only five of them are significant. Table 60.4 shows the relationship between the attributes (the latent variables). Compared with TW results, only four relations are significantly the same: the relation between profitability and uniqueness, the relation between non-debt tax shields and size, the relation between profitability and non-debt tax shields, and the relation between asset structure and industry dummy. Even though the results of attributes' relations are different from TW paper, the estimated correlations between attributes in TW do not show the t-statistic value. Thus we cannot conclude which results represent correct and convincing relationships between the latent variables.

Moreover, there are only two significant estimates of structural coefficients (uniqueness and asset structure) in Table 60.5. The results inconsistent with TW paper may result from the several reasons as below. First, insignificant and incorrect latent variables may cause the wrong outcome of structural model. Besides, too many latent variables and the lack of using indicators with unique weights corresponding to their attributes may also cause the week results (Maddala and Nimalendran 1996). The other conjecture of inconsistent results is that the sampling fluctuation during financial crisis may aggravate the problem of measurement error.

Although the results of estimates of structural coefficients seem very week, the evidence of negative relationship between debt ratio and uniqueness is consistent the statement of Titman (1984) that the high costs of liquidation are imposed on the customers, workers, and suppliers of firms with high uniqueness products. Besides, the evidence of the attribute of asset structure negatively related to debt ratios corresponds to the supposition of Grossman and Hart (1982). They indicate that in order to avoid the threat of bankruptcy and closely monitor, managers in firms with higher debt are less likely to consume excessive perquisites.

60.5 Conclusion

This chapter utilizes the structure equation modeling (SEM) approach to estimate the impact of unobservable attributes on the capital structure. We use the sample period from 2002 to 2010 as same as the duration in TW paper to investigate whether the influences of accounting factors on capital structure are consistent with TW's results and whether the important factors are associated with the previous literature. In SEM framework, the debt ratios as indicators of capital structure choice to present the dependent variables and the observable accounting data from the financial statements used to calculate the indicators of attributes to form the latent variables.

Compared with the results of TW, our empirical work still cannot support the influence of most attributes on the decision of capital structure. The main reason of weak finding in our chapter and TW paper is too many latent variables as indicated in Maddala and Nimalendran (1996). And, another possible reason is the problem of sampling fluctuation during financial crisis which may cause serious measurement error in SEM approach. However, our empirical results show the significantly negative relationship between debt ratio and uniqueness which is consistent with the statement of Titman (1984) and TW results. Finally, in contrast with the unclear and insignificant relationship between the attribute of asset structure and debt ratios in TW paper, our finding of the significant negative relationship supports the argument of Grossman and Hart (1982) that firms with less collateral assets may choose higher debt levels to limit managers' consumption of perquisites.

Appendix: Codes of Structure Equation Modeling (SEM) in LISREL System

```
SEM Model-Titman and Wessels Paper
Observed Variables:
LT_MVE ST_MVE C_MVE LT_BVE ST_BVE C_BVE GTA CE_TA RD_S
SE_S D_TA NDT_TA INT_TA IGP_TA LnS OI_TA OI_S SIGOI
IDUM
Covariance Matrix from File TW0904.COV
Asymptotic Covariance Matrix from File TW0904.ACM
Sample Size: 125
Latent Variables: Growth Uniqueness
Non_Debt_Tax_Shields Asset_Structure Size
Profitability Volatility Industry_Dummy
Relationships:
LT_MVE = Growth Uniqueness
Non_Debt_Tax_Shields Asset_Structure Size
Profitability Volatility Industry_Dummy
ST_MVE = Growth Uniqueness
Non_Debt_Tax_Shields Asset_Structure Size
Profitability Volatility Industry_Dummy
C_MVE = Growth Uniqueness
Non_Debt_Tax_Shields Asset_Structure Size
Profitability Volatility Industry_Dummy
LT_BVE = Growth Uniqueness
Non_Debt_Tax_Shields Asset_Structure Size
Profitability Volatility Industry_Dummy
ST_BVE = Growth Uniqueness
Non_Debt_Tax_Shields Asset_Structure Size
Profitability Volatility Industry_Dummy
C_BVE = Growth Uniqueness
Non_Debt_Tax_Shields Asset_Structure Size
Profitability Volatility Industry_Dummy
Growth = GTA CE_TA RD_S
Uniqueness = RD_S SE_S
Non_Debt_Tax_Shields = D_TA NDT_TA
```

```
Asset_Structure = INT_TA IGP_TA
Size = LnS
Profitability = OI_TA OI_S
Volatility = 1.0*SIGOI
Industry_Dummy = 1.0*IDUM
Set the Error Variance of SIGOI to 0.0
Set the Error Variance of IDUM to 0.0
LISREL Output: PS = SY,FR TD = DI,FR ND = 3 SL = 0.05 SC SE
SS TV AL EF RS MI
Path Diagram
End of Problem
```

References

Bevan, A. A., & Danbolt, J. (2002). Capital structure and its determinants in the UK – A decompositional analysis. *Applied Financial Economics, 12*, 159–170.

Bowen, R. M., Daley, L. A., & Huber, C. C., Jr. (1982). Evidence on the existence and determinants of inter-industry differences in leverage. *Financial Management, 4*, 10–20.

Chang, C., Lee, A., & Lee, C. F. (2009). Determinants of capital structure choice: A structural equation modeling approach. *The Quarterly Review of Economic and Finance, 49*, 197–213.

DeAngelo, H., & Masulis, R. (1980). Optimal capital structure under corporate and personal taxation. *Journal of Financial Economics, 8*, 3–27.

Fama, E. F., & French, K. R. (2002). Testing tradeoff and pecking order predictions about dividends and debt. *Review of Financial Studies, 15*, 1–33.

Grossman, S., & Hart, O. (1982). Corporate financial structure and managerial incentives. In *The economics of information and uncertainty*. Chicago: University of Chicago Press.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: Managerial behavior, agency costs and ownership structure. *Journal of Financial Economics, 3*, 305–360.

Joreskog, K. G. (1977). Structural equation models in the social sciences: Specification estimation and testing. In P. R. Krishnaiah (Ed.).

Joreskog, K. G., & Goldberger, A. S. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of American Statistical Association, 70*, 631–639.

Joreskog, K. G., & Sorbom, D. (1981). *LISREL V, analysis of linear structural relationships by the method of maximum likelihood*. Chicago: National Educational Resources.

Joreskog, K. G., & Sorbom, D. (1989). *LISREL 7: A guide to the program and applications* (2nd ed.). Chicago: SPSS.

Maddala, G. S., & Nimalendran, M. (1996). Error-in-variables problems in financial models. *Handbook of statistics* 14 (pp. 507–528). Elsevier Science Publishers B. V.

Miller, M. (1977). Debt and taxes. *Journal of Finance, 32*, 261–275.

Myers, S. C. (1977). Determinants of corporate borrowing. *Journal of Financial Economics, 5*, 146–175.

Myers, S. C., & Majluf, N. (1984). Corporate financing and investment decision when firms have information investors do not have. *Journal of Financial Economics, 13*, 187–221.

Titman, S. (1984). The effect of capital structure on a firm's liquidation decision. *Journal of Financial Economics, 13*, 137–151.

Titman, S., & Wessels, R. (1988). The determinants of capital structure choice. *Journal of Finance, 43*, 1–19.

Yang, C. C., Lee, C. F., Gu, Y. X., & Lee, Y. W. (2010). Co-determination of capital structure and stock returns – A LISREL approach: An empirical test of Taiwan stock markets. *The Quarterly Review of Economics and Finance, 50*, 222–233.

Evidence on Earning Management by Integrated Oil and Gas Companies

61

Raafat R. Roubi, Hemantha Herath, and John S. Jahera Jr.

Contents

Abstract

The objective of this chapter is to demonstrate specific test methodology for detection of earnings management in the oil and gas industry. This study utilized several parametric and nonparametric statistical methods to test for such earnings management. The oil and gas industry was used given the earlier evidence where such firms manage earnings in order to ease the public view of the significant price swings that occur in oil and gas prices. In this chapter, our

R.R. Roubi (✉) • H. Herath
Department of Accounting, Faculty of Business, Brock University, St. Catharines, ON, Canada
e-mail: rroubi@brocku.ca; hherath@brocku.ca

J.S. Jahera Jr.
Department of Finance, College of Business, Auburn University, Auburn, AL, USA
e-mail: jahers@auburn.edu; jjahera@business.auburn.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_61,

focus is on total accruals as the primary means of earnings management. The prevailing view is that total accruals account for a greater amount of earnings management and should be more readily detected. The model to be considered is the Jones model (*Journal of Accounting Research 29*, 193–228, 1991) which projects the expected level of discretionary accruals. By comparing actuals vs. projected accruals, we are able to compute the total unexpected accruals. Next, we correlate unexpected total accruals with several difficult to manipulate indicators that reflect company's level of activities. The significant positive correlations between unexpected total accruals and these variables are an indication that oil and gas firms do not manage income before extraordinary items and discontinued operations. A second test is conducted by focusing on the possible use of special items to reduce reported net income by comparing mean levels of several special items pre-2008 and 2008. The test results indicate significant difference between 2008 means and the pre-2008 period.

Keywords

Earnings management • Jones model (1991) • Discretionary accruals • Income from operations • Nonrecurring items • Special items • Research and development expense • Write-downs • Political cost • Impression management • Oil and gas industry

61.1 Introduction

Repeated oil crises have often resulted in above normal profits for integrated oil and gas companies. Such above normal profits typically attract attention in the news media and consequently in the political environment. In this circumstance, the political environment in North America typically provides a short-lived threat of higher taxes and/or regulations. In the latest crude oil price hike in 2008 that reached \$140 per barrel, oil and gas companies reported unusually high profits. This was true even though, by the end of 2008, prices were about 50 % of their peak. Reporting high income numbers represent a short-term threat that focuses public attention on these companies for a period of time until consumers adapt to new and higher prices (i.e., \$2.89 per gallon in 2010 is way above \$1.90 per gallon in 2006–2007, but it is significantly lower than a near \$5.00 per gallon in 2008–2009).

The agency model predicts that faced with higher taxes and stricter regulations, management tends to use accounting accruals and/or special items to decrease reported income. Because of the possibility of reversals of all or of these accruals in future years, management has to time their responses to such crises to make its point to influence public opinion. Thus, the purpose of earnings management by management of these companies is to buy time until the public attitudes adjust to the new pricing and profitability levels. Earnings management is by no means a tool that can be used over a long period (i.e., multiple years) given the nature of the financial accounting accruals (i.e., firms cannot decrease sales by delaying revenue recognition over several years or keep reporting higher depreciation expense year after year). It is an important issue however for both regulators and those in the accounting profession, particularly auditors.

61.2 Literature Review

The earnings management literature can be broadly divided into two categories: (i) accrual management and (ii) operating decisions (real activity manipulation or economic earnings management). Accruals management is primarily the accounting choices available to management under generally accepted accounting practices (GAAP) that obscure true economic performance (Dechow and Skinner 2000). On the contrary, real earnings management takes place when managers intentionally change the timing or structuring of an operation, investment, or financing transaction to influence the output of an accounting system (Gunny 2010). Accruals management does not have any direct cash flow consequences as it is the manipulation of accounting numbers by using different accounting procedures and/or revising some specific accounting items to obtain desired reported earnings. In real earnings management, managers tend to take actions that affect cash flows and eventually earnings (Gupta et al. 2010). Over production by firms building up inventories when demand is falling allows management to report increased earnings because GAAP mandated the use of absorption or full costing for reporting purposes.

Much has been written in the area of earnings management. Earnings management may arise in several contextual settings where it can be identified that there exist conditions in which managers' incentive to manage earning is large (Healy and Wahlen 1999; Marquardt and Weidman 2004a among others). The empirical research has investigated many different incentives and settings conducive for earnings management. Some of settings investigated in Marquardt and Weidman (2004a) include the following: (1) equity offerings, where the motivation to manage earnings around equity offerings is viewed as increasing the stock price to benefit the firm; (2) management buyouts which is an opposite goal to reduce the stock price; and (3) firms attempting to avoid earning decreases. In general, as indicated above, earnings management could be either income increasing or income decreasing.

The incentives as per Healy and Whalen (1999) include the following: (1) Capital market expectations and valuation. Here the argument made is that the widespread use of accounting information by investors and financial analysts to help value stock can create an incentive for managers to manipulate earning. (2) Contracts written in terms of accounting numbers. Compensation contracts which are based on accounting numbers are used to align the incentives of management with that of external stakeholders. Watts and Zimmerman (1979) argue that contacts create incentives for earnings management, because it is costly for compensation committees and creditors to "undo" earnings management. (3) Antitrust or other government regulation. The idea is that accounting discretion is used to manage industry-specific regulatory constraints.

In a more recent study, Ang (2011) examines the effect of the Sarbanes-Oxley Act (SOX), specifically Section 404, on earnings management among large firms as defined by the Fortune 500. Ang's findings are that firms tend to filter information that investors receive as well as disguise the true earnings of the firm. Overall, Ang concludes that earnings management activity was diminished with the enactment of Sarbanes-Oxley. In another recent study, Ota (2011) examines earnings forecast in a global setting, more specifically for Japanese firms. This study considers earnings management and its determinants with the conclusion that issues related to distress, growth, size, as well as prior forecasting errors are all related to earnings management in Japan. A further part of this study concludes that analysts are indeed knowledgeable of earnings management practice and take such practices into consideration when preparing their own independent forecasts of earnings.

Habib and Hansen (2008) provide an updated literature review regarding earnings management. Their work essentially extends and updates the earlier work of Healey and Wahlen. In an earlier work, Burgstahler and Dichev (1997) consider the distribution of earnings changes and find a lower frequency of decreases in earnings but higher frequencies of increases. The implication is of course that management seeks to avoid such decreases in reported earnings.

Earnings management has potential costs which may be high or low depending on the situation which can be grouped into (1) costs of detected earnings management and (2) the cost of undetected earnings management (Marquardt and Wiedman 2004b). If a firm's use of earnings management becomes publicly known through the release of SEC enforcement actions, earnings restatements, shareholder litigation, qualified audit opinion, and negative business coverage by press, then it is termed detected earnings management. In undetected earnings management, there is no obvious event of public announcement of its occurrence. In terms of the research methodologies employed in earnings management studies, work by McNichols (2000) critically reviews several approaches from an empirical standpoint. In summary, McNichols considers three approaches typically seen in such analyses. Her work contends that further research should consider alternative specifications as opposed to the more traditional aggregate accrual models.

61.3 Empirical Methodology and Data

61.3.1 Integrated Companies

Integrated oil and gas companies are in the public spotlight whenever there is a spike in consumer prices. Other companies within the industry (e.g., refineries, transportation including pipelines, equipment, and service companies) may or may not report higher profits but are not subject to the same degree of media scrutiny and criticism (i.e., price gouging). In addition, large integrated companies have the resources needed to manage their earnings. Smaller oil and gas producers, with limited resources and limited operations, may find it more difficult to be involved in earnings management. At the time this study has been conducted, we consider 2008 to be a target year for earnings management. Our contention is based on the fact that crude oil prices reached a record price of $140 per barrel in 2008.

61.3.2 Tools of Earnings Management and Associated Costs

Tools of earnings management may include total accruals (A/R, inventory, accounts payable, and depreciation). They may also use special items including write-downs, restructuring charges, and discontinued operations. Companies involved in earnings management are always subject to scrutiny by regulatory agencies (i.e., SEC), external auditors, financial analysts, as well as media attention. Management of such firms has to judge each situation to determine whether it is wise to get involved in this type of activities. According to Marquardt and Wiedman (2004b), the use of special items carries a small cost if discovered, while managing of revenue (A/R) carries a very high cost if discovered. In the current study, we do not presume that management is biased for/against any tool. The argument we use is that a significant reduction of an extremely high profitability situation may prove appropriate. Also, the nature of the industry and the diversity of these companies' businesses (i.e., search, exploration, development, production, transportation, and marketing of oil and gas products) would allow a very wide range of possibilities for management of these companies to influence reported income. For instance, assessing goodwill for impairment provides an excellent opportunity to drive reported profits down because of the conservative nature of impairment losses. Likewise, accelerating the process of expensing research and development costs is another conservative accounting tool that would not raise auditor's concern but yet could affect reported profit. Also, the use of special items such as an upward revision of the provision for site restoration and environmental protection is a very strong tool in management's hands to drive down reported profits.

61.3.3 Measures of Earnings Management

Complete discussion of the empirical methodology is included in the Appendix. In general, the detection of earnings management is based on a variety of measures including (for more details, please see Ronen and Yaari 2008) serial correlation of income streams, the standard deviation of earnings relative to the standard deviation of cash flows, and the correlation between discretionary accruals and change in operating cash flows. In this chapter, our focus is on total accruals as the primary means of earnings management. The prevailing view is that total accruals account for a greater amount of earnings management and should be more readily detected. The model to be considered is the Jones model (1991) which projects the expected level of discretionary accruals.

61.3.4 Data

The sample is drawn from the North America COMPUSTAT – Fundamentals Annual through the Wharton Research Database and consists of international integrated oil and gas companies that carry a GSUBIND code 10102010. The initial sample consists of 1,428 annual observations covering the period 1950–2010. Companies included in the final sample of 836 annual observations met the following conditions: (1) information is available for the 2008 event year, (2) there is a minimum number of 6 annual consecutive observations per company, and (3) all data needed for analysis is available. After excluding out-of-range unusable observations, the final sample consists of 28 integrated oil and gas companies from Argentina, Austria, Brazil, Canada, China, France, Italy, Kazakhstan, the Netherlands, Norway, Russia, South Africa, Spain, the UK, and the USA.

61.3.5 Data Items

The following data items (COMPUSTAT – Fundamentals Annual) were collected for all firms with the GSUBIND code 10102010 and used in the empirical part of this study. The following mnemonics are collected:

AT	Total assets for
GDWLIP	Goodwill impairment (pretax)
IB	Income before extraordinary items
OANCF	Cash flow from operations – net
PPEGT	Property, plant, and equipment – gross
SALE	Sales – net
SPIOP	Special items – pretax
XIDO	Extraordinary items and discontinued operations
XRD	Research and development expense
WDP	Write-downs – pretax

61.4 Empirical Findings

61.4.1 Does Management Employ Total Accruals to Manage Earnings?

The presence of earnings management can be detected by observing a negative correlation between unexpected total accruals ($UTACC_{2008}$) and a change in operating cash flows for the same year ($\Delta OANCF_{2008}$) (see Leuz et al. 2003). Table 61.1 provides parametric and nonparametric correlation coefficients for $UTACC_{2008}$ and several variables including $\Delta OANCF_{2008}$. The statistics are presented for all companies, North American companies and the non-North American companies. The results that are presented indicate no evidence of

Table 61.1 Correlation coefficients for $UTACC_{2008}$

Sample	$\Delta OANCF_{2008}$	$OANCF_{2007}$	$OANCF_{2008}$	$SALES_{2008}$
All Observations (28)				
Pearson corr.	−0.020	0.120	0.097	0.102
Kendall's tau	0.058	0.058	0.101	0.122
Spearman's Rho	0.096	0.102	0.123	0.193
North American (14)				
Pearson corr.	0.427*	0.245	0.284	0.295
Kendall's tau	0.275*	0.231	0.319*	0.341*
Spearman's Rho	0.473**	0.314	0.455**	0.504**
Non-North American (14)				
Pearson corr.	−0.075	0.159	0.096	0.098
Kendall's tau	−0.011	0.011	0.033	0.033
Spearman's Rho	−0.051	0.064	0.055	0.064

*, **, *** significant at the 10 %, 5 %, and 1 %, respectively

Table 61.2 T-test results – special items (2008 versus pre-2008)

Item	Observations 2008	Observations pre-2008	Mean 2008	Mean pre-2008	T-value	One-tail significance
SPI	24	660	1,492.30	48.96	5.419	0.000
WDP	4	32	1,862.47	410.39	1.994	0.027
GDWLIP	5	7	5,507.40	328.10	1.250	0.120
XRD	13	415	535.33	200.20	5.075	0.000

earnings management. In fact, the significant positive correlation coefficients between UTACC2008 and $\Delta OANCF_{2008}$ suggest that North American integrated oil and gas companies actually do not use total accruals to manage earnings.

61.4.2 Does Management Employ Special Items to Manage Earnings?

Using special items to manage earnings is (1) less costly, i.e., unlike the case with manipulating revenues and/or operating expenses which are closely monitored by auditors, and (2) has no reversal problem, i.e., the timing of write-downs and/or spending is under management control. The results presented in Table 61.2 indicate that special items SPI which include goodwill impairment (GDWLIP), in-progress research and development (RDIP), restructuring costs (RCP), and other write-downs before taxes were significantly higher in 2008 ($1,492.30 million) versus an average of $48.96 million for the pre-2008 fiscal year. The t-test value of 5.419 is significant at the 0.000 level. The results in Table 61.2 also provide t-test statistics on some individual components of SPI. For example, the 2008 mean write-down before taxes of $1,862.47 million is significantly higher than the pre-2008 mean write-down of $410.39. The t-test statistic here shows a t-value of 1.994 which is significant at the

0.027 level. The component GDWLIP does not reveal a significant increase in recognizing goodwill impairment though these companies reported on average $5,507.40 million impairment losses in 2008 compared to only $328.10 mean losses in the pre-2008 period. One should notice that individual means do not add up to the total SPI because of the different number of observations used in computing individual components' means. Other components of SPI are missing either because they are not available in the data base or because they do not display significant t-values.

Research and development expenditures represent another discretionary item that management is able to control (see Shehata, October 1991). As the results reported in Table 61.2 reveal, average spending on research and development (XRD) is approximately $535.33 million in 2008 compared to a pre-2008 average of $200.20 million. The t-test statistic indicates a t-value of 5.075 at the 0.000 level.

Given the results reported in Table 61.2, one may conclude that management of integrated oil and gas companies does indeed employ special items as a tool to manage earnings.

61.5 Conclusion

Earnings management, which is an extensively researched topic in accounting, could be either income increasing or income decreasing. While there are several motivations for managing earnings and the associated costs depending whether earnings management is detected through public announcements or not, there are implications for incentive design and standard setting in a firm. In this chapter we explore the income decreasing earnings management by integrated oil and gas companies. Our findings provide evidence that integrated oil and gas companies do indeed use special items to manage earnings downwards. While there could be several reasons, one main reason is impression management. Although we selected earnings management as our topic, our primarily objective in this chapter is to demonstrate the important application of standard statistical analysis to the issue of detecting earnings management practices.

Appendix: Methodology

Detecting Earnings Management

The paper uses the Jones model (1991) as the basis for projecting the expected level of discretionary accruals. The steps are as follows:

Step 1: Define total accruals ($TACC$) for as the difference between income (NI) before extra ordinary items and discontinued operations (EOI) and cash flows from operations (OCF). Compute the $TACC_{it} = NI_{it} - EOI_{it} - OCF_{it}$ for each year (t) for firm (i).

Step 2: Divide the data available into two periods, namely, the estimation period $t = 1, \ldots, T_i$ and the prediction period $p = 1, \ldots, P$.

Step 3: Use ordinary least squares to obtain the coefficient estimates a_i, b_i, and c_i of α_i, β_i, and γ_i, respectively. The linear regression expectation model (Jones 1991) for total accruals after controlling for changes in the economic circumstances of the firm is

$$\frac{TACC_{it}}{TA_{it-1}} = \alpha_i \left[\frac{1}{TA_{it-1}}\right] + \beta_i \left[\frac{\Delta REV_{it}}{TA_{it-1}}\right] + \gamma_i \left[\frac{PPE_{it}}{TA_{it-1}}\right] + \varepsilon_{it} \tag{61.1}$$

where

$TACC_{it}$	is the total accruals in year (t) for firm (i)
ΔREV_{it}	is the revenues in year (t) less revenues in year ($t-1$) for firm (i)
PPE_{it}	is the gross property, plant, and equipment in year (t) for firm (i)
TA_{it-1}	is the total assets in year ($t-1$) for firm (i)
ε_{it}	is the error term in year (t) for firm (i)
$i = 1, \ldots, N$	is the firm index ($N = XX$)
$t = 1, \ldots, T_i$	is year index for the years included in the estimation period for firm (i), where T_i ranges between 6 and 22 years.

The gross property, plant, and equipment and change in revenue are included to control for changes in nondiscretionary accruals due to changing conditions. All variables in the accruals expectation model are scaled by lagged assets to reduce heteroscedasticity. The lagged assets are assumed to be positively associated with the variance of the disturbance term.

Model (1) was used to calculate the coefficients α_i, β_i, and γ_i for the estimation period ending in 2007. This procedure is conducted on a company-by-company basis and produced 28 individual models that fit the *TACC* history of each of the sample companies. Each of these models is used in step 4 below to estimate *TACC* for the event year 2008 for each of the 28 firms.

Step 4: Compute the discretionary accruals for the event year (2008) for firm (i) as follows.

The coefficients α_i, β_i, and γ_i obtained from running model (1) are used to predict $TACC_{2008}$ for each of the 28 sample firms. This procedure produced an estimated $ETACC_{2008}$ which is compared to the actual $ATACC_{2008}$. This comparison produced the unexpected total accruals ($UTACC_{2008}$) for the event year.

$$UTACC_{2008} = ETACC_{2008} - ATACC_{2008} \tag{61.2}$$

where

$UTACC_{2008}$	is the unexpected total accruals for the event year 2008
$ETACC_{2008}$	is the estimated expected total accruals for the event year 2008
$ATACC_{2008}$	is the actual total accruals for the event year 2008.

Step 5: Test for earnings management. This study ran several parametric and nonparametric statistics to test for earnings management as follows:

1. Correlations: We measured correlation coefficients (Pearson, Kendall's tau, and Spearman's Rho) to test for significant correlation between $UTACC_{2008}$ and $\Delta ONACF$ (a change in operating cash flows). A significant negative correlation between the two variables would represent a good sign of earnings management, i.e., a higher level of a change in operating cash flows negatively correlates with higher $UTACC_{2008}$ (Leuz et al., September 2003).
2. T-test: We compared means $UTACC_{2008}$ for North American companies against all other international companies included in the sample.
3. Regression analysis: We ran the following regression model

$$UTACC_{2008} = \beta_i \text{ Country} + \xi \tag{61.3}$$

where

$UTACC_{2008}$	is the unexpected total accruals for 2008
Country	is a partition variable coded 1 for North American companies, 0 otherwise
ξ	is an error term.

Step 6: Testing for use of special items to manage earnings.

This study assumes that integrated oil and gas companies will find it easier and less costly to manage earnings using special items such as *SPI*, *WDP*, *GDWLIP*, and *XRD*. T-test is used to compare means of each of these items for 2008 fiscal year and pre-2008 means. Unlike total accruals which require a fitting period and an event period, special items' expected level is always zero (Marquardt and Wiedman 2004a). Consequently, the unexpected amount of any of the above items is

$$\begin{aligned} USPI_{2008} &= SPI - 0, \\ UWDP_{2008} &= UWDP - 0, \\ UGDWLIP_{2008} &= GDWLIP - 0, \\ UXRD_{2008} &= XRD - 0 \end{aligned}$$

where

$USPI_{2008}$	is unexpected special items for fiscal year 2008
$UWDP_{2008}$	is unexpected write-downs for fiscal year 2008
$UGDWLIP_{2008}$	is unexpected goodwill impairment loss for fiscal year 2008
$UXRD_{2008}$	is unexpected research and development for fiscal year 2008.

References

Ang, R. (2011). Sarbanes-Oxley effectiveness on the earnings management. Available at SSRN http://ssrn.com/abstract=1791766

Burgstahler, D., & Dichev, I. (1997). Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics, 24*, 99–126.

Dechow, P. M., & Skinner, D. J. (2000). Earnings management: Reconciling the views of accounting academics, practitioners, and regulators. *Accounting Horizons, 14*, 235–250.
Gunny, K. (2010). The relation between earnings management using real activities manipulation and future performance: Evidence from meeting earnings benchmarks. *Contemporary Accounting Research, 27*, 855–888.
Gupta, M., Pevzner, M., & Seethamraju, C. (2010). The implications of absorption cost accounting and production decisions for future firm performance and valuation. *Contemporary Accounting Research, 27*, 889–922.
Habib, A., & Hansen, J. (2008). Target shooting: Review of earnings management around earnings benchmarks. *Journal of Accounting Literature, 27*, 25–70.
Healy, P., & Wahlen, J. (1999). A review of the earnings management literature and its implications for standard setting. *Accounting Horizons, 13*, 365–383.
Jones, J. (1991). Earnings management during import relief investigations. *Journal of Accounting Research 29*, 193–228.
Leuz, C., Nanda, D., & Wysocki, P. (2003). Earnings management and investor protection: An international comparison. *Journal of Financial Economics, 69*, 505–527.
Marquardt, C., & Wiedman, C. (2004a). How are earnings managed? An examination of specific accruals. *Contemporary Accounting Research, 21*, 461–491.
Marquardt, C., & Wiedman, C. (2004b). The effect of earnings management on the value relevance of accounting information. *Journal of Business, Finance and Accounting, 31*, 297–332.
McNichols, M. (2000). Research design issues in earnings management studies. *Journal of Accounting and Public Policy, 19*, 313–345.
Ota, K. (2011). Analysts' awareness of systematic bias in management earnings forecasts. *Applied Financial Economics, 21*, 1317–1330.
Ronen, J., & Yaari, V. (2008). Earnings management: Emerging insights in theory, practice, and research. Springer series in accounting scholarship, New York.
Shehata, M. (1991). Self selection bias and the economic consequences of accounting regulations: An application of two-stage switching regression to SFAS no. 2. *The Accounting Review, 66*, 768–787.
Watts, R. L., & Zimmerman, J. L. (1979). The demand for and supply of accounting theories: The market for excuses. *Accounting Review, 54*, 273–306.

A Comparative Study of Two Models SV with MCMC Algorithm

62

Ahmed Hachicha, Fatma Hachicha, and Afif Masmoudi

Contents

A. Hachicha (✉)
Department of Economic Development, Faculty of Economics and Management of Sfax, University of Sfax, Sfax, Tunisia
e-mail: hachicha.ahmed@fsegs.rnu.tn

F. Hachicha
Department of Finance, Faculty of Economics and Management of Sfax, Sfax, Tunisia
e-mail: hachicha_fatma@yahoo.fr

A. Masmoudi
Department of Mathematics, Faculty of Sciences of Sfax, Sfax, Tunisia
e-mail: Afif.Masmoudi@fss.rnu.tn

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_62,

Abstract

This paper examines two asymmetric stochastic volatility models used to describe the volatility dependencies found in most financial returns. The first is the autoregressive stochastic volatility model with Student's t-distribution (ARSV-t), and the second is the basic Svol of JPR (*Journal of Business and Economic Statistics 12*(4), 371–417, 1994). In order to estimate these models, our analysis is based on the Markov Chain Monte Carlo (MCMC) method. Therefore, the technique used is a Metropolis-Hastings (Hastings, *Biometrika 57*, 97–109, 1970), and the Gibbs sampler (Casella and George *The American Statistician 46*(3) 167–174, 1992; Gelfand and Smith, *Journal of the American Statistical Association 85*, 398–409, 1990; Gilks et al. 1993). The empirical results concerned on the Standard and Poor's 500 Composite Index (S&P), CAC 40, Nasdaq, Nikkei, and Dow Jones stock price indexes reveal that the ARSV-t model provides a better performance than the Svol model on the mean squared error (MSE) and the maximum likelihood function.

Keywords

Autoregression • Asymmetric stochastic volatility • MCMC • Metropolis-Hastings • Gibbs sampler • Volatility dependencies • Student's t-distribution • SVOL • MSE • Financial returns • Stock price indexes

62.1 Introduction

Stochastic volatility (SV) models are workhorses for the modelling and prediction of time-varying volatility on financial markets and are essential tools in risk management, asset pricing, and asset allocation. In financial mathematics and financial economics, stochastic volatility is typically modelled in a continuous time setting which is advantageous for derivative pricing and portfolio optimization. Nevertheless, since data is typically only observable at discrete points in time, in empirical applications, discrete-time formulations of SV models are equally important.

Volatility plays an important role in determining the overall risk of a portfolio and identifying hedging strategies that make the portfolio neutral with respect to market moves. Moreover, volatility forecasting is also crucial in derivatives trading.

Recently, SV models allowing the mean level of volatility to "jump" have been used in the literature; see Chang et al. (2007), Chib et al. (2002), and Eraker et al. (2002). The volatility of financial markets is a subject of constant analysis movements in the price of financial assets which directly affects the wealth of individual, companies, charities, and other corporate bodies. Determining whether there are any patterns in the size and frequency of such movements, or in their cause and effect, is critical in devising strategies for investments at the micro level and monetary stability at the macro level. Shephard and Pitt (1997) used improved and efficient Markov Chain Monte Carlo (MCMC) methods to estimate the volatility process "in block" rather than one point of time such as highlighted by Jacquier et al. (1994), for a simple SV model. Furthermore, Hsu and Chiao (2011) analyze the time patterns of individual analyst's relative accuracy ranking in earnings forecasts using a Markov chain model by treating two levels of stochastic persistence.

Least squares and maximum likelihood techniques have long been used in parameter estimation problems.

However, those techniques provide only point estimates with unknown or approximate uncertainty information. Bayesian inference coupled with the Gibbs sampler is an approach to parameter estimation that exploits modern computing technology. The estimation results are complete with exact uncertainty information. Section 62.2 presents the Bayesian approach and the MCMC algorithms. The SV model is introduced in Sect. 62.3, whereas empirical illustrations are given in Sect. 62.4.

62.2 The Bayesian Approach and the MCMC Algorithm

The Bayesian approach is a classical methodology where we assume that there is a set of unknown parameters. Alternatively, in the Bayesian approach the parameters are considered as random variables with given prior distributions. We then use observations (the likelihood) to update these distributions and obtain the posterior distributions.

Formally, let X = ($X_1, \ldots, X_T$) denote the observed data and θ a parameter vector:

$$P\left(\frac{\theta}{X}\right) \propto P\left(\frac{X}{\theta}\right) * P(\theta)$$

The posterior distribution $P(\theta/X)$ of a parameter $\theta/$ given the observed data X, where $P(X/\theta)$ denotes the likelihood distribution of X and $P(\theta)$ denotes the prior distribution of θ.

It would seem that in order to be as subjective as possible and to use the observations as much as possible, one should use priors that are non-informative. However, this can sometimes create degeneracy issues and one should choose a different prior for this reason. Markov Chain Monte Carlo (MCMC) includes the Gibbs sampler as well as the Metropolis-Hastings (M-H) algorithm.

62.2.1 The Metropolis-Hastings

The Metropolis-Hastings is the baseline for MCMC schemes that simulate a Markov chain $\theta^{(t)}$ with $P(\theta/Y)$ as the stationary distribution of a parameter θ given a stock price index X. For example, we can define θ_1, θ_2, and θ_3 such that $\theta = (\theta_1, \theta_2, \theta_3)$ where each θ_1 can be scalar, vectors, or matrices. MCMC algorithms are iterative, and so at iteration t we will sample in turn from the three conditional distributions. Firstly, we update θ_1 by drawing a value ${\theta_1}^{(t)}$ from $p(\theta_1/Y, \theta_2^{(t-1)}, \theta_3^{(t-1)})$. Secondly, we draw a value for ${\theta_2}^{(t)}$ from $p(\theta_2/Y, \theta_1^{(t)}, \theta_3^{(t-1)})$, and finally, we draw ${\theta_3}^{(t)}$ from $p(\theta_3/Y, \theta_1^{(t)}, \theta_2^{(t)})$.

We start the algorithm by selecting initial values, ${\theta_i}^{(0)}$, for the three parameters. Then sampling from the three conditional distributions in turn will produce a set of Markov chains whose equilibrium distributions can be shown to be the joint posterior distributions that we require.

Following Hastings (1970), a generic step from a M-H algorithm to update parameter θ_i at iteration t is as follows:

1. Sample θ_i^* from the proposal distribution $p_t(\theta_i/\theta_i^{(t-1)})$.
2. Calculate $f = p_t(\theta_i^{(t-1)}/\theta_i^*)/p_t(\theta_i^*/\theta_i^{(t-1)})$ which is known as the Hastings ratio and which equals 1 for symmetric proposals as used in pure Metropolis sampling.
3. Calculate $s_t = fp(\theta_i^*/Y, \phi_i)/p(\theta_i^{(t-1)}/Y, \phi_i)$, where ϕ_i is the acceptance ratio and gives the probability of accepting the proposed value.
4. Let $\theta_i^{(t)} = \theta_i^*$ with probability min(1, s_t); otherwise let $\theta_i^{(t)} = \theta_i^{(t-1)}$.

A popular and more efficient method is the acceptance-rejection (A-R) M-H sampling method which is available. Whenever the target densities are bounded by a density from which it is easy to sample.

62.2.2 The Gibbs Sampler

The Gibbs sampler (Casella and Edward 1992; Gelfand and Smith 1990; Gilks et al. 1992) is the special M-H algorithm whereby the proposal density for updating θ_j equals the full conditional $p(\theta_j^*/\theta_j)$ so that proposals are acceptance with probability 1.

The Gibbs sampler involves parameter-by-parameter or block-by-block updating, which when completed from the transaction from $\theta^{(t)}$ to $\theta^{(t+1)}$:

1. $\theta_1^{(t+1)} \approx f_1(\theta_1/\theta_2^t, \theta_3^{(t)}, \ldots \theta_D^{(t)})$
2. $\theta_2^{(t+1)} \approx f_2(\theta_2/\theta_1^{t+1}, \theta_3^{(t)}, \ldots \theta_D^{(t)})$

.

.

.

.

D. $\theta_D^{(t+1)} \approx f_D(\theta_D/\theta_1^{t+1}, \theta_2^{(t+1)}, \ldots \theta_{D-1}^{(t+1)})$

Repeated sampling from M-H samplers such as the Gibbs samplers generates an autocorrelated sequence of numbers that, subject to regularity condition (ergodicity, etc.), eventually "forgets" the starting values $\theta^0 = (\theta_1^{(0)}, \theta_2^{(0)}, \ldots\ldots, \theta_D^{(0)})$ used to initialize the chain and converges to a stationary sampling distribution $p(\theta/y)$.

In practice, Gibbs and M-H algorithms are often combined, which results in a "hybrid" MCMC procedure.

62.3 The Stochastic Volatility Model

62.3.1 Autoregressive SV Model with Student's Distribution

In this paper, we will consider the pth order ARSV-t model, ARSV(p)-t, as follows:

$$\begin{cases} Y_t = \sigma\xi\exp(V_t/2) \\ V_t = \phi_1 V_{t-1} + \ldots. + \phi_p V_{t-p} + \eta_{t-1} \end{cases}$$

$$\xi_t = \frac{\varepsilon_t}{\sqrt{\kappa_t/(v-2)}}$$

$$\kappa_t \approx \chi^2(v)$$

where κ_t is independent of (ε_t, η_t), Y_t is the stock return for market indexes, and V_t is the log-volatility which is assumed to follow a stationarity AR(p) process with a persistent parameter $|\phi| \prec 1$. By this specification, the conditional distribution, ξ_t, follows the standardized t-distribution with mean zero and variance one. Since κ_t is independent of (ε_t, η_t), the correlation coefficient between ξ_t and η_t is also ρ.

If $\phi \approx N(0, 1)$, then

$$\phi_1 = \frac{\left(\sum_{t=1}^{T} V_t V_{t-1}\right) - \phi_2 \left(\sum_{t=1}^{T} V_{t-1} V_{t-2}\right) + \overline{\phi_1}}{\left(\sum_{t=1}^{t} V_{t-1}^2\right) - 1}$$

and

$$\phi_2 = \frac{\left(\sum_{t=2}^{T} V_t V_{t-2}\right) - \phi_1 \left(\sum_{t=2}^{T} V_{t-1} V_{t-2}\right) + \overline{\phi_2}}{\left(\sum_{t=2}^{T} V_{t-2}^2\right) - 1}$$

The conditional posterior distribution of the volatility is given by

$$p(V/\Theta, Y) \propto e^{\left(\frac{1}{2*\sigma^2}\left(\sum_{t=1}^{T} Y_t^2 e^{-V_t}\right) - \frac{1}{2}\sum_{t=1}^{T} (V_t - \phi_1 V_{t-1} - \phi_2 V_{t-2})^2 - \frac{1}{2}\sum_{t=1}^{T} (V_{t+1} - \phi_1 V_t - \phi_2 V_{t-2})^2\right)}$$

The representation of the SV-t model in terms of a scale mixture is particularity useful in a MCMC context since it allows for sampling a non-log-concave sampling problem into a log-concave one. This allows for sampling algorithms which guarantee convergence in finite time (see Frieze et al. 1994). Allowing log returns to be student-t-distributed naturally changes the behavior of the stochastic volatility process; in the standard SV model, large value of $|Y_t|$ induces large value of the V_t.

62.3.2 Basic Svol Model

Jacquier, Polson, and Rossi (1994), hereafter JPR, introduced Markov chain technique (MCMC) for the estimation of the basic Svol model with normally distributed conditional errors:

$$\begin{cases} Y_t = \sqrt{V_t}\,\varepsilon_t^s \\ \log(V_t) = \alpha + \delta \log(V_{t-1}) + \sigma_v\,\varepsilon_t^v \end{cases}$$
$$\left(\varepsilon_t^s, \varepsilon_t^v\right) \approx N(0, I_2)$$

Let $\Theta = (\alpha, \delta, \sigma_v)$ be the vector of parameters of the basic SVOL, and $V = (V_t)_{t=1}^{T}$, where α is the intercept. The parameter vector consists of a location α, a volatility persistence δ, and a volatility of volatility σ_v.

The basic Svol specifies zero correlation, the errors of the mean, and variance equations.

Briefly, the Hammersley-Clifford theorem states that having a parameter-set Θ, a state V_t, and an observation Y_t, we can obtain the joint distribution $p(\Theta, V/Y)$ from p(Θ, V/Y) and $p(V/\Theta, Y)$, under some mild regularity conditions. Therefore by applying the theorem iteratively, we can break a complicated multidimensional estimation problem into many sample one-dimensional problems.

Creating a Markov chain $\Theta^{(t)}$ via a Monte Carlo process, the ergodic averaging theorem states that the time average of a parameter will converge towards its posterior mean.

The formula of Bayes factorizes the posterior distribution likelihood function with prior hypotheses:

$$P(\Theta, V/Y)\alpha P(Y/V, \Theta)P(V/\Theta)P(\Theta)$$

where α is the intercept, δ the volatility persistence, and σ_v is the standard deviation of the shock to log V_t.

We use a normal-gamma prior, so, the parameters α, $\delta \approx N$, and $\sigma_v^2 \approx IG$, (Appendix 1)

Then

$$P(\alpha, \delta/\sigma_v, V, Y) \approx \prod P(V_t/V_{t-1}, \alpha, \delta, \sigma_v)P(\alpha, \delta)\alpha N$$

And for σ_v, we obtain

$$P\left(\sigma^2/\alpha, \sigma_v, V, Y\right)\alpha \prod P(V_t/V_{t-1}, \alpha, \delta, \sigma_v)P\left(\sigma_v^2\right)\alpha IG$$

62.4 Empirical Illustration

62.4.1 The Data

Our empirical analysis focuses on the study of five international financial indexes: the Dow Jones Industrial, the Nikkei, the CAC 40, the S&P500, and the Nasdaq. The indexes are compiled and provided by Morgan Stanley Capital International. The returns are defined as $y_t = 100 * (\log S_t - \log S_{t-1})$. We used the last 2,252 observations for all indexes except the Nikkei, when we have only used 2,201 observations due to lack of data. The daily stock market indexes are for five different countries over the period 1 January 2000 to 31 December 2008.

Table 62.1 reports the mean, standard deviation, median, and the empirical skewness as well as kurtosis of the five series. All series reveal negative skewness and overkurtosis which is a common finding of financial returns.

62.4.2 Estimation of SV Models

The standard SV model is estimated by running the Gibbs and A-R M-H algorithm based on 15,000 MCMC iterations, where 5,000 iterations are used as burn-in period.

Tables 62.2 and 62.3 show the estimation results in the basic Svol model and the SV-t model of the daily indexes. α and δ are independent priors.

The prior in δ is essentially flat over [0, 1]. We impose stationarity for $\log(V_t)$ by truncating the prior of δ. Other priors for δ are possible.

Geweke (1994a, b) proposes alternative priors to allow the formulation of odds ratios for non-stationarity. Whereas Kim et al. (1998) center an informative Beta prior around 0.9.

Table 62.1 Summary statistics for daily returns

	Mean	SD	Median	Skewness	Kurtosis
CAC 40	3.7E-04	0.013	5.0e-4	−0.295	5.455
Dow Jones	2.8e-04	0.015	4.0e-4	−0.368	4.522
Nasdaq	2.5e-04	0.014	5.5e-4	−0.523	6.237
Nikkei	3.5e-04	0.005	3.2e-4	−0.698	3.268
S&P	2.8e-04	0.008	4.5e-4	−0.523	5.659

Table 62.2 Estimation results for the Svol model

	CAC 40	Dow Jones	Nasdaq	Nikkei	S&P
σ	0.4317(0.0312)	0.4561(0.0421)	0.5103(0.0393)	0.5386(0.0523)	0.4435(0.0623)
α	−0.1270(0.0421)	0.0059(0.0534)	0.1596(0.0332)	0.1966(0.0493)	−0.1285(0.0593)
δ	−0.7821(0.0621)	0.0673(0.0317)	0.6112(0.0429)	0.8535(0.0645)	0.7224(0.0423)

Table 62.3 Estimation results for the SV-t model

	CAC 40	Dow Jones	Nasdaq	Nikkei	S&P
Φ_1	0.4548(0.0037)	0.40839(0.0021)	0.5225(0.0065)	0.4348(0.0059)	0.2890(0.0046)
Φ_2	0.5544(0.1524)	0.6437(0.1789)	0.4473(0.1326)	0.4865(0.1628)	0.6133(0.1856)
σ	0.0154(0.0294)	0.0205(0.0367)	0.0131(0.0524)	0.0148(0.0689)	0.0135(0.0312)
ρ	−0.02191(0.0625)	−0.0306(0.0346)	−0.0489(0.0498)	−0.0751(0.0255)	−0.0235(0.0568)

Table 62.2 shows the results for the daily indexes. The posterior of δ are higher for the daily series. The highest means are 0.782, 0.067, 0.611, 0.85, and 0.722, for the full sample Nikkei.

This result is not a priori curious because the model of Jacquier et al. (1994) can lead to biased volatility forecast.

Well, as the basic SVOL, there is no apparent evidence of unit of volatility. There are other factors that can deflect this rate such exchange rate (O'Brien and Dolde 2000).

Deduced from this model, against the empirical evidence, positive and negative shocks have the same effect in volatility.

Table 62.3 shows the Metropolis-Hastings estimates of the autoregressive SV model.

The estimates of ϕ are between 0.554 and 0.643, while those of σ are between 0.15 and 0.205.

Against, the posterior of ϕ for the SV-t model are located higher.[1] This is consistent with temporal aggregation (as suggested by Meddahi and Renault 2000). This result confirms the typical persistence reported in the GARCH literature. After the result, the first volatility factors have higher persistence, while the small values of Φ_2 indicate the low persistence of the second volatility factors.

The second factor Φ_2 plays an important role in the sense that it captures extreme values, which may produce the leverage effect, and then it can be considered conceivable.

The estimates of ρ are negative in most cases. Another thing to note is that these estimates are relatively higher than that observed by Asai et al. (2006) and Manabu Asai (2008). The estimated of ρ for index S&P using Monte Carlo simulation is −0.3117, then it is −0.0235 using Metropolis-Hasting. This implies that for each data set, the innovations in the mean and volatility are negatively correlated.

Negative correlations between mean and variance errors can produce a "leverage" effect in which negative (positive) shocks to the mean are associated with increases (decreases) in volatility.

The return of different indexes not only is affected by market structure (Sharma 2011) but also is deeply influenced by different crises observed in international market, i.e., the Asian crises detected in 1987 and the Russian one in 2002.

[1] We choose $p = 2$ because if $p = 1$ and $v \rightarrow \infty$, the ARSV-t model declined to the asymmetric SV model of Harvey and Shephard (1996).

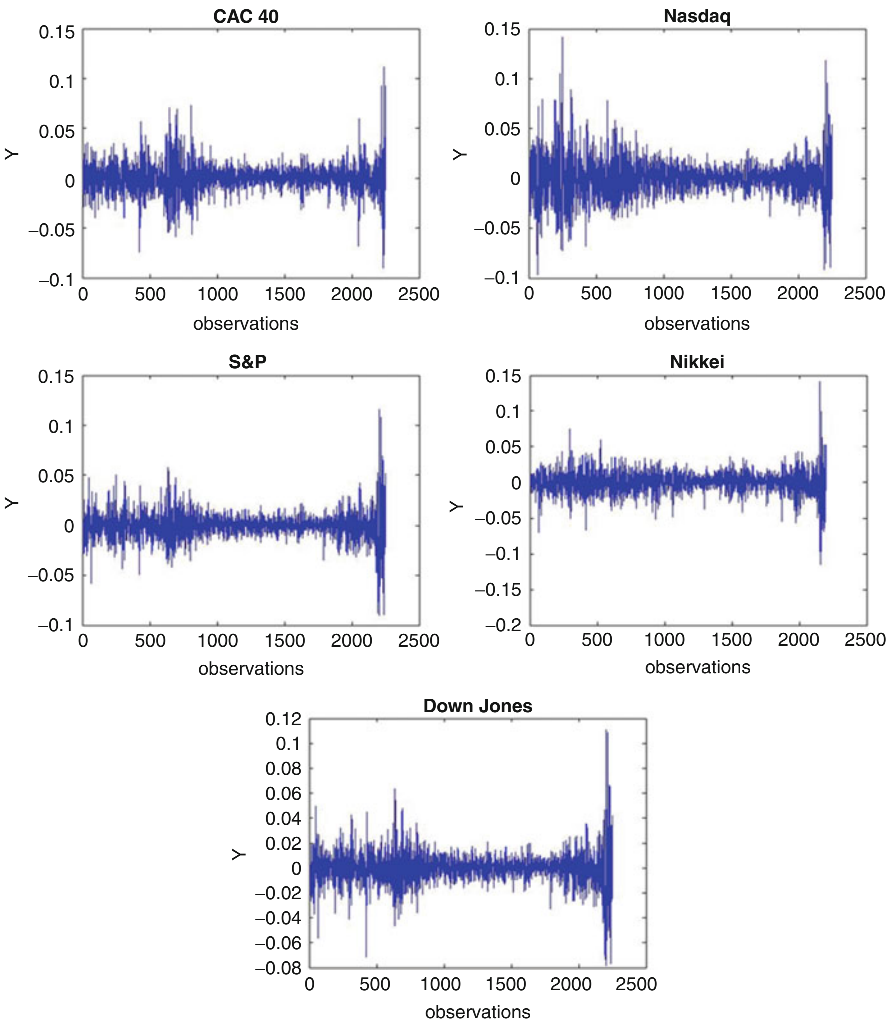


Fig. 62.1 Return for indexes

The markets in our sample are subject to several crises that directly affect the evolution of the return indexes. The event of 11 September 2002 the Russian crisis and especially the beginning of the subprime crisis in the United States in July 2007 justify our results. These results explored in Fig. 62.1 suggest that periods of market crisis or stress increase the volatility. Then the volatility at time (t) depends on the volatility at $(t-1)$ (Engle 1982).

When the new information comes in the market, it can be disrupted and this affects the anticipation of shareholders for the evolution of the return.

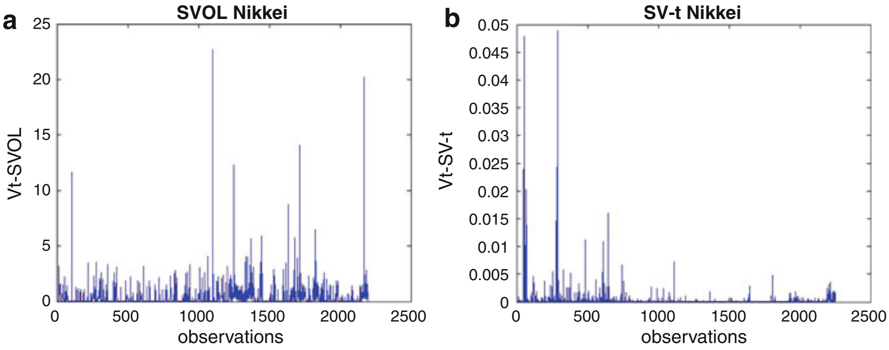


Fig. 62.2 Smoothed estimates of Vt, basic SVOL, and SV-t model

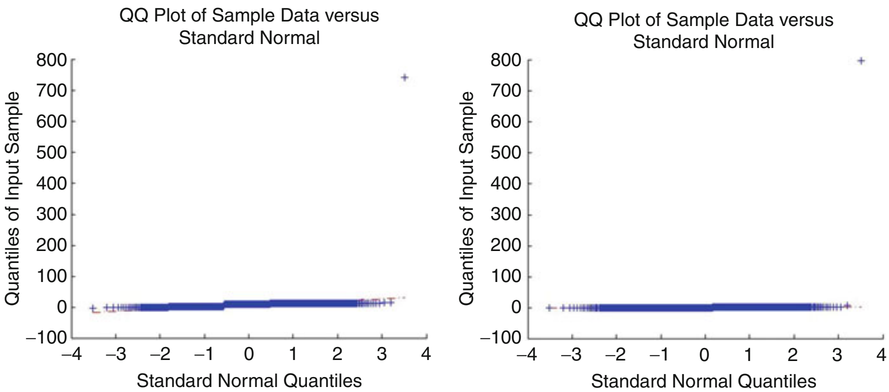


Fig. 62.3 QQ plot of normalized innovation based on the basic Svol model (*left*) and the SV-t model (*right*)

The resulting plots of the smoothed volatilities are shown in Fig. 62.2. We take our analysis in the Nikkei indexes, but the others are reported in Appendix 2.

The convergence is very remarkable for the Nikkei, like Dow Jones, Nasdaq, and the CAC 40 indexes. This enhances the idea that the algorithm used for estimated volatility is a good choice.

The basic Svol model mis-specified can induce substantial parameter bias and error in inference about V_t; Geweke (1994a, b) showed that the basic Svol has the same problem with the largest outlier, October 1987 "Asiatique crisis." The V_t for the model Svol reveal a big outlier on period crises.

The corresponding plots of innovation are given by Fig. 62.3 for two models basic Svol and SV-t for Nikkei indexes. Appendix 3 shows the QQ plot for the other indexes, respectively, for the Nasdaq, S&P, Dow Jones, and CAC 40 for the two models. The standardized innovation reveals a big outlier when the market in stress (Hwang and Salmon 2004).

Table 62.4 MSE and likelihood for two models

	CAC 40		Dow Jones		Nasdaq		Nikkei		S&P	
	SVOL	SV-t	SVOL	SV-t	SVOL	SV-t	SVOL	SV-t	SVOL	SV-t
MSE	0.0.210	0.0296	0.1277	0.0040	0.2035	0.0229	0.0241	0.0168	0.0248	0.0395
Likelihood	$-2.595.10^{-4}$	−0.0046	−0.0374	−0.0012	−0.2257	−0.0158	$-0.2257.10^{-4}$	−0.0094	−0.0054	0.02016

The advantages of asymmetric basic SV is able to capture some aspects of financial market and the main properties of their volatility behavior (Danielsson 1994; Chaos 1991; Eraker et al. 2000).

It is shown that the inclusion of student-t errors improves the distributional properties of the model only slightly. Actually, we observe that basic Svol model is not able to capture extreme observation in the tail of the distribution. In contrast, the SV-t model turns out to be more appropriate to accommodate outliers. The corresponding plot of innovation for the basic model is unable to capture the distribution properties of the returns. This is confirmed by the Jarque-Bera normality test and the QQ plot revealing departures from normality, mainly stemming from extreme innovation.

Finally, in order to detect which of the two models is better, we opt for two indicators of performance, such as the likelihood and the MSE. Likelihood is a function of the parameters of the statistical model that plays a preponderant role in statistical inference. MSE is called squared error loss, and it measures the average of the square of "error." Table 62.4 reveals the results for this measure and indicates that the SV-t model is much more efficient than the other. Indeed, in terms of comparison, we are interested in the convergence of two models. We find that convergence to the SV-t model is fast.

Table 62.4 shows the performance of the algorithm and the consequence of using the wrong model on the estimates of volatility. The efficiency is at 60 %.

The MCMC is more efficient for all parameters used in these two models. In a certain threshold, all parameters are stable and converge to a certain level. Appendices 4 and 5 show that the α, δ, σ, ϕ converge and stabilize; this shows the power for MCMC.

The results for both simulated show that the algorithm of SV-t model is fast and converges rapidly with acceptable levels of numerical efficiency. Then, our sampling provides strong evidence of convergence of the chain.

62.5 Conclusion

We have applied these MCMC methods to the study of various indexes. The ARSV-t models were compared with Svol models of Jacquier et al. (1994) models using the S&P, Dow Jones, Nasdaq, Nikkei, and CAC 40.

The empirical results show that SV-t model can describe extreme values to a certain extent, but it is more appropriate to accommodate outliers. Surprisingly, we have frequently observed that the best model is the Student's t-distribution (ARSV-t) with their forecast performance. Our result confirms the finding from Manabu Asai (2008), who indicates, first, that the ARSV-t model provides a better fit than the MFSV model and, second, the positive and negative shocks do not have the same effect in volatility. Our result proves the efficiency of Markov chain for our sample and the convergence and stability for all parameters to a certain level. This paper has made certain contributions, but several extensions are still possible. To find the best results, opt for extensions of SVOL.

Appendix 1

The posterior volatility is

$$P(V/\Theta, V) \propto P(Y/\Theta, V)P(V/\Theta) \propto \prod_{t=1}^{T} P(V_t/V_{t-1}, V_{t+1}, \Theta, Y_t)$$

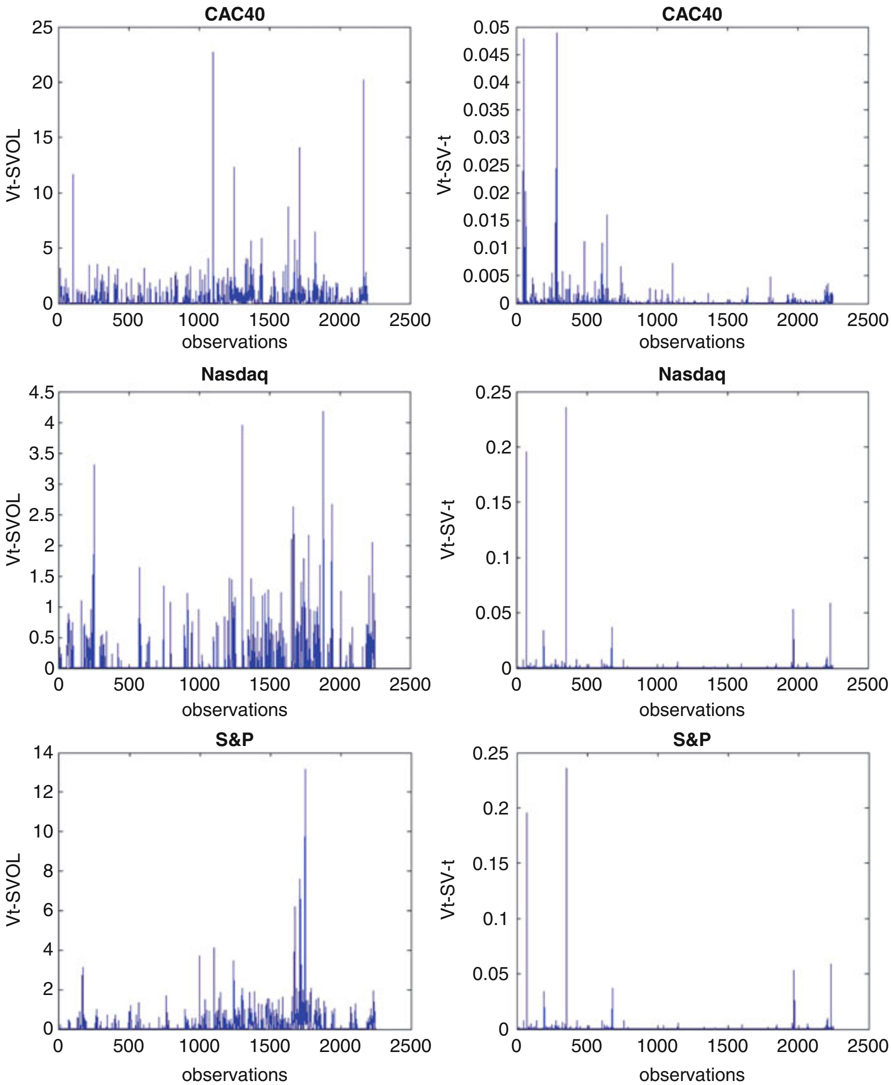


Fig. 62.4 (continued)

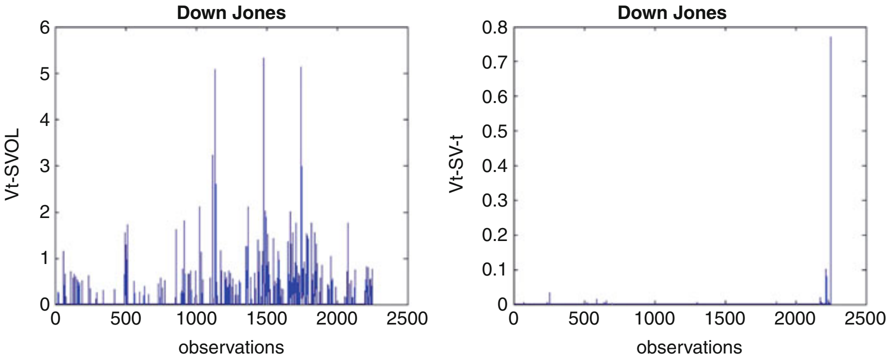


Fig. 62.4 Smoothed estimates of Vt, basic SVOL, and SV-t model

with

$$P(V/V_{t-1}, V_{t+1}, \Theta, Y_t) \propto P(Y_t/V_t, \Theta)P(V_t/V_{t-1}, \Theta)P(V_{t+1}/V_t, \Theta)$$

A simple calculation shows that

$$\prod(V_t) = P(V_t/V_{t-1}, V_{t+1}, \Theta, Y_t) \propto \frac{1}{V_t^{0.5}} \exp\left(\frac{-Y_t^2}{2V_t}\right) \frac{1}{V_t} \exp\left(-\frac{(\log V_t - \mu_t)^2}{2\sigma^2}\right)$$

with

$$\mu_t = \frac{\alpha(1-\beta) + \beta(\log V_{t+1} + \log V_{t-1})}{1+\beta^2}$$

and

$$\sigma^2 = \frac{\sigma_v^2}{1+\beta^2}$$

The MCMC algorithm consists of the following steps:

$$P(\alpha, \delta/\sigma_v, V, Y) \approx N$$

$$P\left(\sigma_v^2/\alpha, \delta, V, Y\right) \approx IG$$

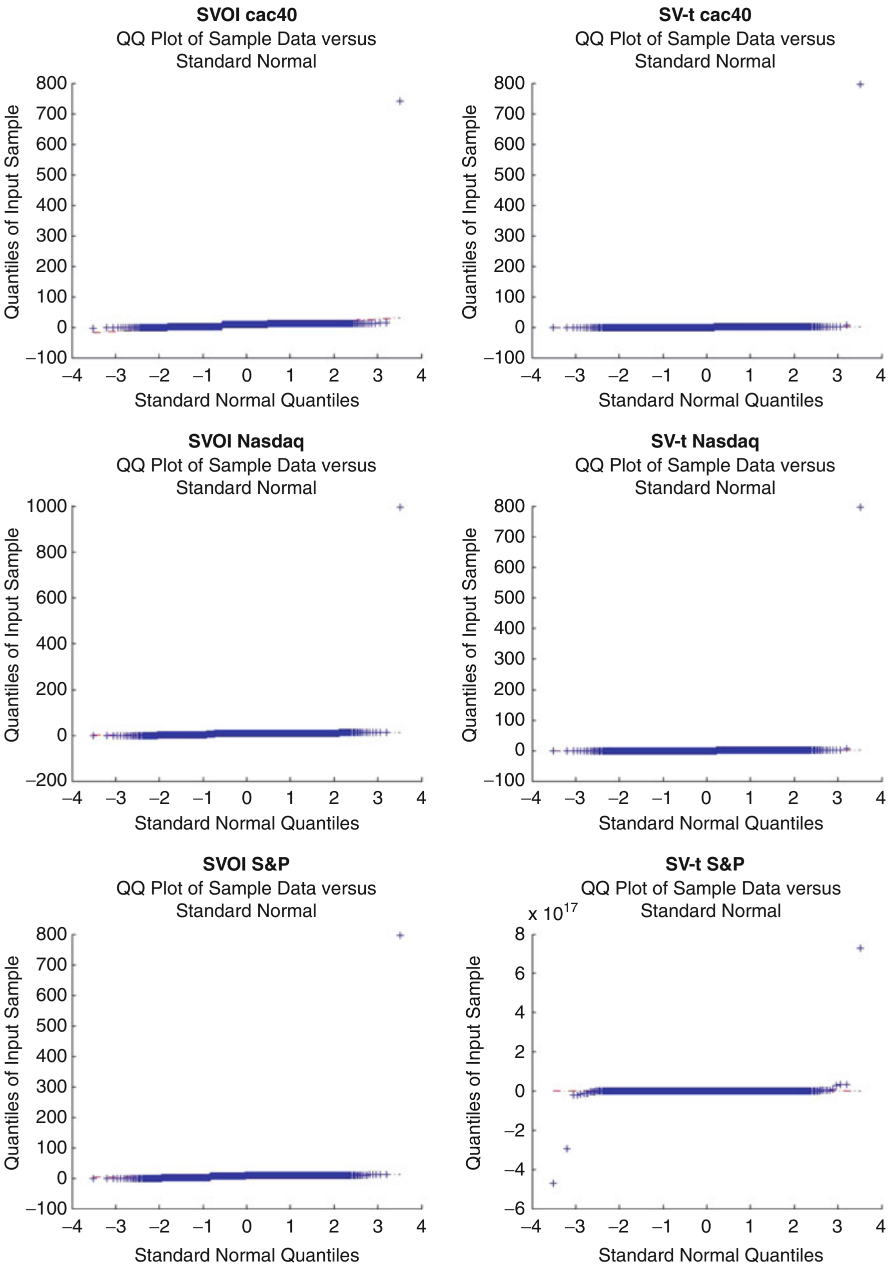


Fig. 62.5 (continued)

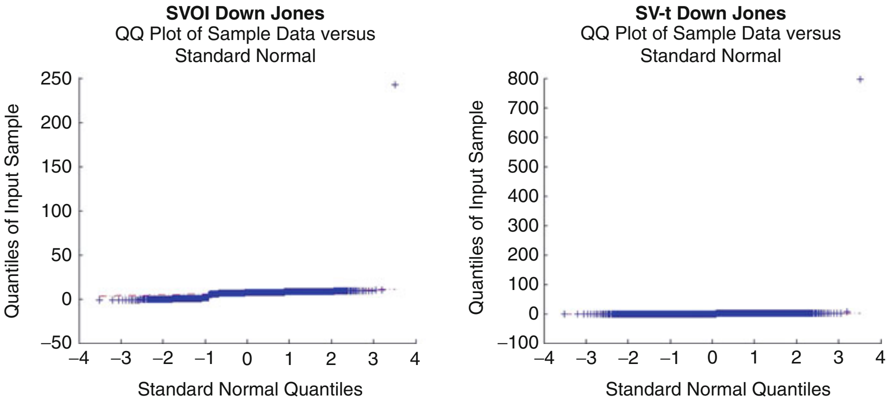


Fig. 62.5 Smoothed estimates of Vt, basic SVOL, and SV-t model

$$P(V_t/V_{t-1}, V_{t+1}, \Theta, Y_t) : \text{Metropolis-Hastings}$$

An iteration (j),

$$\alpha^{(j)} = \frac{\sum_{t=1}^{T} \log V_t^{(j-1)} - \beta^{(j-1)} \sum_{t=1}^{T} \log V_{t-1}^{(j-1)}}{\left(\sigma_v^2\right)^{(j-1)} + T}$$

By following the same approach, the estimator δ at step (j) is given by

$$\delta^{(j)} = \frac{\sum_{t=1}^{T} \left[\log V_{t-1}^{(j-1)} \left(\log V_t^{(j-1)} - \alpha^{(j)}\right)\right]}{\left(\sigma_v^2\right)^{(j-1)} + \sum_{t=1}^{T} \left(\log V_{t-1}^{(j-1)}\right)^2}$$

For parameter σ_v^2, the prior density is an inverse gamma (IG (a, b)). The expression of the estimator parameter σ_v^2 at step (j) is given by

$$\left(\sigma_v^2\right)^{(j)} = \frac{\frac{1}{2} \sum_{t=1}^{T} \left(\log V_t^{(j-1)} - \alpha^{(j)} - \delta^{(j)} \log V_{t-1}^{(j-1)}\right)^2 + b}{T/2 + a - 1}$$

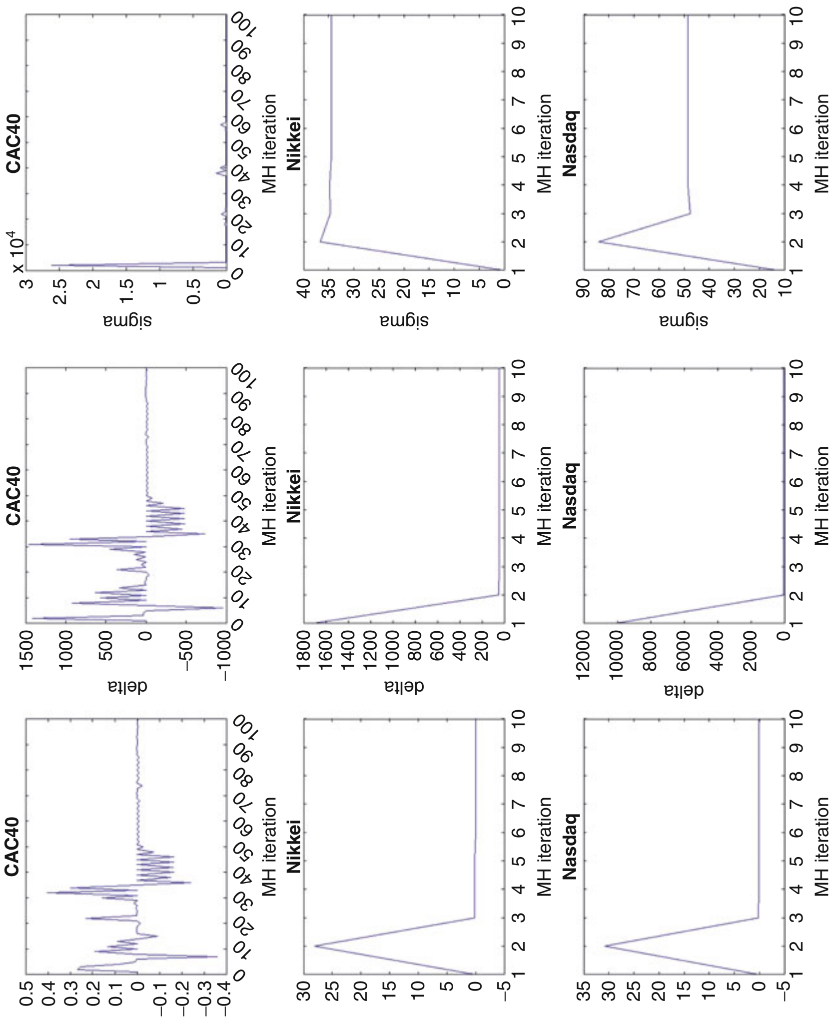


Fig. 62.6 (continued)

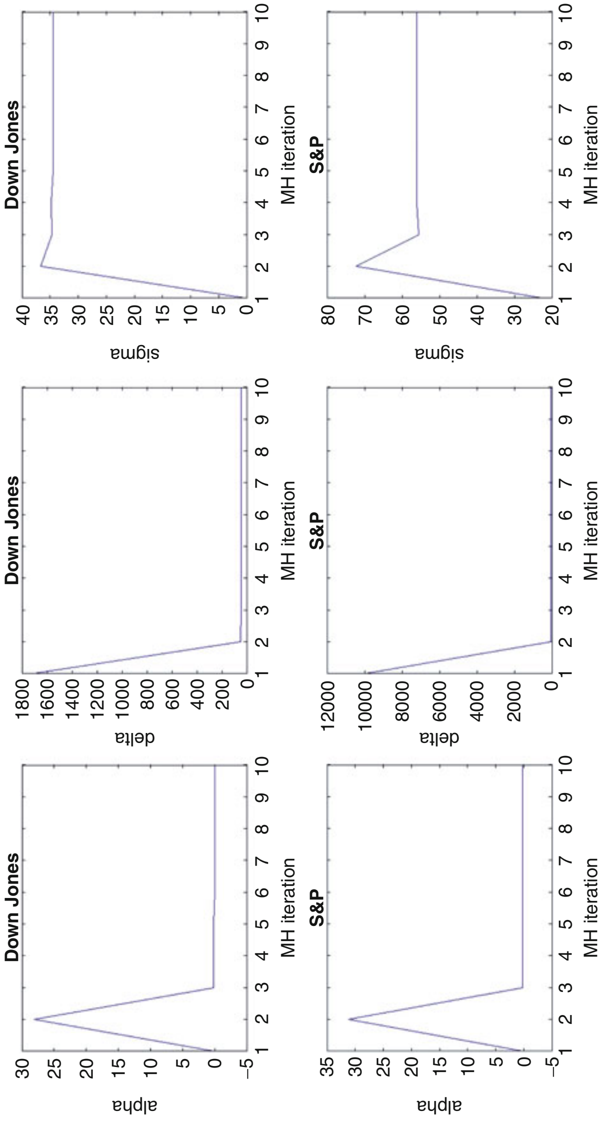


Fig. 62.6 Behavioral of parameters of basic Svol model

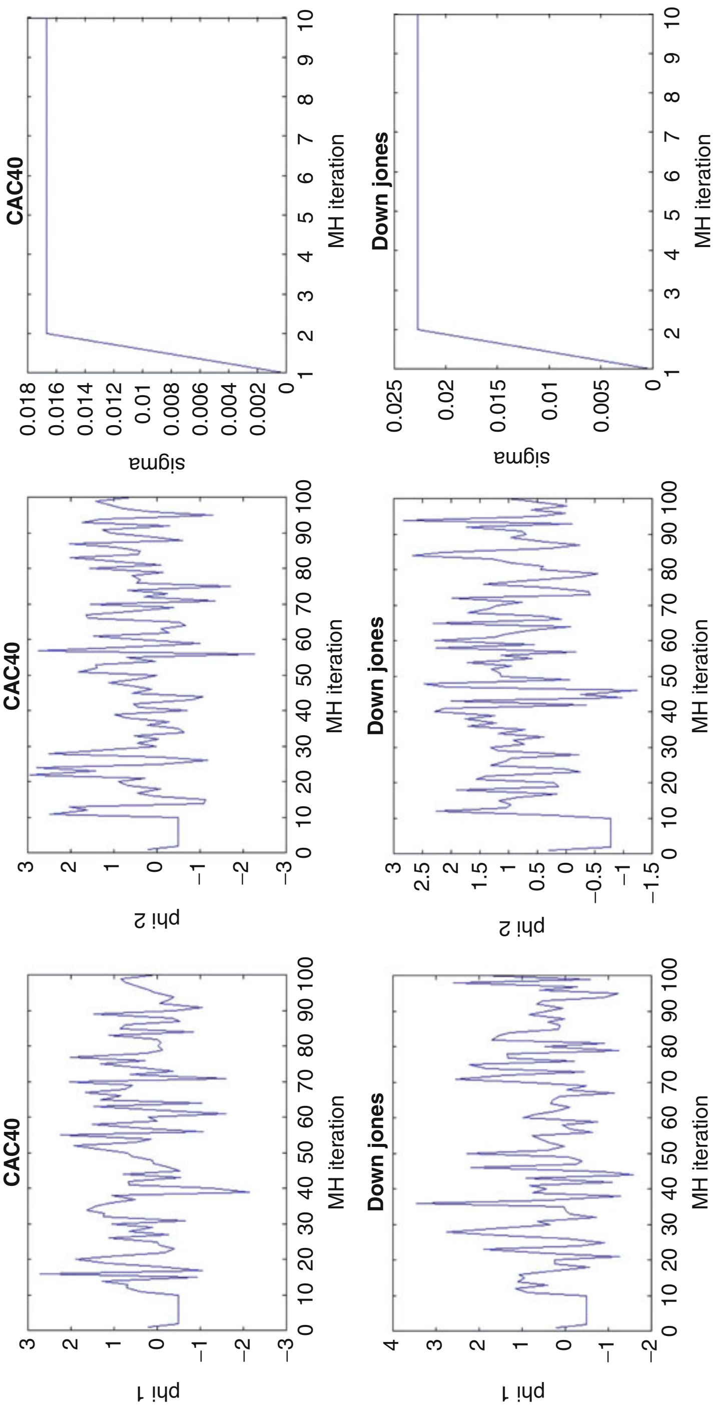


Fig. 62.7 (continued)

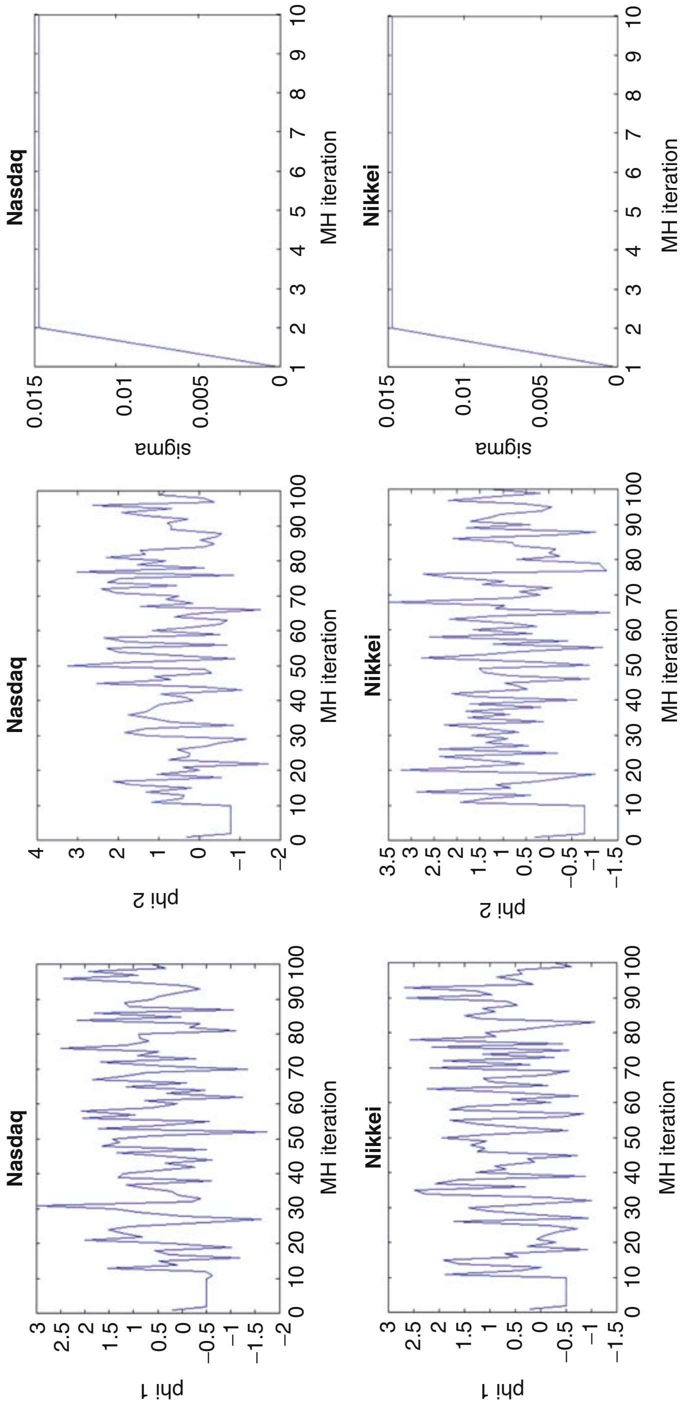


Fig. 62.7 Behavioral of parameters of SV-t model

Appendix 2

See Fig. 62.4

Appendix 3

See Fig. 62.5

Appendix 4

See Fig. 62.6

Appendix 5

See Fig. 62.7

References

Asai, M. (2008). Autoregressive stochastic volatility models with heavy-tailed distributions: A comparison with multifactors volatility models. *Journal of Empirical Finance, 15*, 322–345.

Asai, M., McAleer, M., & Yu, J. (2006). Multivariate stochastic volatility. *Econometric Reviews, 25*(2–3), 145–175.

Casella, G., & Edward, G. (1992). Explaining the Gibbs sampler. *The American Statistician, 46*(3), 167–174.

Chang, R., Huang, M., Lee, F., & Lu, M. (2007). The jump behavior of foreign exchange market: Analysis of Thai Baht. *Review of Pacific Basin Financial Markets and Policies, 10*(2), 265–288.

Chib, S., Nardari, F., & Shephard, N. (2002). Markov chain Monte Carlo methods for stochastic volatility models. *Journal of Econometrics, 108*, 281–316.

Danielsson, J. (1994). Stochastic volatility in asset prices: Estimation with simulated maximum likelihood. *Journal of Econometrics, 64*, 375–400.

Engle, R. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica, 50*, 987–1007.

Eraker, B., Johannes, M., & Polson, N. (2000). The impact of jumps in volatility and return. Working paper, University of Chicago.

Frieze, A. M., Kannan, R., & Polson, N. (1994). Sampling from log-concave distributions. *Annals of Applied Probability, 4*, 812–837.

Gelfand, A. E., & Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association, 85*, 398–409.

Geweke, J. (1994a). Priors for macroeconomic time series and their application. *Econometric Theory, 10*, 609–632.

Geweke, J. (1994b) Bayesian comparison of econometric models. Working paper, Federal Reserve Bank of Minneapolis Research Department.

Gilks, W. R., & Wild, P. (1992). Adaptive rejection sampling for Gibbs sampling. *Journal of the Royal Statistical Society, Series C, 41*(41), 337.

Harvey, A. C., & Shephard, N. (1996). Estimation of an asymmetric stochastic volatility model for asset returns. *Journal of Business and Economic Statistics, 14*, 429–434.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika, 57*, 97–109.

Hsu, D., & Chiao, C. H. (2011). Relative accuracy of analysts' earnings forecasts over time: A Markov chain analysis. *Review of Quantitative Finance and Accounting, 37*(4), 477–507.

Hwang, S., & Salmon, M. (2004). Market stress and herding. *Journal of Empirical Finance, 11*, 585–616.

Jacquier, E., Polson, N., & Rossi, P. (1994). Bayesian analysis of stochastic volatility models (with discussion). *Journal of Business and Economic Statistics, 12*(4), 371–417.

Kim, S. N., Shephard, N., & Chib, S. (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models. *Review of Economic Studies, 65*, 365–393.

Meddahi, N., & Renault, E. (2000). Temporal aggregation of volatility models. Document de Travail CIRANO, 2000-22.

O'Brien, T. J., & Dolde, W. (2000). A currency index global capital asset pricing model. *European Financial Management, 6*(1), 7–18.

Sharma, V. (2011). Stock return and product market competition: Beyond industry concentration. *Review of Quantitative Finance and Accounting, 37*(3), 283–299.

Shepherd, P. (1997). Likelihood analysis of non-Gaussian measurement time series. *Biometrika, 84*(3), 653–667.

Internal Control Material Weakness, Analysts Accuracy and Bias, and Brokerage Reputation

63

Li Xu and Alex P. Tang

Contents

L. Xu (✉)
Washington State University, Richland, WA, USA
e-mail: li.xu@tricity.wsu.edu

A.P. Tang
Morgan State University, Baltimore, MD, USA
e-mail: Alex.Tang@morgan.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_63,

Abstract

We examine the impact of internal control material weaknesses (ICMW hereafter) on sell-side analysts. Using matched firms, we find that ICMW reporting firms have less accurate analyst forecasts relative to non-reporting firms when the reported ICMWs belong to the Pervasive type. ICMW reporting firms have more optimistically biased analyst forecasts compared to non-reporting firms. The optimistic bias exists only in the forecasts issued by the analysts affiliated with less highly reputable brokerage houses. The differences in accuracy and bias between ICMW and non-ICMW firms disappear when ICMW disclosing firms stop disclosing ICMWs. Collectively, our results suggest that the weaknesses in internal control increase forecasting errors and upward bias for financial analysts. However, a good brokerage reputation can curb the optimistic bias.

We use the Ordinary Least Squares (OLS) methodology in the main tests to examine the impact of internal control material weaknesses (ICMW hereafter) on sell-side analysts. We match our ICMW firms with non-ICMWs based on industry, sales, and assets. We reestimate the models using rank regression technique to assess the sensitivity of the results to the underlying functional form assumption made by OLS. We use Cook's distance to test the outliers.

Keywords

Internal control material weakness • Analyst forecast accuracy • Analyst forecast bias • Brokerage reputation • Sarbanes-Oxley act • Ordinary least squares regressions • Rank regressions • Fixed effects • Matching procedure • Cook's distance

63.1 Introduction

As part of the Sarbanes-Oxley Act of 2002 (SOX), SEC registrants' executives are now required to certify that they have evaluated the effectiveness of their internal controls over financial reporting (Section 302, effective in August 2002) and to provide an annual report to assess the effectiveness of the internal control structure and procedures (Section 404, effective in November 2004).[1,2] These assessments of internal control requirements have arguably been the most controversial aspect of SOX. On the one hand, many firms complain that internal control problems are inconsequential for financial statement users, and hence the high compliance costs of assessing internal control are not justified (Solomon 2005). On the other hand, a growing chorus of investors claims that good internal controls result in much more reliable corporate financial statements, which benefit financial statement users by reducing their information collection and interpretation costs.[3]

In this paper we test whether internal control problems are inconsequential for financial statement users by examining the impact of internal control material weaknesses (ICMWs hereafter) on one group of the most important financial statement users – sell-side analysts. We examine the association between the analyst forecast accuracy and bias and the disclosed ICMWs under SOX Sections 302 and 404 and how the brokerage reputation will influence this association. Additionally, we investigate the impact of the types and severity of ICMWs on forecast accuracy.

Based on a sample of 727 firms that have disclosed ICMWs since August of 2002, we find that the analysts' forecasts are less accurate among ICMW reporting firms relative to matched non-reporting firms. When we classify ICMW reporting firms into Pervasive or Contained ICMW reporting firms, the accuracy is significantly lower among Pervasive ICMW reporting firms. Our findings suggest that Pervasive ICMWs significantly increase the complexity of the forecasting task for analysts. In contrast, Contained ICMWs *alone* do not significantly increase the complexity of the forecasting task for analysts. When we investigate the association between analyst forecast bias and ICMWs,

[1]Key points of Section 302 include the following: (1) The signing officers must certify that they are responsible for establishing and maintaining internal control and have designed such internal controls to ensure that material information relating to the registrants and its consolidated subsidiaries is made known to such officers by others within those entities, particularly during the period in which the periodic reports are being prepared. (2) The officers must query "have evaluated the effectiveness of the registrant's internal controls" as of a date within 90 days prior to the report and have presented in the report their conclusions about the effectiveness of their internal controls based on their evaluation as of that date.

[2]Key points of Section 404 include the following: (1) Management is required to produce an internal control report as part of each annual Exchange Act report. (2) The report must affirm the responsibility of management for establishing and maintaining an adequate internal control structure and procedures for financial reporting. (3) The report must also contain an assessment, as of the end of the most recent fiscal year of the registrant, of the effectiveness of the internal control structure and procedures of the issuer for financial reporting. (4) External auditors are required to issue an opinion on whether effective internal control over financial reporting was maintained in all material respects by management. This is in addition to the financial statement opinion regarding the accuracy of the financial statements.

[3]For example, Donald J. Peters, a portfolio manager at T. Rowe Price Group, says: "The accounting reforms [of SOX] have been a win. It is [now] much easier for financial statement users to have a view of the true economics" of a company (Wall Street Journal, January 29, 2007).

we find that the analysts' forecasts are more optimistically biased toward ICMW reporting firms compared to matched non-reporting firms. Taken together, our overall findings on accuracy and bias suggest that ICMWs add a unique dimension to forecasting complexity.

In addition, we separate the analysts' brokerage houses into two groups: highly reputable and less highly reputable brokerage houses. Highly reputable brokerage houses value analysts' reports more than less highly reputable brokerage houses. Hence, analysts are likely to feel constrained from adding an arbitrarily high optimistic bias to their private estimates for fear of hurting the brokerage house's reputation. In addition, more reputable brokerage firms tend to spend significant resources in collecting information, and thus the access to firms' private information is relatively less important for analysts from highly reputable brokerage houses (compared to those from less highly reputable brokerage houses). We predict and find that analysts from less highly reputable brokerage houses are more likely to issue optimistic forecasts for ICMW reporting firms.

We also examine the association between ICMW and forecast accuracy (and bias) in the post-reporting periods. We find that the differences in accuracy and bias between ICMW reporting firms and control firms disappear when firms stop reporting material weaknesses.

This paper makes four major contributions to the accounting literature. First, sell-side analysts are among the most important users of financial reports. Researchers have long been interested in learning about analysts' use of accounting information (Schipper 1991). While prior studies provide evidence of the link between earnings quality and weaknesses in internal control, exactly how weakness in internal control affects the users of earnings reports directly has been largely ignored.[4] This study adds to this research by *directly* documenting a relation between ICMW and accuracy along with bias of analyst forecasts. The evidence presented in this paper shows that internal control deficiencies can influence the quality of analysts' forecasts.

Secondly, this study finds that not all ICMWs are created equal. The association between ICMW and forecast accuracy depends upon the severity of the reported ICMWs. When we separate material weaknesses into Contained or Pervasive types of weaknesses (based on the severity of the weaknesses), we find that firms identified with Pervasive ICMWs are more likely to be associated with forecast errors. In contrast, the relation between ICMW and accuracy is insignificant among firms identified with Contained ICMWs.

Thirdly, we are able to link brokerage reputation to the analysts' optimistic bias. Conceptually, it makes sense that the optimistic bias should be related to the reputation of the brokerage houses since highly reputable brokerage houses

[4]A recent working paper by Kim et al. (2009) confirms our results. They find that internal control quality is inversely; associated with analysts' error and forecast dispersion.

are more concerned with the analysts' forecasts. In addition, highly reputable houses have the resources to conduct more sophisticated analyses. This study presents the evidence to demonstrate that a good brokerage reputation can curb the optimistic bias.

Finally, the study also shows that the differences in accuracy and bias between ICMW reporting firms and control firms disappear when firms stop reporting material weaknesses. We provide evidence that ICMWs indeed contribute to lower forecast accuracy and more positive forecast bias.

The chapter is organized as follows. Section 63.2 develops hypotheses, and Sect. 63.3 describes sample selection process and forecast properties measurement. Sections 63.4 and 63.5 present the empirical tests and additional analysis, respectively. Conclusions are presented in Sect. 63.6.

63.2 Hypothesis Development

63.2.1 Internal Control Material Weakness and Forecast Accuracy

The extant literature on the relation between ICMW, accruals quality, and management's earnings guidance implies that ICMWs could affect analysts' forecast accuracy.

Doyle et al. (2007a) and Ashbaugh et al. (2008) find that ICMW reporting firms have lower accruals quality. Lobo et al. (2012) link accruals quality with analysts' forecast accuracy by documenting that firms with lower accruals quality tend to have larger forecast errors. They conclude that analysts are unable to fully resolve the uncertainty in firms with lower accruals quality.

Extant literature also suggests that ICMWs could affect management's earnings guidance. In a recent study, Feng et al. (2009) discover that ICMW reporting firms have less accurate management earnings guidance. Management's earnings guidance has been shown to be directly related to forecast accuracy, since the earnings guidance provides valuable information for analysts (Chen et al. 2011). Thus, we conjecture that ICMW disclosing firms are associated with less accurate analysts' earnings forecast. This leads to our first hypothesis in the alternative form as the following:

H1 Analysts' earnings forecast for ICMW reporting firms will be less accurate relative to non-reporting firms (in alternative form).

Internal control material weaknesses vary widely with respect to severity and underlying reasons. (See page 196 of Doyle et al. 2007b). Doyle et al. (2007b) find that the type of internal control problem is an important factor when examining determinants of ICMWs. They recommend that the type and severity of ICMWs should be considered by future research on internal control.

If a firm's management lacks the abilities or resources to exercise efficient internal control within the firm, the firm tends to have ICMWs about the overall control environment (defined as Pervasive ICMWs).[5] Alternatively, even if management has sufficient capabilities and resources to prepare accurate and adequate financial statements, a firm may still have internal control deficiencies over financial reporting. Such internal control deficiencies may be related to controls over specific account balances or transaction-level processes (defined as Contained ICMWs).[6,7] Doyle et al. (2007a) find that among all ICMW reporting firms, the earnings quality is significantly lower for firms reporting Pervasive ICMWs. In contrast, Contained ICMWs have no impact on earnings quality.

In making their forecasts, analysts can use earnings-related information, disaggregated segmental information, and information provided by management (Previts and Bricker 1994; Bouwman and Frishkoff 1995; Rogers and Grant 1997). If analysts use firms' earnings-related and segmental information to make forecasts, they might have more difficulty in predicting earnings for firms with Pervasive ICMWs. This is because earnings quality of Pervasive type ICMW firms is low (Doyle et al. 2007a). On the contrary, analysts should have no difficulty in predicting earnings for firms with Contained ICMWs, because auditors would be able to mitigate the errors in reported earnings associated with Contained ICMWs. Consistent with this argument, Doyle et al. (2007a) find that account-specific material weaknesses are not associated with lower earnings quality. If analysts use the guidance provided by management or other unaudited reports to make forecasts, it is possible that Contained ICMWs might still introduce errors into the reports since these reports are unaudited. In this case, both Pervasive and Contained ICMWs may increase the forecasting difficulties for analysts. Taken together, the forecast errors are expected to be larger for firms with Pervasive ICMWs than for firms with Contained ICMWs.

To compare the forecast accuracy between Pervasive and Contained ICMWs, we classify firms into two groups based on the company's stated reasons for material weaknesses. The first group of firms discloses *only* Contained ICMWs

---

[5]For example, DynTek Inc. disclosed the following deficiencies in their 2004, 10-K: "The material weaknesses that we have identified relate to the fact that our overall financial reporting structure and current staffing levels are not sufficient to support the complexity of our financial reporting requirements. We have experienced employee turnover in our accounting department including the position of Chief Financial Officer. As a result, we have experienced difficulty with respect to our ability to record, process and summarize all of the information that we need to close our books and records on a timely basis and deliver our reports to the Securities and Exchange Commission within the time frames required under the Commission's rules."

[6]For example, Westmoreland Coal Inc. disclosed the following deficiencies in their 2005, 10-K: "The company's policies and procedures regarding coal sales contracts with its customers did not provide for a sufficiently detailed, periodic management review of the accounting for payments received. This material weakness resulted in a material overstatement of coal revenues and an overstatement of amortization of capitalized asset retirement costs." Moody suggests that these types of material weaknesses are "auditable" and thus do not represent as serious a concern regarding the reliability of the financial statements.

[7]The detailed classification of Pervasive and Contained ICMWs is provided in Appendix 1.

(defined as G1 firms), which are related to controls over specific accounting choices. The second group of firms discloses Pervasive ICMWs (with or without Contained ICMWs) (defined as G2 firms). Note that unlike G1 firms that disclose only one type of ICMWs (i.e., Contained ICMWs). G2 firms may disclose two types of ICMWs (i.e., Pervasive as well as Contained ICMWs).[8] We have the following hypothesis based on the types of ICMWs:

H2 The analysts' forecast errors are more pronounced for forecasts issued for G2 firms (compared to those issued for G1 firms) (in alternative form).

63.2.2 Internal Control Material Weaknesses and Optimistic Forecast Bias

Extant literature suggests that sell-side analysts have an inclination to issue optimistic forecasts for several reasons. First, the compensation of analysts is tied to the amount of trade they generate for their brokerage firms. Given widespread unwillingness or inability to sell short, more trades will result from more optimistic forecasts. Moreover, mutual funds, the client group with resources to generate large trades, are precluded by regulation from selling short. Hence, without reputation concerns, analysts will prefer to issue more optimistic forecasts.

Secondly, a positive outlook improves the chances of analysts' brokerage houses winning investment banking deals. A number of prior studies have suggested that initial public offering (IPO) activities may compromise the quality of analysts' research. For example, Womack (1996) argues that analysts are reluctant to issue unfavorable forecasts if there is an IPO underwriting relationship.

Thirdly, prior studies show that analyst forecasts contain private information in addition to a statistical model based only on public information. Hence, access to management is crucial for analysts, as evidenced by the reports from Institutional Investor (a firm that compiles annual analyst rankings) showing that analysts rank the access to management as the sixth most valuable attribute out of 13 attributes (ahead of accuracy of earnings estimates, written reports, stock selection, and financial modeling). As evidenced by Huang et al. (2005), being optimistic has historically helped analysts maintain good relations with management.[9]

---

[8]The conclusions of this paper remain the same if we classify G2 firms as firms disclosing only Pervasive ICMWs.

[9]Our sample period starts after the introduction of Regulation Fair Disclosure (Reg. FD). The goal of Reg. FD is to prohibit management from selectively disclosing private information to analysts. However, as pointed out by recent research (see, e.g., Ke and Yu 2006; Kanagaretnam et al. 2012), there is no empirical evidence of management relations incentive weakening after Reg. FD. In the post-Reg. FD period, there are other incentives for analysts to please management, such as to gain favored participation in conference calls (Libby et al. 2008; Mayew 2008). Anecdotal evidence also shows that Reg. FD does not prevent a company from more subtle forms of retaliations against analysts who issue negative research reports (Solomon and Frank 2003).

The optimistic bias in analyst forecasts is more pronounced when earnings are less predictable (e.g., Lim 2001; Das et al. 1998). Feeling less accountable in uncertain environments, analysts are inclined to issue more optimistic forecasts. Consistently, Zhang (2006) concludes that greater information uncertainty predicts more positive forecast bias.

Since both less predictable earnings (Doyle et al. 2007a) and an uncertain information environment (Beneish et al. 2008) are more prevalent among ICMW firms, we expect that analysts will issue positively biased forecasts when firms have ICMWs.[10] We, therefore, offer our next hypothesis regarding the relation between forecast bias and ICMWs:

H3 Analysts' earnings forecasts are more positively biased among ICMW reporting firms relative to non-reporting firms (in alternative form).

63.2.3 The Impact of the Reputation of Brokerage Houses

As we discussed before, the magnitude of the bias is held in check by reputational concerns. We hypothesize that highly reputable brokerage houses value analysts' reports more than less highly reputable brokerage houses. Hence, their analysts are likely to feel constrained from adding an arbitrarily high optimistic bias to their private estimates by the fear of hurting the brokerage houses' reputations. In addition, highly reputable brokerage firms tend to have significant resources to collect information, and thus the access to firms' private information is relatively less important for analysts from highly reputable brokerage houses.[11]

Alternatively, analysts from less highly reputable brokerage houses tend to have limited resources in research and thus have more incentives to issue more biased forecasts. If ICMWs indeed increase the cost of information collection and research, these extra costs will exacerbate the need of analysts from less highly reputable brokerage houses to access firms' private information. Hence, we expect that analysts from less highly reputable brokerage houses will issue more upwardly biased forecasts for ICMW firms (compared to analysts from highly reputable brokerage houses). We, therefore, offer our last hypothesis:

H4 The analysts' optimistic biases associated with ICMW firms are more pronounced for forecasts issued by analysts from less highly reputable brokerage houses (compared to analysts from highly reputable brokerage houses) (in alternative form).

---

[10]Note that our hypothesis is still valid if the forecasts have been made before the disclosure of ICMWs. Doyle et al. (2007a) argue that Sarbanes-Oxley has led to the disclosure of ICMWs that might have existed for some time. Indeed, they find that accrual quality has been lower for ICMW firms relative to non-ICMW firms even in the periods prior to the disclosure of ICMWs.

[11]Brown et al. (2009) document stock market response to an analyst's recommendation change and the difference between the analyst's recommendation and the consensus recommendation. The market's reaction is strongly influenced by the analyst's reputation.

63.3 Sample Selection and Descriptive Statistics

63.3.1 The Sample Selection and Matching Procedure

We first use the key words "internal control" and "material weakness" to search the 8-K, 10-Q, and 10-K filings in Lexis/Nexis during the period of August 2002 to December 2006 and obtain 1,275 firms which disclose at least one material weakness (ICMW firms).[12] We then cross-check our 1,275 firms against Doyle et al.'s (2007a) 1,210 sample firms, which are obtained by Doyle et al. through searching 10Kwizard.com (10-Ks, 10-Qs, and 8-Ks) from August 1, 2002, to October 31, 2005.[13] We find that there are 181 firms in Doyle et al.'s (2007a) sample but not in ours. We subsequently add these 181 firms into our sample to create an initial sample of 1,456 firms. Out of the 1,456 initial sample firms, 952 firms have the required firm characteristics variables (for regressions) in the Compustat and CRSP annual database. Among these 952 firms, 745 firms have analyst forecast data in IBES database.

Next we identify a sample of matched firms that do not disclose internal control material weaknesses (non-ICMW firms) with similar IBES, CRSP, and Compustat requirements as the ICMW firms. We match ICMW firms with non-ICMW firms by industry, firm size, and sales performance, as measured during the fiscal year in which the ICMW is disclosed. Industry is defined by using the 48 industry codes identified by Fama and French (1997), firm size is measured as total assets (Compustat #6), and sales performance is measured as total sales (Compustat #12).

The matching algorithm is similar to that used by Francis et al. (2006). In particular, matches are identified by an algorithm that calculates the distance between each ICMW firm k and its matched non-ICMW counterpart j. Specifically, for each non-ICMW firm j in the same Fama-French industry as ICMW firm k, we calculate the percentage difference in assets, $\text{AssetsDIS} = \left|\frac{\text{Assets}_j - \text{Assets}_k}{\text{Assets}_k}\right|$, and the percentage difference in sales, $\text{SalesDIS} = \left|\frac{\text{Sales}_j - \text{Sales}_k}{\text{Sales}_k}\right|$. The sum of the two distance measures yields a matching score for each non-ICMW firm j that is in the same industry as ICMW firm k. From the set of matching scores that are less than two, we choose the non-ICMW firm with the smallest matching score for each ICMW firm; we then remove the matched pair (the ICMW and its non-ICMW counterpart) from the lists of ICMW and non-ICMW firms. In some cases, a single non-ICMW firm is the best match for several ICMW firms. In this case, we control for the order in which we match a non-ICMW firm to an ICMW firm by first

[12]It could be argued that our sample might miss some firms which have ICMWs but choose not to disclose them. However, discovery and disclosure of material weaknesses are mandatory according to 2004 SEC FAQ #11. We, therefore, use the disclosure of ICMW as a proxy for the existence of ICMW.

[13]We thank Sarah McVay for making the data available on her website (http://pages.stern.nyu.edu/~smcvay/research/Index.html).

calculating all possible matching scores and then assigning the non-ICMW firm j to the ICMW firm k whose matching score is the smallest among the candidate ICMWs. For the remaining candidate non-ICMWs, we repeat the above steps using the remaining ICMW firms. In total, the application of these procedures produces a final sample of 727 pairs of ICMW and non-ICMW firm-year observations.[14]

63.3.2 Descriptive Statistics

Panels A and B in Table 63.1 show the number and percentage of sample firms by industry and by stock exchange, respectively. Industry groups with the largest representations in the sample include durable manufacturers (20.4 %), computers (19.1 %), retail (12.9 %), and financial (12.5 %). Our sample distribution is similar to that of Ghosh and Lubberink (2006) and Beneish et al. (2008). As a comparison, in the last two columns of Table 63.1 we also present the number and percentage of 2003 Compustat population by industry. The industry distributions of our sample firms and 2003 Compustat firms are similar. The industry groups with the largest representations in 2003 Compustat population are also durable manufacturers, computers, and financial. The retail industry has a larger weight in our sample relative to the 2003 Compustat population. In terms of stock exchange, the majority of ICMW firms are listed on NASDAQ (436 firm-years) and NYSE (245 firm-years).

Table 63.2 Panel A presents the descriptive statistics for sample firms and matched firms. The median number of analysts following sample firms is smaller than that of analysts following matched firms. The firm size is measured as the natural logarithm of the market value of equity. The median firm size suggests that sample firms are smaller than matched firms (significant at the 0.1 significance level). There are no significant mean and median differences in leverage between sample firms and matched firms. The mean and median profitability of sample firms, measured by return on assets (ROA), are significantly lower than those of matched firms at the 0.01 significance level.

The mean and median of book to market ratios (BM) of sample firms are significantly larger than those of matched firms at the 0.01 significance level. The mean and median of percentage change in earnings of sample firms are significantly smaller than those of matched firms at the 0.01 significance level, which suggests a systematic downward shift in reported earnings for firms disclosing ICMWs. Also, the median number of negative earnings of sample firms is significantly greater than that of matched firms at the 0.01 significance level. Taken together, the statistics imply that sample firms are followed by fewer analysts and have smaller market capitalization, higher book to market ratio, and lower profitability than their matched firms.

[14]If a firm in our final sample reports internal material weaknesses in multiple years, the firm will show up in our sample multiple times. We have 599 distinct firms showing up once in our final sample and 64 firms showing up twice in our sample. The conclusions of this paper remain the same if we exclude these 64 firms from our analyses.

Table 63.1 Industry and exchange distributions of firms reporting internal control material weakness (ICMW)

Panel A: by industry

Industry name	SIC codes	Sample *N*	Sample %	2003 Compustat *N*	2003 Compustat %
Mining and construction	1000–1999 excluding 1300–1399	13	1.79	158	2.56
Food	2000–2111	3	0.41	112	1.82
Textiles and printing/publishing	2200–2799	22	3.03	210	3.41
Chemicals	2800–2824, 2840–2899	9	1.24	135	2.19
Pharmaceuticals	2830–2836	31	4.26	559	9.07
Extractive	1300–1399, 2900–2999	25	3.44	196	3.18
Durable manufactures	3000–3999, excluding 3570–3579, 3670–3679	148	20.36	945	15.34
Computers	3570–3579, 3670–3679	139	19.12	853	13.84
Transportation	4000–4899	42	5.78	333	5.40
Utilities	4900–4999	27	3.71	287	4.66
Retail	5000–5999	94	12.93	460	7.47
Financial	6000–6999	91	12.52	1377	22.35
Services	7000–8999 excluding 7370–7379	83	11.42	537	8.71
Total		727	100	6,162	98.80

Panel B: by stock exchange

Stock exchange	*N*	%
NYSE	245	33.70
NASDAQ	436	59.97
AMEX	29	3.99
OTC	16	2.20
Other	1	0.14
Total	727	100.00

A total of 727 firm-year observations have reported ICMW and have data available from Compustat and IBES. SIC codes and stock exchanges are from the Compustat
The sample period is from 2003 to 2006. All data are from CRSP, Compustat, and IBES databases

63.3.3 The Sample Firms by Types of ICMW

As we discussed in Sect. 63.2, we classify firms disclosing ICMWs into two groups: one group of firms discloses only Contained ICMWs, and the other group of firms discloses Pervasive ICMWs (with or without Contained ICMWs).

The classification of Contained or Pervasive ICMWs is similar to that of Moody's. Contained ICMWs are defined as internal control issues related to controls over specific account balances, or transaction-level processes, or accounting policy interpretations. Pervasive ICMWs are defined as internal control issues related to controls over the control environment or the overall financial reporting process.

Table 63.2 Descriptive statistics for selected variables

Panel A: descriptive statistics for the ICMW sample and the matched sample firms				
	ICMW sample ($N = 727$)		Matched sample ($N = 727$)	
	Mean	Median	Mean	Median
NUMBER	6.183	4.000**	6.614	5.000
MV	6.285	6.165*	6.456	6.343
LEV	0.196	0.157	0.185	0.138
ROA	−0.019***	0.014***	0.019	0.038
BM	0.534***	0.473***	0.488	0.424
ECHG	−0.219***	−0.116***	−0.021	0.081
LOSS	0.354	0.000***	0.228	0.000
Panel B: descriptive statistics for the Contained ICMW (G1) and Pervasive ICMW (G2) firms				
	G1 firms ($N = 348$)		G2 firms ($N = 379$)	
	Mean	Median	Mean	Median
ROA	−0.001***	0.017	−0.035	0.012
LOSS	0.322*	0.000	0.383	0.000
FOREIGN	0.391*	0.000*	0.459	0.000
ECHG	−0.261	−0.101	−0.183	−0.129

Matched firms consist of firms in the same industry based on the 48 industry codes identified by Fama and French (1997) with the closest market value and sales at the end of fiscal year

*, **, *** denote two-tailed significance levels of 10 %, 5 %, and 1 %, respectively, for the differences between the ICMW sample and the matched sample. *T*-test is used to test the difference between the mean of the ICMW sample and the matched sample, and median test is used to test the difference between the median of the ICMW sample and the matched sample

We match ICMW firms with non-ICMW firms by industry, firm size, and sales performance, as measured in the fiscal year in which the ICMW is disclosed. Industry is defined using the 48 industry codes identified by Fama and French (1997), firm size is measured as total assets (Compustat #6), and sales performance is measured as total sales (Compustat #12). The matching algorithm is similar to that used by Francis et al. (2006)

All the variables are defined in Appendix 2. N is the number of firm-year observations

*, **, *** denote two-tailed significance levels of 10 %, 5 %, and 1 %, respectively, for the differences between G1 firms and G2 firms. *T*-test is used to test the difference between the mean of G1 sample and G2 sample, and median test is used to test the difference between the median of G1 sample and G2 sample

G1 firms are firms that disclose *only* Contained ICMW, and G2 firms are firms that disclose *at least* Pervasive ICMW. The Contained and Pervasive internal control material weaknesses are similar to Moody's classification scheme. The Contained internal control material weakness is defined as the internal control issues related to controls over specific account balances, or transaction-level processes, or special accounting policy interpretation; the Pervasive internal control material weakness is defined as the internal control issues related to controls over the control environment or the overall financial reporting process

The detailed classification procedures are as follows: we first provide a breakdown of the firms based on firms' stated reasons for material weaknesses as in Ge and McVay (2005). Firms usually disclose internal control issues in nine areas: Account-Specific, Training and Personnel, Period-End Reporting/Accounting Policies, Revenue Recognition, Segregation of Duties, Account Reconciliation, Subsidiary-Specific, Senior Management, and Technology Issues. We then classify

Contained internal control issues as (1) Account-Specific, (2) Period-End Reporting/Accounting Policies, (3) Revenue Recognition, and (4) Account Reconciliation issues. The rationale for this classification is that these internal control issues are all related to controls over specific account balances, or transaction-level processes, or accounting policy interpretation. We next classify Pervasive internal controls issues as (1) Training and Personnel, (2) Segregation of Duties, (3) Subsidiary-Specific, (4) Senior Management, and (5) Technology Issues. All of these control issues are related to controls over the control environment or the overall financial reporting process.

We record 1,372 distinct deficiencies for our 727 firm-year observations since some firms disclose more than one ICMW. Among these 1,372 deficiencies, 431 are Account-Specific deficiencies; 243 are Period-End Reporting/Accounting Policies deficiencies; 138 are Revenue Recognition deficiencies; 90 are Account Reconciliation issues deficiencies; 165 are Training and Personnel deficiencies; 53 are Segregation of Duties deficiencies; 89 are Subsidiary-Specific deficiencies; 56 are Senior Management deficiencies; and 68 are Technology deficiencies. Examples of our material weakness classification scheme are presented in Appendix 1.

Among 727 ICMW firm-year observations, 348 firm-year observations disclose Contained ICMWs; 318 firm-year observations disclose both Contained and Pervasive ICMWs; and 61 firm-year observations disclose only Pervasive ICMWs. Hence, there are 348 G1 firm-year observations and 379 G2 firm-year observations (318 firm-year observations plus 61 firm-year observations).

We examine and compare G1 and G2 firms' profitability, business complexity, and changes in earnings. We use return on assets (ROA) and a loss indicator (LOSS) to proxy for profitability. ROA is calculated as earnings before extraordinary items (Compustat #18) scaled by average total assets (Compustat # 6); and LOSS is an indicator variable that equals one if earnings are negative and zero otherwise. As in Ge and McVay (2005), we use the existence of a foreign currency adjustment to proxy for the complexity of the business (FOREIGN) (Compustat Data Item #150). Lastly, we examine the percentage change in earnings.

The descriptive statistics provided in Table 63.2 Panel B suggest that G1 firms are, on average, more profitable than G2 firms. The average return on assets (ROA) of G1 firms is significantly higher than that of G2 firms (at the 0.01 significance level); the average loss (LOSS) of G1 firms is significantly lower than that of G2 firms (at the 0.1 significance level). By comparing our business complexity measure, the existence of a foreign currency adjustment (FOREIGN), we find that the business models of G2 firms, on average, are significantly more complex than those of G1 firms (at the 0.1 significance level). Lastly, we find no significant differences in means and medians of earnings changes (ECHG), which suggests that there are no systematic differences in the reporting earnings between G1 and G2 firms.

In summary, we find that G2 firms are less profitable and more complex compared to G1 firms. These results imply that the managements of G2 firms may have limited resources to invest in proper internal control (due to lower profitability) and have more difficulty in establishing efficient internal control (due to higher complexity) than G1 firms.

63.4 Empirical Tests

63.4.1 Variable Measurement

Based on Kanagaretnam et al. (2012), we define forecasts accuracy and bias as follows:

Forecast accuracy (ACCURACY) is calculated as $-\frac{|\text{EPS Forecasted}-\text{EPS Actual}\,|}{\text{Beginning Stock Price}}$; both forecasted and actual earnings per share are from IBES Summary Files. Because we are interested in assessing the impact of a firm's internal control on its financial statements, we try to focus on a particular announcement date: the annual earnings announcement. Forecast accuracy is computed as the absolute difference between the last median forecasted earnings before the annual earnings announcement and the actual earnings for the year in which ICMWs are disclosed.[15] We deflate forecast accuracy by beginning stock price to facilitate comparisons across firms.

Forecast bias (BIAS) is calculated as $\frac{\text{EPS Forecasted}-\text{EPS Actual}}{\text{Beginning Stock Price}}$, both forecasted and actual earnings per share are from IBES Summary Files. Forecast bias is computed as the difference between the last median forecasted earnings before the annual earnings announcement and the actual earnings for the year in which ICMWs are disclosed. We also deflate forecast bias by beginning stock price to facilitate comparisons across firms.[16]

63.4.2 Univariate Analysis

We first examine whether there are significant differences between ICMW firms and their matched firms in forecast accuracy and bias using mean and median tests. The results are reported in Table 63.3.

We find that median forecast accuracy for the ICMW sample is significantly smaller than that for matched firms at the 0.01 level, suggesting that internal control material weaknesses are related to less accurate forecasts. In addition, we find that median forecast bias for the ICMW sample is significantly larger than that for matched firms at the 0.01 level, consistent with the notion that internal control material weaknesses are related to more optimistic forecasts.

A similar pattern exists when comparing forecast accuracy and bias separately for G1 firms and their matched firms and for G2 firms and their matched firms.

[15]Note that the ACCURACY is defined so that larger errors correspond to a lower level of accuracy.

[16]As a sensitivity test, we calculate ACCURACY and BIAS using the simple average of the measures across the 6 or 12 monthly reporting periods on the IBES before the company's fiscal year end. In other words, we choose all median forecasts across the 6 or 12 monthly reporting periods on IBES before the company's fiscal year ends and then average the median forecasts to create our ACCURACY and BIAS variables. The results are similar to what we report in this paper (not tabulated). When we use forecasts from the prior year instead of the current year, we also get similar results (not tabulated).

Table 63.3 Accuracy and bias statistics for all ICMW firms, Contained ICMW (G1) and Pervasive ICMW firms (G2), and their matched firms

	ICMW firms ($N = 727$)		Matched firms ($N = 727$)	
	Mean	Median	Mean	Median
ACCURACY	−0.019	−0.003***	−0.015	−0.002
BIAS	0.007*	0.001***	−0.001	0.000
	G1 firms ($N = 348$)		Matched firms ($N = 348$)	
	Mean	Median	Mean	Median
ACCURACY	−0.017	−0.003***	−0.022	−0.002
BIAS	0.004	0.001***	0.000	0.000
	G2 firms ($N = 379$)		Matched firms ($N = 379$)	
	Mean	Median	Mean	Median
ACCURACY	−0.020***	−0.004***	−0.009	−0.002
BIAS	0.010***	0.001***	0.000	0.000

ACCURACY is forecast accuracy, calculated as the negative of the absolute difference between actual EPS and last median forecasted EPS scaled by stock price. BIAS is forecast bias, calculated as the difference between last median forecasted EPS and actual EPS scaled by stock price. *, **, *** denote two-tailed significance levels of 10 %, 5 %, and 1 %, respectively, for the differences between ICMW firms and matched firms, G1 firms and matched firms, and G2 firms and matched firms. *T*-test is used to test the difference between the mean of the ICMW sample and the matched sample, and median test is used to test the difference between the median of the ICMW sample and the matched sample

Matched firms consist of firms in the same industry based on the 48 industry codes identified by Fama and French (1997) with the closest market value and sales at the end of fiscal year

G1 firms are firms that disclose *only* Contained ICMW, and G2 firms are firms that disclose *at least* Pervasive ICMW. The Contained and Pervasive internal control material weaknesses are similar to Moody's classification scheme. The Contained internal control material weakness is defined as the internal control issues related to controls over specific account balances, or transaction-level processes, or special accounting policy interpretation; the Pervasive internal control material weakness is defined as the internal control issues related to controls over the control environment or the overall financial reporting process

The median forecast accuracy of G1 firms is smaller than that of the matched firms at the 0.01 significance level. For G2 firms, the median forecast accuracy is also smaller than that of matched firms at the 0.01 significance level. In terms of optimistic bias, we find that median forecast biases for G1 and G2 samples are larger than those for their matched firms at the 0.01 significance level.

These findings suggest that internal control material weaknesses are associated with less accurate and more optimistic forecasts for both G1 and G2 firms. In the next two subsections, we examine the association between ICMW and forecast accuracy (and bias) using multivariate regressions.

The definitions of all the independent variables are provided in Appendix 2. Since we have a large number of independent variables incorporated in our regression models, the statistical inference on the variables could be affected by multicollinearity. Multicollinearity is a high degree of correlation (linear dependency) among several independent variables. It commonly occurs when some of the independent variables measure the same concepts or phenomena.

In Table 63.4, we present the pair-wise correlations among all independent variables in the regressions and find that none of the correlations are larger than 0.50 (or smaller than −0.50) except for the correlation between SKEW and EPSVOL. To avoid the multicollinearity issue, we choose not to include these two variables in the same regression model.

63.4.3 Analysis of Forecast Accuracy

The hypotheses to be tested are that analyst accuracy and bias are a function of ICMWs. However, the results will be difficult to interpret if endogeneity is a concern. We include firm-specific fixed effects to control for the possibility that endogeneity arises from omitted unobserved factors (e.g., business models) that may be correlated with both forecast quality and ICMWs.

The model employed to test the association between forecast accuracy and ICMW is (H1 and H2):

$$\begin{aligned}\mathrm{ACCURACY}_{\mathrm{i,t}} = \alpha_0 + \alpha_1 \mathrm{ICMW}_{i,t} + \alpha_2 \mathrm{NUM}_{i,t} + \alpha_3 \mathrm{MV}_{i,t} + \alpha_4 \mathrm{LEV}_{i,t} \\ + \alpha_5 \mathrm{ROA}_{i,t} + \alpha_6 \mathrm{BM}_{i,t} + \alpha_7 \mathrm{EPSVOL}_{i,(t-5,t-1)} + \alpha_8 \mathrm{ABSECHG}_{i,t} \\ + \alpha_9 \mathrm{LOSS}_{i,t} + \alpha_{10} \mathrm{SPECIAL}_{i,t} + \alpha_{11} \mathrm{RET}_{i,(t-3,t-1)} + \alpha_{12} \mathrm{DA}_{i,t} + \varepsilon\end{aligned} \tag{63.1}$$

where ACCURACY is forecast accuracy, calculated as the negative of the absolute difference between actual EPS and last median forecasted EPS scaled by stock price. ICMW is an indicator variable that equals one if a firm discloses a material weakness in internal control and zero otherwise. The definitions of the other variables are provided in Appendix 2.

In our regression model we first control for earnings characteristics. Prior research identifies earnings volatility (EPSVOL), losses (LOSS), and special items (SPECIAL) as earnings characteristics that can negatively affect forecast accuracy. The forecasting task is more difficult for firms with historically more volatile earnings compared to firms with historically more stable earnings (e.g., Kross et al. 1990; Lim 2001), losses, and special items (Brown and Higgins 2001). In addition, we include absolute earnings changes (ABSECHG) to capture any shift in reported earnings. Prior studies show that forecast errors are larger for larger earnings surprises (e.g., Lang and Lundholm 1996; Duru and Reeb 2002). Moreover, we use absolute abnormal accruals (DA) to control for earnings quality. DA is estimated using the modified Jones model of Larcker et al. (2007). We expect a negative relation between forecast accuracy and DA.[17]

[17]Note that the differences in the earnings quality could also be the consequences of the existence of ICMWs. By controlling for absolute abnormal accruals, we may overcontrol the impact of ICMWs. But it will bias against us finding any results.

Table 63.4 Pearson correlations among the variables used in regression analysis

	NUM	MV	LEV	BM	SPE	NECHG	EPSVOL	SKEW	ICMW	ABSECHG	ECHG	LOSS	ROA	RET
NUM	1.00	0.47^{***}	-0.21^{***}	-0.18^{***}	−0.10	-0.17^{***}	0.31^{***}	−0.07	0.03	−0.03	−0.01	-0.11^{**}	0.23^{***}	-0.21^{***}
MV		1.00	0.08	−0.24	-0.07^{***}	-0.28^{***}	0.29^{***}	−0.09	−0.03	−0.09	0.07	-0.18^{***}	0.29^{***}	−0.07
LEV			1.00	−0.05	0.03	0.08	0.09	0.12^{**}	0.03	0.01	−0.08	0.12^{**}	-0.13^{**}	0.01
BM				1.00	0.06	0.19^{***}	−0.01	0.04	0.13^{**}	−0.01	−0.04	0.07	−0.05	0.03
SPE					1.00	−0.06	0.05	−0.02	0.01	−0.05	0.06	−0.03	0.03	0.09
NECHG						1.00	0.05	0.05	0.19^{***}	0.10	-0.35^{**}	0.33^{***}	-0.33^{***}	−0.01
EPSVOL							1.00	-0.67^{***}	0.04	−0.02	0.01	0.11^{**}	-0.12^{**}	0.02
SKEW								1.00	0.05	−0.02	0.01	0.09	−0.05	0.03
ICMW									1.00	−0.01	−0.06	0.11^{**}	−0.09	−0.08
ABSECHG										1.00	-0.49^{***}	0.19^{***}	-0.19^{***}	−0.07
ECHG											1.00	-0.18^{***}	0.17^{***}	0.02
LOSS												1.00	-0.45^{***}	−0.04
ROA													1.00	0.01
RET														1.00

All the variables are described in Appendix 2
**, *** significant at 0.05 level and 0.01 level, respectively

We next control for other firm characteristics such as size, growth, financial leverage, profitability, and risk. Firm size is measured as the natural logarithm of market value at the end of the year (MV) and has been used in the literature as a proxy for a number of factors. To the extent that size reflects information availability about a firm (other than through annual reports), a positive relation with forecast accuracy is expected (Ho 2004). However, firm size could also proxy for a host of other factors, such as managers' incentives, for which predictions for the relation with forecast accuracy are unclear.[18]

We measure growth as the natural logarithm of the ratio of the book value of equity to the market value of equity at the end of the year (BM). Dechow and Sloan (1997) and Richardson et al. (2004) find that forecast accuracy and bias are related to measures of growth. Consistent with prior research, we expect firms with low book to market ratios (i.e., high growth firms) to have more accurate forecasts than firms with high book to market ratios (i.e., turnaround and declining firms). We also include the debt to equity ratio (LEV) to proxy for financial leverage and return on assets (ROA) to proxy for profitability. We do not have predictions for the sign of these two variables. Finally, we use the equally weighted market-adjusted cumulative return over the past 3 years (RET) to proxy for firm risk. We expect a negative relation between forecast accuracy and RET.[19]

Next, we use the natural logarithm of the number of analysts who issue the forecasts in calculating the last median earnings (NUM) to account for the effects of differences in forecast characteristics on forecast accuracy. Lys and Soo (1995) argue that the number of analysts proxies for the intensity of competition in the market. We expect a positive relation between forecast accuracy and analyst following. Finally, we include firm-specific and exchange-specific fixed effects to control for firm-specific and exchange-specific shocks.

The results of the accuracy tests are reported in Table 63.5. Column 3 shows the OLS regression results for the overall sample along with the matched sample; Column 4 shows the OLS regression results for G1 sample firms and their matched firms; and Column 5 shows the OLS regression results for G2 sample firms and their matched firms.

The results presented in Table 63.5 show that the coefficients on ICMW for the overall sample and G1 sample firms are not significantly different from zero. These findings suggest that forecast accuracy is not significantly different for all ICMW disclosing firms and for Contained ICMW disclosing firms relative to their corresponding matched firms when controlling for other independent variables. In contrast, for G2 sample firms, the coefficient of ICMW in Column 4 is significantly negative (at the 0.1 level for the two-tailed test). The negative coefficient suggests that forecast accuracy is significantly lower for Pervasive ICMW disclosing firms compared with their matched firms.

Inferences about the control variables in the regression are generally similar to previous studies. Specifically, firms with lower frequency of negative earnings,

[18]Fan and Yeh (2006) find that forecasting error is a negative function of firm size.

[19]The results are similar if we use value-weighted market-adjusted cumulative returns.

Table 63.5 OLS regression estimations relating ACCURACY to ICMW firm variables
$\mathrm{ACCURACY}_{i,t} = \alpha_0 + \alpha_1 \mathrm{ICMW}_{i,t} + \alpha_2 \mathrm{NUM}_{i,t} + \alpha_3 \mathrm{MV}_{i,t} + \alpha_4 \mathrm{LEV}_{i,t} + \alpha_5 \mathrm{ROA}_{i,t} + \alpha_6 \mathrm{BM}_{i,t} + \alpha_7 \mathrm{EPSVOL}_{i,(t-5,t-1)} + \alpha_8 \mathrm{ABSECHG}_{i,t} + \alpha_9 \mathrm{LOSS}_{i,t} + \alpha_{10} \mathrm{SPECIAL}_{i,t} + \alpha_{11} \mathrm{RET}_{i,(t-3,t-1)} + \alpha_{12} \mathrm{DA}_{i,t} + \varepsilon$

		Dependent variable = ACCURACY		
	Predicted *Sign*	All sample $N = 1454$	G1 sample $N = 696$	G2 sample $N = 758$
(1)	(2)	(3)	(4)	(5)
Independent variables		Coefficient	Coefficient	Coefficient
		(t-statistic)	(t-statistic)	(t-statistic)
INTERCEPT		−0.043	−0.038	−0.046
		(−3.65***)	(−1.92*)	(−3.19***)
ICMW	−	−0.005	0.002	−0.009
		(−1.22)	(0.36)	(−1.77*)
NUM	+	0.003	−0.002	0.007
		(0.92)	(−0.36)	(1.69*)
MV	+	0.003	0.004	0.002
		(1.41)	(1.14)	(0.73)
LEV	+/−	−0.005	−0.016	−0.008
		(−0.44)	(−0.80)	(−0.52)
BM	−	−0.009	−0.006	−0.011
		(−3.02***)	(−0.99)	(−3.12***)
SPECIAL	−	−0.016	−0.024	−0.009
		(−2.57***)	(−2.22**)	(−1.14)
EPSVOL	−	−0.000	0.002	−0.000
		(−1.13)	(0.92)	(−1.10)
ABSECHG	−	−0.000	−0.000	0.000
		(−0.05)	(−0.20)	(0.47)
LOSS	−	−0.010	−0.006	−0.006
		(−1.75*)	(−0.54)	(−0.76)
ROA	+/−	0.120	0.181	0.120
		(5.69***)	(3.96***)	(5.21***)
RET	−	0.002	0.017	−0.005
		(0.79)	(2.93***)	(−1.35)
DA	−	−0.044	−0.099	−0.026
		(−1.49)	(−1.67*)	(−0.84)
R-square		16.11 %	14.21 %	23.32 %

Observations include ICMW firms and matched firms that do not disclose ICMW. Matched firms consist of firms in the same industry based on the 48 industry codes identified by Fama and French (1997) with the closest market value and sales at the end of fiscal year. Regressions in the third, fourth, and fifth columns include all ICMW firms and matched firms, G1 firms and matched firms, and G2 firms and matched firms separately. Regressions control for exchange and firm fixed effects. Outliers are excluded using Cook's (1977) distance statistic. N is the number of firm-year observations. *, **, *** denote two-tailed significance levels of 10 %, 5 %, and 1 %, respectively. All the variables are defined in Appendix 2 lower book to market ratio, smaller abnormal accruals, and a larger number of analyst following have more accurate forecasts (as evidenced by significant P-value [<0.1] in Table 63.5). The results documented in this subsection are consistent with our H2 hypothesis.

One alternative explanation for our findings is that auditors may become more conservative in the year of the ICMW, leading to a systematic downward shift in reported earnings, as evidenced by the results in Table 63.2 Panel A.[20] The systematic downward shift may account for the differences in ACCURACY between ICMW firms and non-ICMW firms. This alternative explanation is unlikely since we have included ABSECHG, absolute changes in reported earnings, in the regression model to capture any systematic shift in reported earnings.

63.4.4 Analysis of Forecast Bias

We estimate the following regression model to test H3 and H4 on forecast bias:

$$\begin{aligned}\mathrm{BIAS}_{\mathrm{i,t}} = {} & \beta_0 + \beta_1\mathrm{ICMW}_{i,t} + \beta_2\mathrm{NUM}_{i,t} + \beta_3\mathrm{MV}_{i,t} + \beta_4\mathrm{LEV}_{i,t} + \beta_5\mathrm{BM}_{i,t} \\ & + \beta_6\mathrm{SKEW}_{i,(t-5,t-1)} + \beta_7\mathrm{ECHG}_{i,t} + \beta_8\mathrm{LOSS}_{i,t} + \beta_9\mathrm{SPECIAL}_{i,t} \\ & + \beta_{10}\mathrm{NECHG}_{i,t} + \beta_{11}\mathrm{RET}_{i,(t-3,t-1)} + \beta_{12}\mathrm{DA}_{i,t} + \varepsilon \end{aligned} \tag{63.2}$$

where BIAS is forecast bias, calculated as the difference between last median forecasted EPS and actual EPS scaled by stock price.[21] The definitions of Eq. 63.2's other variables are provided in Appendix 2.

As in the forecast accuracy tests, we use losses (LOSS) and special items (SPECIAL) to control for earnings characteristics.[22] Unlike the accuracy test, we use earnings changes (ECHG) instead of absolute earnings changes (ABSECHG). We expect that optimistic forecast bias is positively associated with earnings characteristics because these characteristics are positively related to the complexity of forecasting tasks. In addition, we include ECHG and NECHG to control for the anchoring behavior of analysts who tend to anchor their forecasts closely to previous period's actual results.[23] We use the difference between the mean and median of price-scaled earnings from five prior years (minimum of 4 years) to proxy for earnings skewness (SKEW). We expect that optimistic forecast bias is

[20]We thank an anonymous referee for pointing this out to us.

[21]When we use forecasts from the prior year instead of the current year, we also get similar results (not tabulated).

[22]We exclude earnings volatility (EPSVOL) variable because the variable is significantly correlated with SKEW variable. In a sensitivity test, we replace SKEW by EPSVOL; the results are similar to what we report in the paper.

[23]See Shiller (1999) on anchoring behavior of financial analysts.

negatively associated with earnings skewness primarily due to mean-median differences in skewed earnings distributions (Gu and Wu 2003).

As in the forecast accuracy test, we use size (MV), growth (BM), leverage (LEV), profitability (ROA), risk (RET), and earnings quality (DA) to control for differences in firm characteristics. We have no prediction on the signs of these variables. We also include the log of the number of analysts following the firm (NUM) to account for differences in forecasts characteristics that affect forecast bias. In addition, we include firm-specific and exchange-specific fixed effects to control for firm-specific and exchange-specific shocks.

In Table 63.6 we report the results for the bias test. The coefficient on ICMW is significantly positive for the overall sample at the 0.05 significance level, which suggests that analysts tend to be more optimistic toward ICMW disclosing firms. Inferences about the control variables suggest that firms with smaller size, higher leverage, higher frequency of negative earnings, and lower profitability have more optimistically biased forecasts (as evidenced by significant P-value [<0.1] in Table 63.6). The results presented in Table 63.6 are consistent with our H3 hypothesis. Analysts' earnings forecasts are more positively biased among ICMW reporting firms relative to non-reporting firms.

63.4.5 Analysis of Brokerage Reputation

In this subsection, we examine the impact of the brokerage reputation on the positive bias. We use Institutional Investors' ranking of brokerage houses to proxy for brokerage reputation: a brokerage is considered to be of high reputation if it is consistently (for at least three prior years) ranked among the "Leaders" by Institutional Investors during the sample period. The following 12 brokers are identified as highly reputable: Bear Stearns & Co; CS First Boston; Goldman Sachs; Lehman Brothers; Donaldson, Lufkin & Jenrette; J.P. Morgan Securities; Merrill Lynch; Morgan Stanley & Co; Paine Webber; Prudential Securities; Salomon Brothers; and Smith Barney. The other brokerages are classified as less highly reputable brokerages.

An analyst's brokerage affiliation data comes from the IBES Detailed Files. We first separate all forecasts issued for ICMW firms into two subsamples based on whether the forecasts are issued by the analysts from highly reputable brokerage houses or not. All forecasts issued by the analysts affiliated with highly reputable brokerage houses belong to the highly reputable brokerage subsample, and the remaining forecasts belong to the less highly reputable brokerage subsample.

For each analyst, we only keep the last forecast issued for each forecast period. We then calculate the median forecasted earnings for each forecast period for each sample firm in each subsample. Note that, if a firm has forecasts made by the analysts affiliated with both highly reputable and less highly reputable brokerage houses, the firm will show up in both subsamples with different values for the medians. The median forecasted earnings in the highly reputable brokerage (less highly reputable brokerage) subsample are calculated based on the forecasts made

Table 63.6 OLS regression estimations relating bias to ICMW firm variables $\mathrm{BIAS}_{i,t} = \beta_0 + \beta_1 \mathrm{ICMW}_{i,t} + \beta_2 \mathrm{NUM}_{i,t} + \beta_3 \mathrm{MV}_{i,t} + \beta_4 \mathrm{LEV}_{i,t} + \beta_5 \mathrm{BM}_{i,t} + \beta_6 \mathrm{SKEW}_{i,(t-5,t-1)} + \beta_7 \mathrm{ECHG}_{i,t} + \beta_8 \mathrm{LOSS}_{i,t} + \beta_9 \mathrm{SPECIAL}_{i,t} + \beta_{10} \mathrm{NECHG}_{i,t} + \beta_{11} \mathrm{RET}_{i,(t-3,t-1)} + \beta_{12} \mathrm{DA}_{i,t} + \varepsilon$

	Dependent variable = BIAS	
	Predicted Sign	All sample $N = 1454$
(1)	(2)	(3)
Independent variables		Coefficient
		(t-statistic)
INTERCEPT		0.015
		(2.04**)
ICMW	+	0.005
		(2.03**)
NUM	+/−	−0.001
		(−0.33)
MV	+/−	−0.003
		(−1.97**)
LEV	+/−	0.017
		(2.34[b])
BM	+/−	0.001
		(0.77)
SPECIAL	+	0.004
		(1.11)
NECHG	−	0.001
		(0.22)
SKEW	−	−0.000
		(−1.53)
ECHG	+	−0.000
		(−0.03)
LOSS	+	0.013
		(3.34***)
ROA	+/−	−0.059
		(−4.51***)
RET	+/−	−0.001
		(−0.77)
DA	+/−	0.025
		(1.39)
R-square		17.03 %

Observations include ICMW firms and matched firms that do not disclose ICMW. Matched firms consist of firms in the same industry based on the 48 industry codes identified by Fama and French (1997) with the closest market value and sales at the end of fiscal year. Regressions control for exchange fixed effects. Outliers are excluded using Cook's (1977) distance statistic. N is the number of firm-year observations. All the variables are defined in Appendix 2

*, **, *** denote two-tailed significance levels of 10 %, 5 %, and 1 %, respectively

by the analysts affiliated with highly reputable brokerage houses (less highly reputable brokerage houses).

We next compute forecast bias (BIAS) as the difference between the last median forecasted earnings before the annual earnings announcement and the actual earnings for the year in which ICMWs are disclosed. As in the prior section, we also deflate forecast bias by beginning stock price to facilitate comparisons across firms. We reestimate Eq. 63.2 for highly reputable brokerage forecasts of the sample firms and their matched firms (638 observations) and for less highly reputable brokerage forecasts of the sample firms and their matched firms (816 observations), respectively.

The results are presented in Columns 3–4 of Table 63.7. As reported in Column 3, the analysts affiliated with highly reputable brokerage houses are less likely to issue biased forecasts for firms with ICMWs (compared to the analysts from less highly reputable brokerage houses). The significantly positive relation between ICMWs and forecast bias only exists when those forecasts are made by the analysts affiliated with less highly reputable brokerage houses as reported in Column 4 (P-value < 0.1).

There are some sample firms that have forecasts made by analysts affiliated with both highly reputable and less highly reputable brokerage houses. We repeat our tests for this subsample. We reestimate Eq. 63.2 for highly reputable brokerage forecasts of the subsample firms and their matched firms (276 observations) and for less highly reputable brokerage forecasts of the subsample firms and their matched firms (276 observations), respectively.

The results are presented in Table 63.7 Columns 5–6. Similar to what are reported for the whole sample, the significantly positive relation between ICMWs and forecast bias only exists when those forecasts are made by the analysts affiliated with less highly reputable brokerage houses as reported in Column 5 (P-value < 0.1).

Taken together, our regression results show that analysts from less highly reputable brokerage houses are likely to issue more optimistic forecasts. We interpret these findings as that highly reputable brokerage houses value the creditability of analysts' reports more than less highly reputable brokerage houses. Hence, analysts are likely to feel constrained from adding an arbitrarily high optimistic bias to their estimates by a fear of hurting the brokerage houses' reputations.

63.5 Additional Tests

63.5.1 ICMW Resolution and Forecast Accuracy and Bias

Our first additional test probes our findings of the association between ICMWs and forecast accuracy (and bias) for G2 firms. In particular, we investigate whether the association between ICMW and forecast accuracy (and bias) still exists when firms resolve the internal control deficiencies. If there is indeed an association between ICMW and forecast accuracy (and bias), we should find no difference in forecast accuracy (and bias) between G2 firms and their matched firms after ICMW firms resolve their internal control issues. To test this hypothesis, we examine whether the

Table 63.7 OLS regression estimations relating bias to ICMW firm variables based on brokerage reputation partition

	Predicted Signs	Highly reputable $N = 638$	Less highly reputable $N = 816$	Highly reputable reputable $N = 276$	Less highly reputable $N = 276$
(1)	(2)	(3)	(4)	(5)	(6)
Independent variables		Coefficient (t-statistic)	Coefficient (t-statistic)	Coefficient (t-statistic)	Coefficient (t-statistic)
INTERCEPT		0.302	0.058	−0.008	0.013
		(6.99***)	(2.98***)	(−0.63)	(0.18)
ICMW	+	−0.002	0.007	0.003	0.004
		(−0.29)	(1.97**)	(0.85)	(1.84*)
NUM	+/−	−0.007	0.003	−0.002	−0.001
		(−1.16)	(0.98)	(−0.73)	(0.88)
MV	+/−	−0.003	−0.008	−0.001	0.000
		(−0.66)	(−3.63***)	(−0.30)	(−0.43)
LEV	+/−	−0.025	0.020	−0.014	0.001
		(−1.18)	(1.75*)	(1.19)	(1.06)
BM	+/−	0.005	0.001	0.000	−0.001
		(0.78)	(0.44)	(0.10)	(−0.41)
SPECIAL	+	0.037	−0.000	0.010	−0.003
		(3.30***)	(−0.04)	(1.62)	(−1.34)
NECHG	−	0.003	−0.004	0.004	0.000
		(0.37)	(−1.09)	(0.98)	(0.13)
SKEW	−	−0.000	−0.000	−0.000	−0.000
		(0.06)	(1.71*)	(0.04)	(0.11)
ECHG	+	−0.000	0.000	−0.000	−0.000
		(−0.37)	(0.53)	(−0.10)	(−1.47)
LOSS	+	0.005	0.017	0.010	0.008
		(0.40)	(3.09***)	(1.68*)	(3.22***)
ROA	+/−	−0.128	−0.050	0.030	0.003
		(−2.49***)	(−2.81***)	(1.02)	(0.23)
RET	+/−	−0.008	−0.004	0.001	0.002
		(−1.56)	(−1.46)	(0.67)	(1.51)
DA	+/−	0.148	0.012	0.038	0.013
		(1.85*)	(0.51)	(0.85)	(0.70)
R-square		35.24 %	23.75 %	12.16 %	10.57 %

Observations include ICMW firms and matched firms that do not disclose ICMW. Regressions in the third and fourth columns include estimates made by analysts affiliated with highly reputable brokers and the estimates made by analysts affiliated with less highly reputable brokers separately The regressions results in the fifth and sixth columns are for ICMW firms that are covered by both analysts affiliated with highly reputable and less highly reputable brokers. The regressions results on the estimates made by analysts affiliated with highly reputable brokers are reported in the fifth column, and the regressions results on the estimates made by analysts affiliated with less highly reputable brokers are reported in the sixth column. Regressions control for exchange and firm fixed effects. Outliers are excluded using Cook's (1977) distance statistic. *, **, *** denote two-tailed significance levels of 10 %, 5 %, and 1 %, respectively. All the variables are defined in Appendix 2. Dependent variable is BIAS forecast accuracy (and bias) for G2 sample firms is significantly different from their matched firms after the sample firms stop disclosing internal control material weaknesses.

We start by assigning a post-ICMW year to each G2 firm-year observation. The post-ICMW year is defined as the first year after the year in which ICMWs are disclosed in our sample period (i.e. the first year in which our sample firms solve their disclosed internal control issues).[24] We then reestimate Eqs. 63.1 and 63.2 for the post-ICMW firm-year observations. We have 196 post-ICMW firm-year observations for G2 sample, and 183 firm-year observations are lost due to the lack of required data. The un-tabulated results show that the ICMW coefficients in both accuracy and bias tests are no longer significant, which suggests that in the first year in which our sample firms solve their disclosed internal control issues, there are no significant differences in accuracy and bias between G2 firms and their matched firms.

63.5.2 The Differences in the Frequency of Analyst Forecasts Between Highly Reputable Houses and Less Highly Reputable Houses

For our sample period, analysts affiliated with highly reputable houses on average issue 3.49 annual earnings forecasts for each firm each year (the median number of forecasts issued is 3), while analysts affiliated with less highly reputable houses on average issue 3.07 annual earnings forecasts for each firm each year (the median number of forecasts issued is 2). Thus, it appears that analysts affiliated with highly reputable houses issue forecasts more frequently than those affiliated with less highly reputable houses. To further examine this issue, we create one variable STALE, which is calculated as the difference between the date of fiscal year end and the date of the last annual forecast issued by an analyst. In un-tabulated test, we find that the mean and median STALE values for analysts affiliated with highly reputable houses are 99.83 and 68 days, respectively. In contrast, the mean and median STALE values for analysts affiliated with the less highly reputable houses are 110.92 and 75 days. Both mean and median differences for the STALE variable between analysts affiliated with highly reputable houses and those affiliated with less highly reputable houses are statistically significant at the 0.01 levels.

To make sure that our results in Table 63.7 are not caused by the differences in the frequency of analyst forecasts made by highly reputable houses and less highly reputable houses, we include an ADJSTALE variable in Eq. 63.2. The ADJSTALE variable is the median value of the STALE variable for each forecast period for each sample firm. We then reestimate the revised Eq. 63.2 for highly reputable brokerage forecasts of the sample firms and their matched firms (638 observations) and for less highly reputable brokerage forecasts of the sample firms and their

[24] For firms that disclose ICMWs for multiple years, the last year in which ICMWs are disclosed for the firms in our sample period is used.

matched firms (816 observations), respectively. Similar to what are reported in Table 63.7, we find that the significantly positive relation between ICMWs and forecast bias only exists when those forecasts are made by the analysts affiliated with less highly reputable brokerage houses (results are not tabulated). Hence, our results in Table 63.7 are unlikely to be caused by the differences in the frequency of analyst forecasts made by highly reputable houses and less highly reputable houses.

63.5.3 Other Robustness Tests

We conduct several additional tests to probe the robustness of our main results reported in Tables 63.4 and 63.5. Eames and Glover (2003) document that there is no significant relation between forecast error and earnings predictability after controlling for the level of earnings. We, first, investigate whether the level of earnings will affect our results. Our main results remain qualitatively the same after controlling for the level of earnings (not tabulated).

Second, we reestimate the models using rank regression techniques to assess the sensitivity of the results to the underlying functional form assumption made by OLS (Cavanagh and Sherman 1998). The results using rank regressions support the reported results (not tabulated). Lastly, we classify our sample firms into 302 and 404 ICMW firms to capture any differences in ICMW disclosures between Sections 302 and 404. ICMW firms under Section 302 are those that have a market value of less than $75 million or have disclosed their material weakness prior to November 15, 2004. ICMW firms under Section 404 are those that have a market value of at least $75 million and disclose a material weakness on or after November 15, 2004. We reestimate our Eqs. 63.1 and 63.2 separately for 302 and 404 disclosures. The results for 302 and 404 firms are qualitatively similar (not tabulated).

Last, on April 28, 2003, an enforcement agreement, the Global Settlement, was reached between the SEC, NASD, NYSE, and ten of the largest investment firms in the USA to address issues of conflict of interest within their businesses. One of the goals of Global Settlement is to reduce the biases in analyst reports. To make sure that our results are not influenced by the Global Settlement, we include a set of dummy variables in all our regression models to control for year-specific shocks. Our un-tabulated results are similar to what are reported in the paper.

63.6 Conclusion

In this paper we investigate the effects of ICMWs on the accuracy and optimistic bias of financial analysts' earnings forecasts. We find that financial analysts' earnings forecasts are less accurate and more positively biased among ICMW reporting firms relative to non-reporting firms.

When separating all the ICMW reporting firms into Pervasive ICMW and Contained ICMW reporting firms, we find that accuracy is only inversely associated with Pervasive ICMWs. One possible explanation is that the managers of Pervasive ICMW reporting firms cannot effectively control firms' financial reporting process, which leads to much noisier financial statements and thus makes financial analysts' earnings forecasts less accurate.

In addition, we find that analysts from less highly reputable brokerage houses are likely to issue more optimistic forecasts. We interpret these findings as that highly reputable brokerage houses value the creditability of analysts' reports more than less highly reputable brokerage houses. Hence, analysts are likely to feel constrained from adding an arbitrarily high optimistic bias to their estimates by a fear of hurting the brokerage houses' reputations.

Our results suggest that the weaknesses in internal control increase the complexity of the forecasting tasks for financial analysts. Analysts' forecasts become less accurate if the weaknesses belong to the Pervasive type. The weaknesses in internal control are also associated with more positively biased forecasts. However, the good reputation of brokerage firms appears to curb the upward bias. We also show that when Pervasive ICMW reporting firms stop disclosing weaknesses, the documented relation between accuracy (bias) and ICMWs disappears.

Appendix 1: Material Weakness Classification Examples

G1 Contained Material Weaknesses

1. Account-Specific
   e.g., "U.S. Cellular did not maintain effective controls over the completeness, accuracy, presentation and disclosure of its accounting for income taxes." (U.S. Cellular Inc., 2004 10-K report)
2. Period-End Reporting/Accounting Policies
   e.g., "–a lack of an ongoing formal self-assessment process related to internal control over financial reporting." (Ivanhoe energy Inc., 2004 10-K report)
3. Revenue Recognition
   e.g., "The Company's policies and procedures regarding coal sales contracts with its customers did not provide for a sufficiently detailed, periodic management review of the accounting for payments received. This material weakness resulted in a material overstatement of coal revenues and an overstatement of amortization of capitalized asset retirement costs." (Westmoreland Coal Co., 2005 10-K report)
4. Account Reconciliation
   e.g., "The Company did not maintain effective controls over reconciliations of certain financial statement accounts." (SIRVA Inc., 2004 10-K report)

G2 Pervasive material weaknesses

1. Segregation of Duties
   e.g., "Inadequate segregation of duties was noted with respect to the revenue, expenditure and payroll processes as numerous incompatible tasks are performed by the same accounting personnel." (Versant Co., 2004 10-K report)

2. Subsidiary-Specific
   e.g., "We have reported to the SEC that one of our foreign subsidiaries operating in Nigeria made improper payments of approximately $2.4 million to an entity owned by a Nigerian national who held himself out as a tax consultant when in fact he was an employee of a local tax authority. The payments were made to obtain favorable tax treatment and clearly violated our Code of Business Conduct and our internal control procedures." (Halliburton Co., 2002 10-K report)
3. Senior Management
   e.g., "management did not set a culture that extended the necessary rigor and commitment to internal control over financial reporting." (SIRVA Inc, 2004 10-K report)
4. Technology Issues
   e.g., "There are weaknesses in the Company's information technology ("IT") controls which makes the Company's financial data vulnerable to error or fraud; a lack of documentation regarding the roles and responsibilities of the IT function; lack of security management and monitoring and inadequate segregation of duties involving IT functions." (Earthshell Co., 2005 10-K report)
5. Training and Personnel
   e.g., "The Company lacks personnel with adequate expertise in accounting for income taxes in accordance with U.S. GAAP." (Westmoreland Coal Co., 2005 10-K report)

Appendix 2: Variable Definitions

Variables	Definition and data source
NUMBER	The number of analysts who make the forecasts in calculating the last median earnings
NUM	Natural logarithm of the number of analysts who make the forecasts in calculating the last median earnings
SKEW	The skewness of the earnings, calculated as the mean-median difference of price-scaled *earnings* from the prior 5 years (minimum of 4 years)
SPECIAL	An indicator variable equal to one if special items (Compustat # 17) are not equal to zero, zero otherwise
RET	Is the equally weighted market-adjusted cumulative return over the past 3 years
DA	The absolute value of abnormal total accruals. The calculation of DA is estimated using the modified Jones model of Larcker et al. (2007)
LOSS	An indicator variable equal to one if earnings are negative, zero otherwise
ECHG	Is the change in earnings (Compustat #18), calculated as the change in earnings over the previous year scaled by the previous year's earnings
NECHG	Is equal to one if ECHG is negative and zero otherwise

(*continued*)

Variables	Definition and data source
ABSECHG	The absolute change in earnings (Compustat #18), calculated as the absolute value of the change in earnings over the previous year scaled by the previous year's earnings
EPSVOL	The standard deviation of earnings before extraordinary items (Compustat #18) estimated using data from the prior 5 years (minimum of 4 years)
BM	The natural logarithm of book value (Compustat # 60) to the market (Compustat # 25* Compustat #199) of the firms
ROA	The ratio of return to asset, calculated as earnings before extraordinary items (Compustat #18) scaled by average total assets (Compustat # 6)
LEV	The ratio of debt (Compustat # 34 + Compustat # 9) to averaged total assets (Compustat # 6)
MV	Natural logarithm of the market value of equity (Compustat #199*#25)
FOREIGN	An indicator variable equal to one if foreign currency translation (COMPUSTAT Data Item #150) is not zero, zero otherwise
G1	An indicator variable that equals one if a firm discloses only Contained ICMWs in year t, and zero otherwise
G2	An indicator variable equal to one if a firm discloses Pervasive ICMWs in year t, and zero otherwise
ICMW	An indicator variable equal to one if the firm has disclosed internal control material weakness in year t, and zero otherwise
STALE	The days between the date of fiscal year end and the date of the last annual forecast issued (before the fiscal year end) by an analyst
ADJSTALE	The median value of the STALE variable for a firm per fiscal year
AGE	The number of years that a firm has been the CRSP database
BROKERAGE REPUTATION	The Institutional Investors' ranking of brokerage houses is used to proxy for brokerage reputation: a brokerage is considered to be of high reputation if it is consistently (at least for three prior years) ranked among the "Leaders" by Institutional Investors during the sample period. The following 12 brokers are identified as highly reputable: Bear Stearns & Co; CS First Boston; Goldman Sachs; Lehman Brothers; Donaldson, Lufkin & Jenrette; J.P. Morgan Securities; Merrill Lynch; Morgan Stanley & Co; Paine Webber; Prudential Securities; Salomon Brothers; and Smith Barney. The other brokerages are classified as less highly reputable

Appendix 3: Matching Procedure

The matches are identified by an algorithm that calculates the distance between each ICMW firm k and its matched non-ICMW counterpart j. Specifically, for each non-ICMW firm j in the same Fama-French industry as ICMW firm k, we calculate the percentage difference in assets, $\text{AssetsDIS} = \left|\frac{\text{Assets}_j - \text{Assets}_k}{\text{Assets}_k}\right|$, and the percentage difference in sales, $\text{SalesDIS} = \left|\frac{\text{Sales}_j - \text{Sales}_k}{\text{Sales}_k}\right|$. The sum of the two distance measures yields a matching score for each non-ICMW firm j that is in the same industry as ICMW firm k. From the set of matching scores that are less than two, we choose the non-ICMW firm with the smallest matching score for each ICMW firm; we then remove the matched pair (the ICMW and its non-ICMW counterpart) from the lists of ICMW and non-ICMW firms. In some cases, a single non-ICMW firm is the best match for several ICMW firms. In this case, we control for the order in which we match a non-ICMW firm to an ICMW firm by first calculating all possible matching scores, and then assigning the non-ICMW firm j to the ICMW firm k whose matching score is the smallest among the candidate ICMWs. For the remaining candidate non-ICMWs, we repeat the above steps using the remaining ICMW firms.

Appendix 4: Ordinary Least Squares (OLS) Method

For the main tests, we use Ordinary Least Squares (OLS) or linear least squares method to estimate the ACCURACY and BIAS parameters in the following linear regression models (Greene 2011). This method minimizes the sum of squared vertical distances between the observed responses (ICMW and control variables) in the dataset and the responses predicted by the linear approximation.

$$\begin{aligned}\mathrm{ACCURACY}_{\mathrm{i,t}} = {} & \alpha_0 + \alpha_1 \mathrm{ICMW}_{i,t} + \alpha_2 \mathrm{NUM}_{i,t} + \alpha_3 \mathrm{MV}_{i,t} + \alpha_4 \mathrm{LEV}_{i,t} \\ & + \alpha_5 \mathrm{ROA}_{i,t} + \alpha_6 \mathrm{BM}_{i,t} + \alpha_7 \mathrm{EPSVOL}_{i,(t-5,t-1)} + \alpha_8 \mathrm{ABSECHG}_{i,t} \\ & + \alpha_9 \mathrm{LOSS}_{i,t} + \alpha_{10} \mathrm{SPECIAL}_{i,t} + \alpha_{11} \mathrm{RET}_{i,(t-3,t-1)} + \alpha_{12} \mathrm{DA}_{i,t} + \varepsilon\end{aligned} \tag{63.3}$$

$$\begin{aligned}\mathrm{BIAS}_{\mathrm{i,t}} = {} & \beta_0 + \beta_1 \mathrm{ICMW}_{i,t} + \beta_2 \mathrm{NUM}_{i,t} + \beta_3 \mathrm{MV}_{i,t} + \beta_4 \mathrm{LEV}_{i,t} + \beta_5 \mathrm{BM}_{i,t} \\ & + \beta_6 \mathrm{SKEW}_{i,(t-5,t-1)} + \beta_7 \mathrm{ECHG}_{i,t} + \beta_8 \mathrm{LOSS}_{i,t} + \beta_9 \mathrm{SPECIAL}_{i,t} \\ & + \beta_{10} \mathrm{NECHG}_{i,t} + \beta_{11} \mathrm{RET}_{i,(t-3,t-1)} + \beta_{12} \mathrm{DA}_{i,t} + \varepsilon\end{aligned} \tag{63.4}$$

Appendix 5: Cook's Distance

We use Cook's distance to test the influence of outliers when performing least squares regression analysis. Cook's distance is calculated using the following formula (Cook 1977):

$$D_i = \frac{\sum_{j=1}^{n} \left(\widehat{Y_j} - \widehat{Y_{j(i)}}\right)\left(\hat{Y}_j - \hat{Y}_{j(i)}\right)^2}{p \cdot MSE}$$

where $\hat{Y}_J$ is the prediction from Eqs. 63.1 and 63.2 for observation j; $\hat{Y}_{j(i)}$ is the prediction for observation j from a refitted regression model in which observation i has been omitted; *MSE* is the mean square error of Eqs. 63.1 and 63.2. *P* is the number of fitted parameters in Eqs. 63.1 and 63.2. The higher the Cook's D is, the more influential the point. We choose 4/Nas our cutoff point. N is the number of observations in the regression tests. We exclude all observations with Cook's D higher than 4/N.

Appendix 6: Rank Regression

As a robust test, we reestimate the models using rank regression techniques to assess the sensitivity of the results to the underlying functional form assumption made by OLS (Cavanagh and Sherman 1998). Rank regression is one type of nonparametric tests that are widely used for studying populations that take on a ranked order. The use of nonparametric methods may be necessary when data have a ranking but no clear numerical interpretation. We first performance rank transformation (quintile rank) on all continuous variables used in Eqs. 63.1 or 63.2 and then performing OLS test on the ranks of the data instead of the data themselves.

References

Ashbaugh-Skaife, H., Collins, D. W., Kinney, W. R., & LaFond, R. (2008). Internal control deficiencies, remediation and accrual quality. *The Accounting Review, 83*, 217–250.

Beneish, D., Billings, M., & Hodder, L. (2008). Internal control weaknesses and information uncertainty. *The Accounting Review, 83*, 665–704.

Bouwman, M., & Frishkoff, P. (1995). The relevance of GAAP-based information: A case study exploring some uses and limitations. *Accounting Horizons, 9*, 22–47.

Brown, L. D., & Higgins, H. N. (2001). Managing earnings surprises in the U.S. versus 12 other countries. *Journal of Accounting and Public Policy, 20*, 371–398.

Brown, R., Chan, H., & Ho, Y. (2009). Analysts' recommendations: From which signal does the market take its lead? *Review of Quantitative Finance and Accounting, 33*, 91–111.

Cavanagh, C., & Sherman, R. P. (1998). Rank estimators for monotonic index models. *Journal of Econometrics, 84*(1), 351–381.

Chen, S., Matsumoto, D., & Rajgopal, S. (2011). Is silence golden? An empirical analysis of firms that stop giving quarterly earnings guidance. *Journal of Accounting and Economics, 51*(1–2), 134–150.

Cook, R. D. (1977). Detection of influential observations in linear regression. *Technometrics, 19*(1), 15–18.

Das, S., Levine, C. B., & Sivaramakrishnan, K. (1998). Earnings predictability and bias in analysts' earnings forecasts. *The Accounting Review, 73*, 277–294.

Dechow, P. M., & Sloan, R. G. (1997). Returns to contrarian investment: Tests of the naïve expectations hypotheses. *Journal of Financial Economics, 43*, 3–28.

Doyle, J., Ge, W., & McVay, S. (2007a). Accruals quality and internal control over financial reporting. *The Accounting Review, 82*, 1141–1170.

Doyle, J., Ge, W., & McVay, S. (2007b). Determinants of weaknesses in internal control over financial reporting and the implications for earnings quality. *Journal of Accounting and Economics, 44*, 193–223.

Duru, A., & Reeb, D. (2002). International diversification and analysts' forecast accuracy and bias. *The Accounting Review, 77*, 415–433.

Eames, M., & Glover, S. (2003). Earnings predictability and the direction of analysts' earnings forecast errors. *The Accounting Review, 78*, 707–724.

Fama, E., & French, K. (1997). Industry costs of equity. *Journal of Financial Economics, 43*, 153–194.

Fan, D., & Yeh, J. (2006). Analyst earnings forecasts for publicly traded insurance companies. *Review of Quantitative Finance and Accounting, 26*, 105–136.

Feng, M., Li, C., & McVay, S. (2009). Internal control and management guidance. *Journal of Accounting and Economics, 48*, 190–209.

Francis, J., Nanda, D., & Wang, X. (2006). Re-examining the effects of regulation fair disclosure using foreign listed firms to control for concurrent shocks. *Journal of Accounting and Economics, 41*, 271–292.

Ge, W., & McVay, S. (2005). On the disclosure of material weaknesses in internal control after the Sarbanes-Oxley act. *Accounting Horizon, 19*, 137–158.

Ghosh, A., & Lubberink, M. (2006). Timeliness and mandated disclosures on internal controls under section 404. Working paper, City University of New York.

Greene, W. H. (2011). *Econometric analysis* (11th ed.). New York: Macmillan.

Gu, Z., & Wu, J. (2003). Earnings skewness and analyst forecast bias. *Journal of Accounting and Economics, 35*, 5–29.

Ho, J. (2004). Analysts' forecasts of Taiwanese firms' earnings: Some empirical evidence. *Review of Pacific Basin Financial Markets and Policies, 7*, 571–597.

Huang, H., Willis, R., & Zang, A. (2005). Bold security analysts' earnings forecasts and managers' information flow. Working paper, Duke University.

Kanagaretnam, K., Lobo, G., & Mathieu, A. (2012). CEO compensation mix and analysts forecast accuracy and bias. *The Review of Quantitative Finance and Accounting*, forthcoming.

Ke, B., & Yu, Y. (2006). The effect of issuing biased earnings forecasts on analysts' access to management and survival. *Journal of Accounting Research, 44*, 965–1000.

Kim, J. B., Song, B. Y., & Zhang, L. (2009). Internal control quality and analyst forecast behavior: Evidence from SOX section 404 disclosures. Working paper, City University of Hong Kong.

Kross, W., Ro, B. T., & Schroeder, D. (1990). Earnings expectations: The analysts' information advantage. *The Accounting Review, 65*, 461–476.

Lang, M. H., & Lundholm, R. J. (1996). Corporate disclosure policy and analyst behavior. *The Accounting Review, 71*(4), 467–492.

Larcker, D. F., Richardson, S. A., & Tuna, A. I. (2007). Corporate governance, accounting outcomes, and organizational performance. *The Accounting Review, 82*, 963–1008.

Libby, R., Hunton, J., Tan, H., & Seybert, N. (2008). Relationship incentives and the optimistic/pessimistic pattern in analysts' forecasts. *Journal of Accounting Research, 46*, 173–198.

Lim, T. (2001). Rationality and analysts' forecast bias. *Journal of Finance, 56*, 369–385.

Lobo, G., Song, M., & Stanford, M. (2012). Accruals quality and analyst coverage. *Journal of Banking and Finance, 36*, 497–508.

Lys, T., & Soo, L. (1995). Analysts' forecast precision as a response to competition. *Journal of Accounting, Auditing and Finance, 10*, 751–765.

Mayew, W. (2008). Evidence of management discrimination among analysts during earnings conference calls. *Journal of Accounting Research, 46*, 627–659.

Previts, G., & Bricker, R. (1994). A content analysis of sell-side financial analyst company reports. *Accounting Horizons, 8*, 55–70.

Richardson, S., Teoh, S., & Wysocki, P. (2004). The walkdown to beatable analyst forecasts: The role of equity issuance and insider trading incentives. *Contemporary Accounting Research, 21*, 885–924.

Rogers, R., & Grant, J. (1997). Content analysis of information cited in reports of sell-side financial analysts. *Journal of Financial Statement Analysis, 3*, 17–30.

Sarbanes-Oxley Act. (2002). Public Law 107-204, 107th Cong., 2nd session. Washington, DC: GPO.

Shiller, R. J. (1999). Human behavior and the efficiency of the financial system. In J. B. Taylor & M. Woodford (Eds.), *Handbook of macroeconomics* (1st ed., Vol. 1, pp. 1305–1340). Amsterdam: Elsevier. Chapter 20.

Schipper, K. (1991). Commentary on analysts' forecasts. *Accounting Horizons, 5*, 105–121.

Solomon, D., & Frank, R. (2003, June 19). Stock analysis: 'You don't like our stock? You are off the list' – SEC sets new front on conflicts by taking aim at companies that retaliate against analysts. *Wall Street Journal*, C.1.

Solomon, D. (2005, May 2). Accounting rule exposes problems but draws complaints about cost. *Wall Street Journal*, A.1.

Womack, K. (1996). Do brokerage analysts' recommendations have investment value? *Journal of Finance, 54*, 137–157.

Zhang, F. (2006). Information uncertainty and analyst forecast behavior. *Contemporary Accounting Research, 23*, 1–28.

What Increases Banks Vulnerability to Financial Crisis: Short-Term Financing or Illiquid Assets?

64

Gang Nathan Dong and Yuna Heo

Contents

Abstract

It is not clear whether short-term financing increases banks' vulnerability to financial crisis or just reflects the weakness of banks' balance sheets. This chapter examines the role of short-term financing and assets with deteriorated quality in the financial crisis 2007–2009.

We apply logit and OLS econometric techniques to analyze the Federal Reserve Y-9C report data. We show that short-term financing is a response to the adverse economic shocks rather than a cause of the recent crisis. The likelihood of financial crisis actually stems from the illiquidity and low

G.N. Dong (✉)
Columbia University, New York, NY, USA
e-mail: gd2243@columbia.edu

Y. Heo
Rutgers Business School, Rutgers, The State University of New Jersey, Newark-New Brunswick, NJ, USA
e-mail: yunaheo@pegasus.rutgers.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_64,

creditworthiness of the investment. Our results are robust to endogeneity concerns when we use a difference-in-differences (DiD) approach with the Lehman bankruptcy in 2008 proxying for an exogenous shock.

Keywords
Financial crisis • Short-term financing • Debt maturity • Liquidity risk • Deterioration of bank asset quality

Banks will be required to hold emergency stocks of easy-to-sell assets . . . 'liquidity coverage ratio' that will require banks to hold buffers against a 30-day market crisis.
–Financial Times, January 8, 2012

64.1 Introduction

Does short-term financing increase banks' vulnerability to financial crisis? Banks have a special feature, what we called "lend long and borrow short," that the illiquid loans which typically constitute assets be financed by volatile and demandable deposits. This means that banks which want to finance illiquid investments have to borrow short-term debt; hence, the increasing illiquidity of the investment possibly causes higher exposure of liquidity risk and the susceptibility to crises. Recently, the role of short-term financing to financial crises is being debated again in depth.[1] In particular, short-term financing and its rollover risks have been recognized as a distinct characteristic of the financial crisis 2007–2009 since, during recent financial crisis, bank runs were incited by short-term creditors who were concerned about liquidity and solvency, unlike old-style bank runs instigated by uninsured depositors. This indicates that short-term financing exposes banks to rollover risk and can amplify financial crises. However, it is not clear whether the short-term financing indeed causes the vulnerability of financial crisis or just reflects the weakness of banks' balance sheets.

Why is short-term financing an important source of financing for banks? One well-known answer is that short-term financing is an equilibrium response to the agency problems.[2] Unless banks are fully equity financed, they have the wrong incentives when it comes to continuing or liquidating its project. Similar to a risk-shifting problem, banks have an incentive to continue excessively risky projects at the cost of debt holders. Therefore, banks' choice of maturity structure and the

[1]See Brunnermier (2009), Diamond and Rajan (2009), Gorton (2009), Gorton and Metrick (2012), Ivashina and Scharfstein (2009), Eisenbach (2010), Kashyap (2010), Benmelech and Dvir (2011), and Brunnermeier et al. (2011).

[2]See Diamond and Rajan (2001), Diamond (2004), Diamond and Rajan (2009), Eisenbach (2010), and Kashyap (2010).

implied exposure to rollover risk can play an important role of risk takings and bank runs in financial crisis. Banks can choose any combination of long-term and short-term debt to finance its investment.[3] While long-term debt has the same maturity as the projects' final payoff, short-term debt has to be rolled over after the additional information about the project's expected payoff and the liquidation value becomes available. Rollover risk arises since it may not be possible to satisfy all withdrawals of short-term creditors, even by liquidating all of the bank's assets.

However, the likelihood of financial crisis can stem not from short-term financing, but from the illiquidity and low creditworthiness of the investment being financed.[4] How assets with low quality and short-term financing affect the financial crisis 2007–2009? The answer can be summarized as follows: A substantial amount of mortgage-backed securities with exposure to subprime risk were kept on bank balance sheets, and banks financed these securities and other risky assets with short-term debt. As the housing market deteriorated, the perceived risk of mortgage-backed securities increased, and it became difficult to roll over short-term borrowing against those securities. The funding problems led to fire sales of assets, and these fire sales made banks obtain short-term financing even harder. These difficulties of bank funding spilled over to the entire economy, causing bank runs and financial crisis. Therefore, short-term financing could be a symptom of adverse economic shocks rather than a cause. In this setting, the investments being financed are becoming illiquid and as a result banks will increase short-term financing.[5]

Given the ambiguous causality, this chapter examines the role of short-term financing and assets with deteriorated quality in the financial crisis 2007–2009. Specifically, we try to answer two questions: (1) Does short-term financing predicts the financial crisis? (2) Do assets with deteriorated quality increase banks' vulnerability to financial crisis? Empirically we test whether short-term financing and assets with low creditworthiness are related to a bank's risk and profitability during

[3]Given the complexity of bank-risk taking, investors would have demanded a very high premium for financing the bank long term. Creditors would have been willing to hold short-term debt on the bank since that would give them the option to exit or get a higher premium if banks were appeared to be getting into trouble. So, creditors would have demanded lower premium for holding short-term secured debt in light of potential agency problems at banks. Thus, from the banker's perspective, financing with short-term debt claims is more attractive to the banks than issuing long-term claims. Clearly, banks should have been worried about the possibility that they could become illiquid and incapable of rolling over financing (Diamond and Rajan 2009).

[4]Diamond and Rajan (2001) argue that "It is no surprise that illiquid or poor quality investment when a bank or banking system is close to its debt capacity will result in a buildup of short-term debt."

[5]Eisenbach (2010) points out similar view, saying that "While short-term debt acts as an effective disciplining device when banks only face idiosyncratic risk, it is severely undermined when aggregate risk is added. The problem is that the disciplining effect is too weak in good sates and too powerful in bad states. This leads to a two-sided inefficiency: In good aggregate states the banks take excessive risks in the form of projects with negative net present value. Bad aggregate states suffer from fire sales as projects with positive net present value are liquidated."

the crisis.[6] The following results are found: (1) Short-term financing is a response, rather than a cause, of the recent financial crisis; (2) Assets that experienced asset deterioration are positively related to financial distress. The results are consistent to recent findings of banks' risk-taking behaviors and short-term financing literatures.[7] Finally, we check if the above results are driven by endogeneity concerns, namely, that significant omitted variables are correlated with both dependent and independent variables.

Section 64.2 summarizes the related literatures and hypotheses development. Section 64.3 explains the data and descriptive statistics. Section 64.4 presents the empirical results. Section 64.5 conducts robustness checks, and Sect. 64.6 concludes.

64.2 Related Literatures

What caused the financial crisis?[8] What keeps asset prices and lending depressed? According to Diamond and Rajan (2009), there are some consensus on the proximate causes of the crisis: (1) The US financial sector misallocated resources to real estate, financed through the issuance of exotic new financial instruments; (2) a significant portion of these instruments found their way, directly or indirectly, into commercial and investment bank balance sheets; (3) these investments were largely financed with short-term debt.

There are two approaches linking short-term financing to financial crisis. According to the first approach, taking on short-term debt increases banks' exposure to bank runs, and the bank is therefore more likely to fail. In the second approach, short-term financing is endogenous and is potentially the only financing available for lower-quality banks. Hence, the likelihood of failure is not necessarily driven by short-term debt itself but rather is a consequence of the bank's underlying economic conditions.

The first approach describes that short-term financing exposes banks to rollover risk and thus can cause and increase financial crisis. There is a growing agreement that an excessive buildup of short-term debt was a proximate cause of the financial crises. In the recent crisis, many studies point out the fragility embedded in

[6]Brunnermeier et al. (2011) suggest a more comprehensive measure of short-term financing. But in this paper we define short-term financing as short-term debt subtracted by long-term debt due in 1 year.

[7]Benmelech and Dvir (2011) and Bhattacharyya and Purnanandam (2011).

[8]According to Bhattacharyya and Purnanandam (2011), the analysis of the causes of the financial crisis in the USA during 2007–2009 has followed three distinct sources. One track focuses on the important role played by the shadow banking system in the securitization of mortgage loans (Gorton and Metrick 2012). A second track examines the incentives provided by the securitization process to skimp on adequate due diligence in the origination process (Keys et al. 2010). A third track focuses on the role of banks in originating unduly risky loans and also examines the incentives of bank management to originate risky loans (Fahlenbrach and Stulz 2011; Acharya and Richardson 2010).

short-term debt and rollover risk on the collapse of the housing and mortgage-backed securities market[9] as well as the increased short-term repurchase agreements (Brunnermeier and Oehmke 2013; He and Xiong 2009). As a result, some of them argue to regulate the use of short-term debt in the shadow banking system (Gotron and Metrick 2012). Borrowing with large amounts of short-term debt can lead to the threat of runs on banks, because there may be an externality across lenders (Diamond 2004). These runs on firms are very similar to the bank runs analyzed in Diamond and Dybvig (1983).[10]

In contrast to this view, some financial economists argue that short-term financing may be the optimal choice for borrowers who experience the deterioration in asset quality (Diamond and Rajan 2001; Eisenbach 2010; Benmelech and Dvir 2010). Diamond and Rajan (2001) suggest that maturity mismatch may be an optimal ex ante capital structure for banks when they cannot commit to fully repay investors once a project has been completed. In their model, if the projects being financed are seen as becoming less liquid due to an adverse shock to fundamentals, banks will find it harder to secure long-term financing from creditors and as a result will increase short-term financing. Short-term financing is hence a result of adverse fundamental economic shock rather a cause.[11]

In addition to the relation between short-term financing and financial crises, the finance literature abounds with attempts to quantify and explain risk-taking behavior of banks. Many studies have pointed out that risk-taking incentives among banks cause the financial crisis (see Bernanke 1983). However, the difficulty to accurately measure the banks' risk limits the access to information needed to evaluate the risk factors of banks.

Among recent studies, Bhattacharyya and Purnanandam (2011) document remarkable changes in the composition of risk taking by banks from 2006 to 2006.

[9]Changes in asset prices show up immediately on balance sheets and have an instant impact on the net worth of all constituents of the financial system. The net worth of financial intermediaries is especially sensitive to fluctuations in asset prices given the highly leveraged nature of such intermediaries' balance sheets. Procyclical leverage can be seen as a consequence of the active management of balance sheets by financial intermediaries who respond to changes in prices and measured risk. For financial intermediaries, their models of risk and economics capital dictate active management of their overall Value-at-Risk (VaR) through adjustments of their balance sheets (Adrian and Shin 2010).

[10]When banks do not have cash on hand to pay all depositors and must liquidate assets at a loss to pay those who withdraw first, this can lead all depositors to withdraw whenever they expect enough others to withdraw, even though this makes them collectively worse off. They all withdraw because the payments to those who withdraw impose losses on those who wait to withdraw after the bank runs out of money.

[11]Theories of imperfect capital markets (Bernanke and Gertler 1989; Kiyotaki and Moore 1997) argue that the banking sector is especially vulnerable to adverse selection and moral hazard, both caused by asymmetric information. The US financial sector is indeed becoming more vulnerable to systemic risk. Rajan (2005) argues that compensation structure at many financial firms may induce additional risk taking against underperforming their peers. This type of common exposure may increase the possibility of a severe tail event, if exacerbated by liquidity and informational frictions.

According to Bhattacharyya and Purnanandam (2011), the systematic risk doubled between 2000 and 2006, and banks with heavy involvement in residential mortgage lending and securitization have higher earnings per share. Banks heavily engaged in residential mortgage lending started exhibiting lower earnings and stock returns even prior to the crisis.

In addition, many recent studies have suggested that governance structures are related to risk-taking behavior of banks. Cheng et al. (2010) find evidence that excess compensation is correlated with risk taking and suggest that institutional investors both pushed managers towards a risky business model and rewarded them for it through higher compensation. Fahlenbrach and Stulz (2011) show that banks where the incentives of CEOs were better aligned with those of shareholders did not perform better during the crisis. Beltratti and Stulz (2011) investigate a sample of banks across the world and show that banks with more vulnerable financing with better governance performed worse during the crisis. Ellul and Yerramilli (2010) find that bank holding companies with strong and independent risk management functions tend to have lower enterprise-wide risk. However, none of these studies provide an answer to the question whether short-term financing or deteriorated assets contributed to the crisis.

64.3 Data and Summary Statistics

64.3.1 Data

This study focuses on commercial banks in the USA with SIC code 60 or 61. The sample is from 1993 to 2009, consisting of a panel database. Banks' monthly stock return from CRSP and financial statement data from Compustat and Federal Reserve from FR Y-9C are obtained.

Table 64.1 presents the variables' definition. Short-term financing is defined as short-term debt subtracted by long-term debt due in 1 year. Assets with deteriorated quality is defined as loans secured by real estate (scaled by total assets) subtracted by short-term financing. *Crisis* is an indicator variable that equals 1 for year 2007–2009 and zero otherwise. Std(ROA) is defined as a standard deviation of net income divided by total assets in a given firm and a given year. Earning per shares (EPS) is defined as net income divided by numbers of shares outstanding. Stock returns are holding period returns from CRSP monthly.

64.3.2 Descriptive Statistics

Table 64.2 provides the average and standard deviation of main variables from year 1993 to 2009. Short-term financing increased before 2007 and drastically decreased year between 2008 and 2009. Assets with deteriorated quality also increased and peaked at 2008. During the crisis period, the EPS plummeted and the standard deviation of earnings increased dramatically.

Table 64.1 Variable definitions

Variable	Calculation	Sources
Short-term financing	Short-term debt subtracted by long-term debt due in 1 year/total assets	Compustat
Assets with deteriorated quality	(Loan secured by real estate/total assets) – short-term financing	FR Y-9C report Compustat
Std(ROA)	Standard deviation of return on assets (net income/total assets)	Compustat FR Y-9C report
Crisis	An indicator variable that equals one for years 2007–2009, zero otherwise	
Ln (total assets)	Log of total assets	FR Y-9C report
Market to book equity	Market value of equity/book value of equity	CRSP indices monthly
Leverage	debt/equity	Compustat
Earning per share	Net income/numbers of outstanding shares	FR Y-9C report CRSP indices monthly
Stock return	Holding period of return	CRSP indices monthly

Table 64.2 Mean and standard deviation by year

	Short-term financing		Assets with deteriorated quality		Earnings per share	
Year	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
1993	0.03	0.05	0.34	0.15	1.77	1.74
1994	0.04	0.05	0.35	0.15	1.81	1.82
1995	0.04	0.05	0.38	0.15	1.67	0.98
1996	0.05	0.20	0.37	0.24	1.85	1.57
1997	0.05	0.25	0.38	0.28	1.77	1.15
1998	0.06	0.18	0.34	0.23	1.71	1.98
1999	0.08	0.17	0.35	0.24	1.65	2.17
2000	0.06	0.13	0.39	0.20	1.52	1.38
2001	0.06	0.14	0.40	0.21	1.49	1.14
2002	0.06	0.10	0.41	0.18	1.66	1.15
2003	0.06	0.09	0.41	0.18	1.69	1.13
2004	0.06	0.09	0.44	0.19	1.55	1.09
2005	0.06	0.09	0.46	0.19	1.67	1.20
2006	0.06	0.09	0.47	0.19	1.73	1.37
2007	0.07	0.07	0.47	0.18	1.44	1.36
2008	0.06	0.06	0.49	0.17	−0.27	3.26
2009	0.04	0.05	0.48	0.17	−0.92	4.99

Table 64.3 shows the summary statistics of variables used in study. The average of short-term financing is 0.056 and the average of assets with deteriorated quality 0.412. During the sample period, the average EPS is $1.39 and the average stock return is 1.1 %. Bhattacharyya and Purnanandam (2011) report the average EPS of $1.82 and the average return of 18 % during the peak period of 2000–2006.

Table 64.3 Summary statistics

Variable	Mean	Standard deviation	Minimum	Maximum
Short-term financing	0.056	0.122	0.000	4.992
Assets with deteriorated quality	0.412	0.205	−4.452	0.907
Std(ROA)	0.015	0.014	0.000	0.251
Crisis	0.175	0.380	0.000	1.000
Ln (total asset)	14.218	1.552	11.898	21.612
Market to book equity	1.470	1.220	−4.084	14.382
Leverage	10.995	5.125	−0.400	152.340
Earnings per share	1.391	2.105	−62.572	33.760
Stock return	0.011	0.097	−0.750	1.756

The correlation matrix is reported in Table 64.4. There is no large correlation between variables, except the correlation between short-term financing and assets with deteriorated quality.

64.4 Empirical Results

Table 64.5 shows the logit analysis results of short-term financing and assets with deteriorated quality. The dependent variable is a dummy variable that equals one if a bank bankrupted, filed for bankruptcy protection, or closed and received by the FDIC during year 2007–2009 and zero otherwise. The independent variables are short-term financing (short-term debt subtracted by long-term debt due in 1 year) and assets with deteriorated quality (loans secured by real estate subtracted by short-term financing). The control variables include log of total asset, market to book equity, and leverage. The results show that there are increasing assets with deteriorated quality, in contrast to decreasing short-term financing during the financial crisis years, indicating that short-term financing may not have contributed to the crisis, but reflect the weakness of banks' balance sheets.

Table 64.6 reports the regression results of standard deviation of ROA on short-term financing and assets with deteriorated quality. The dependent variable is the standard deviation of ROA and independent variables include short-term financing and assets with deteriorated quality. The control variables include log total asset, market to book equity ratio, and leverage ratio. The results show that short-term financing is not significantly related to the bank risk as measured by the standard deviation of ROA. Instead, assets with deteriorated quality during crisis years are positively related to the bank risk. This suggests that it is the assets with quality deterioration rather than short-term financing that actually caused bank distress. However, this finding does not exclude the possibility that exposures to rollover risk by short-term financing might also cause bank distress during the crisis.

Table 64.7 reports the regression results of earnings per share on short-term financing and assets with deteriorated quality. The dependent variable is EPS and the independent variables include short-term financing and assets with deteriorated

Table 64.4 Correlation table

	Short-term financing	Assets with deteriorated quality	Std(ROA)	Crisis	Log(total asset)	Market to book equity	Leverage	Earnings per share	Stock return
Short-term financing	1								
Assets with deteriorated quality	−0.682	1							
Std(ROA)	−0.018	−0.018	1						
Crisis	−0.003	0.152	0.130	1					
Log(total asset)	0.244	−0.315	0.094	0.204	1				
Market to book equity	0.069	−0.107	0.068	−0.158	0.308	1			
Leverage	−0.031	0.031	0.206	0.146	0.045	−0.025	1		
Earnings per share	0.120	−0.223	−0.339	−0.282	0.147	0.130	−0.355	1	
Stock return	0.021	−0.064	−0.032	−0.104	−0.006	0.094	0.050	0.044	1

Table 64.5 Logit analysis. This table shows the logit analysis results of short-term financing and assets with deteriorated quality. The dependent variable is a dummy variable that equals one if a bank bankrupted, filed for bankruptcy protection, or closed and received by the FDIC during year 2007–2009 and zero otherwise, indicating crisis years. Independent variables include short-term financing, which is defined as short-term debt subtracted by long-term debt due in 1 year and assets with deteriorated quality, which is defined as loans secured by real estate subtracted by short-term financing. The control variables include log total asset, market to book equity, and leverage

Dependent variable: crisis	(1)	(2)	(3)	(4)
Short-term financing	−0.0542		-2.469^{***}	
	(−0.23)		(−4.74)	
Assets with deteriorated quality		2.752^{***}		5.075^{***}
		(15.69)		(23.83)
Log(total asset)			0.477^{***}	0.677^{***}
			(23.72)	(30.11)
Market to book			-0.520^{***}	-0.579^{***}
			(−21.09)	(−21.69)
Leverage			0.0646^{***}	0.0705^{***}
			(9.74)	(9.36)
Intercept	-1.549^{***}	-2.768^{***}	-8.355^{***}	-13.63^{***}
	(−51.00)	(−32.02)	(−29.11)	(−34.88)
N	9,060	9,060	9,060	9,060

The parenthesis with ***, **, and * indicates its statistical significant level of 0.1 %, 1 %, and 5 %, respectively

quality. The control variables include log total asset, market to book equity ratio, and leverage ratio. The results suggest that short-term financing is positively related to banks' profitability. During the crisis years, short-term financing's contribution to bank profitability is significantly negative, indicating that short-term financing could be a symptom of adverse economic shocks rather than a cause of financial crisis, whereas assets with quality deterioration actually reduced banks' profits.

Table 64.8 reports the regression results of stock returns on short-term financing and assets with deteriorated quality. The results suggest that short-term financing is positively related to bank performance as measured by stock returns; however, during the crisis years, the effect turns to negative. The relation between assets with deteriorated quality and stock returns is always negative, regardless of the sample period.

64.5 Robustness Checks

As a robust check, we add time fixed effects to the regressions and report the results in Table 64.9. The dependent variables are standard deviation of ROA, EPS, and stock returns. Independent variables include short-term financing and assets with deteriorated quality. Consistent with the previous results, Table 64.9 shows that short-term financing during crisis is negatively related to standard deviation of ROA, suggesting that short-term financing might not actually contribute to financial crisis. The positive relation between assets with deteriorated quality and bank

Table 64.6 Regression of standard deviation of ROA on short-term financing. This table reports the regression results of standard deviation of ROA on short-term financing and assets with deteriorated quality. The dependent variable is a standard deviation of return on assets and independent variables include short-term financing and assets with deteriorated quality. The control variables include log total asset, market to book equity, and leverage

Dependent variable: Std(ROA)	(1)	(2)	(3)	(4)	(5)	(6)
Short-term financing	−0.00225					
	(−1.83)					
Assets with deteriorated quality				−0.00118		
				(−1.61)		
Short-term financing × crisis		0.000310	−0.00860			
		(0.07)	(−1.84)			
Assets with deteriorated quality × crisis					0.0107***	0.00868***
					(14.12)	(11.31)
Log(total asset)			0.000681***			0.000443***
			(6.65)			(4.44)
Market to book			0.000570***			0.000916***
			(4.47)			(7.15)
Leverage			0.000567***			0.000517***
			(19.93)			(18.06)
Intercept	0.0155***	0.0154***	−0.00136	0.0159***	0.0145***	0.00126
	(93.39)	(97.47)	(−0.96)	(46.86)	(89.03)	(0.90)
N	9,060	9,060	9,060	9,060	9,060	9,060
Adjusted R-square	0.0003	0.0001	0.052	0.000	0.021	0.065

The parenthesis with ***, **, and * indicates its statistical significant level of 0.1 %, 1 %, and 5 %, respectively

performance suggests that the low creditworthiness of the investment assets might have caused the bank distress during the crisis.

We now examine if the above results are driven by endogeneity concerns. Specifically, are significant omitted variable(s) correlated with both dependent and independent variables driving our results spuriously? In order to do so, we consider the bankruptcy filing of Lehman Brothers on September 15, 2011 (2008Q3), as an exogenous shock. We employ a difference-in-differences approach to analyze whether banks with different short-term financing and deteriorated assets react differently in terms of equity returns when they face the unexpected shock of Lehman Brothers bankruptcy. Accordingly, banks with more short-term financing and deteriorated assets are defined as the treatment group, and banks with less short-term financing and deteriorated assets are the control or non-treated group. We rank all commercial banks based on their short-term financing and deteriorated assets in the year 2007Q2 and Q3. The dummy variable of post-Lehman bankruptcy is set to one if the date is 2008Q4 and zero if the date is 2007Q4. Table 64.10 reports the DiD regression results.

In columns 1 and 2, the dummy variable of *top-quartile short-term financing* is set to one if a bank's short-term financing is in the top quartile and zero if it is in the bottom

Table 64.7 Regression of earnings per share on short-term financing. This table reports the regression results of earnings per share on short-term financing and assets with deteriorated quality. The dependent variable is earnings per share and independent variables include short-term financing and assets with deteriorated quality. The control variables include log total asset, market to book equity, and leverage

Dependent variable: EPS	(1)	(2)	(3)	(4)
Short-term financing		1.511***		
		(9.04)		
Assets with deteriorated quality				−0.929***
				(−8.94)
Short-term financing × crisis	−9.345***	−9.917***		
	(−14.58)	(−15.46)		
Assets with deteriorated quality × crisis			−3.079***	−2.765***
			(−30.18)	(−25.72)
Log(total asset)	0.244***	0.218***	0.260***	0.214***
	(17.51)	(15.33)	(19.66)	(15.17)
Market to book	0.0944***	0.0931***	0.0238	0.0326
	(5.44)	(5.39)	(1.41)	(1.93)
Leverage	−0.145***	−0.144***	−0.129***	−0.129***
	(−37.32)	(−37.05)	(−33.91)	(−34.09)
Intercept	−0.532**	−0.249	−0.658***	0.336
	(−2.75)	(−1.27)	(−3.58)	(1.57)
N	9,060	9,060	9,060	9,060
Adjusted R-square	0.176	0.183	0.233	0.240

The parenthesis with ***, **, and * indicates its statistical significant level of 0.1 %, 1 %, and 5 %, respectively

quartile A third dummy variable of *top-quartile short-term financing × post-Lehman bankruptcy* is the cross product of the previous two dummy variables. In columns 3 and 4, the dummy variable of *top-quartile deteriorated assets* is set to one if a bank's deteriorated asset is in the top quartile and zero if it is in the bottom quartile. A third dummy variable of *top-quartile deteriorated assets × post-Lehman bankruptcy* is the cross product of the previous two dummy variables. The coefficient estimates of the cross-product dummy for short-term financing in DiD specifications (1) and (2) are not statistically significant, whereas the coefficient estimates of the cross-product dummy for deteriorated assets in (3) and (4) are significantly negative, suggesting that a bank with asset experiencing deterioration of quality after the Lehman shock have much lower equity returns. This suggests that our results are not driven by omitted variables(s) that happen to be correlated with both dependent and independent variables.

64.6 Conclusion

It is not clear whether short-term financing indeed caused the banks' vulnerability to financial crisis or simply reflects the weakness of banks' balance sheets. Given the ambiguous causality, in this chapter we examine the role of short-term financing

Table 64.8 Regression of stock returns on short-term financing. This table reports the regression results of stock returns on short-term financing and assets with deteriorated quality. The dependent variable is earnings per share and independent variables include short-term financing and assets with deteriorated quality. The control variables include log total asset, market to book equity, and leverage

Dependent variable: stock return	(1)	(2)	(3)	(4)
Short-term financing		0.0271**		
		(3.19)		
Assets with deteriorated quality				−0.0199***
				(−3.65)
Short-term financing × crisis	−0.234***	−0.244***		
	(−7.23)	(−7.51)		
Assets with deteriorated quality × crisis			−0.0555***	−0.0488***
			(−10.42)	(−8.66)
Log(total asset)	−0.00116	−0.00164*	−0.00126	−0.00224**
	(−1.65)	(−2.28)	(−1.83)	(−3.03)
Market to book	0.00764***	0.00762***	0.00663***	0.00682***
	(8.71)	(8.69)	(7.50)	(7.70)
Leverage	0.00104***	0.00107***	0.00133***	0.00132***
	(5.28)	(5.42)	(6.69)	(6.68)
Intercept	0.00706	0.0121	0.00917	0.0305**
	(0.72)	(1.22)	(0.95)	(2.71)
N	9,060	9,060	9,060	9,060
Adjusted R-square	0.018	0.019	0.024	0.025

The parenthesis with ***, **, and * indicates its statistical significant level of 0.1 %, 1 %, and 5 %, respectively

and assets with deteriorated quality during the crisis of 2007–2009. We find that short-term financing is a response or a symptom of the adverse economic shocks rather than a cause of the recent crisis; instead, assets that experienced quality deterioration are positively related to bank distress. These results might suggest that financial crisis can stem not from short-term financing, but from the illiquidity and low creditworthiness of the investment. However, this finding does not rule out the possibility that exposures to rollover risk by short-term financing might also contribute to the distress in the banking sector during the crisis.

Appendix

We focus on all publicly traded commercial banks in the USA, namely, with SIC codes 60 and 61 and filing Federal Reserve Y-9C report in each quarter. FR Y-9C report collects basic financial data from a domestic bank holding company (BHC) on a consolidated basis in the form of a balance sheet, an income statement, and detailed supporting schedules, including a schedule of off balance-sheet items. By focusing on commercial banks, we do not include insurance companies, investment

Table 64.9 Fixed effects regression. This table presents the fixed effect regression results for robust checks. Dependent variables are standard deviation of ROA and earnings per share and stock returns. Independent variables include short-term financing and assets with deteriorated quality. The control variables include log total asset, market to book equity, and leverage

Dependent variables	Std(ROA) (1)	Std(ROA) (2)	EPS (3)	EPS (4)	Stock return (5)	Stock return (6)
Short-term financing	−0.00115		0.901***		0.0251**	
	(−0.94)		(5.57)		(3.12)	
Short-term financing × crisis	−0.0557***		0.551		−0.00521	
	(−9.11)		(0.68)		(−0.13)	
Assets with deteriorated quality		−0.00310***		−0.832***		−0.0200***
		(−3.72)		(−7.65)		(−3.67)
Assets with deteriorated quality × crisis		0.0143***		−1.847***		−0.0647***
		(6.69)		(−6.61)		(−4.63)
Log(total asset)	0.000520***	0.000357**	0.305***	0.258***	−0.000218	−0.00156*
	(4.87)	(3.15)	(21.69)	(17.40)	(−0.31)	(−2.10)
Market to book	0.00108***	0.00110***	0.0202	0.0196	0.00653***	0.00652***
	(8.16)	(8.32)	(1.16)	(1.13)	(7.54)	(7.53)
Leverage	0.000465***	0.000465***	−0.122***	−0.120***	0.000916***	0.000989***
	(16.09)	(16.05)	(−31.96)	(−31.57)	(4.82)	(5.20)
Intercept	0.00183	0.00357*	−1.696***	−0.485*	−0.00708	0.0262*
	(1.23)	(2.03)	(−8.66)	(−2.11)	(−0.73)	(2.28)
N	9,060	9,060	9,060	9,060	9,060	9,060
Adjusted R-square	0.049	0.044	0.155	0.166	0.008	0.013

The parenthesis with ***, **, and * indicates its statistical significant level of 0.1 %, 1 %, and 5 %, respectively

Table 64.10 DiD regressions of a bank's equity return prior and post-Lehman bankruptcy. We consider the bankruptcy filing of Lehman Brothers on September 15, 2011, (2008Q3) as an exogenous shock. We employ a difference-in-differences (DiD) approach (see Meyer 1995; Angrist and Krueger 1999 for detailed explanations of this methodology). Banks with more short-term financing and deteriorated assets are defined as the treatment group, and banks with less short-term financing and deteriorated assets are the control or non-treated group We rank the banks' short-term financing and deteriorated assets in 2007 separately. The dummy variable of *post-Lehman bankruptcy* is set to one if the date is 2008Q4 (the quarter after the bankruptcy filing of Lehman Brothers) and zero if the date is 2007Q4 (1 year before the bankruptcy filing of Lehman Brothers). In models (1) and (2), the dummy variable of *top-quartile short-term financing* is set to one if a bank's short-term financing is in the top quartile (75 percentile and above) and zero if it is in the bottom quartile (25 percentile and below). A third dummy variable of *top-quartile short-term financing* × *post-Lehman bankruptcy* is the cross product of the previous two dummy variables. In models (3) and (4), the dummy variable of *top-quartile deteriorated assets* is set to one if a bank's IBVC income is in the top quartile (75 percentile and above) and zero if it is in the bottom quartile (25 percentile and below). A third dummy variable of *top-quartile deteriorated assets* × *post-Lehman bankruptcy* is the cross product of the previous two dummy variables

Dependent variable: Stock return	(1)	(2)	(3)	(4)
Post-Lehman bankruptcy dummy	-0.926^{**}	-0.912^{**}	-1.199^{***}	-1.029^{***}
	(−2.15)	(−2.16)	(−4.16)	(−3.61)
Top-quartile short-term financing dummy	-1.173^{***}	−0.528		
	(−2.74)	(−1.26)		
Top-quartile short-term financing × post-Lehman bankruptcy dummy	−0.993	−1.039		
	(−1.32)	(−1.58)		
Top-quartile deteriorated asset dummy			-0.726^{**}	0.500^{*}
			(−2.01)	(1.83)
Top-quartile deteriorated asset × post-Lehman bankruptcy dummy			-1.059^{**}	-1.078^{**}
			(−2.02)	(−2.16)
Log(total asset)		−0.656		-2.639^{***}
		(−0.57)		(−2.82)
Market to book		−0.163		0.175
		(−0.62)		(0.93)
Leverage		−0.0216		−0.00908
		(−0.65)		(−0.28)
Intercept	-1.292^{***}	7.714	-1.853^{***}	23.13^{***}
	(−4.29)	(0.83)	(−9.19)	(3.03)
N	250	250	250	250
Adjusted R-square	0.16	0.26	0.13	0.24

A *t*-test is shown in the parenthesis with ***, **, and * indicating its statistical significant level of 1 %, 5 %, and 10 %, respectively

banks, investment management companies, and brokers. Our sample is from 1993 to 2009 and consisted of an unbalanced panel of 538 unique banks. We obtain a bank's monthly equity returns from CRSP and financial statement data from Compustat and FR Y-9C filed by a bank with the Federal Reserve.

Short-term financing (*STF*) is defined as short-term debt subtracted by long-term debt due in 1 year. Assets with deteriorated quality (*ADQ*) are defined as loans secured by real estates and scaled by total assets (*AT*) subtracted by short-term financing. *Crisis* is an indicator variable that equals one if a bank bankrupted, filed for bankruptcy protection, or closed and received by the FDIC during year 2007–2009 and zero otherwise. Bank risk is defined as the standard deviation of net income which is divided by total assets in a given firm and a given year. Earnings per shares are defined as net income divided by numbers of shares outstanding. Stock returns are holding period of returns from CRSP monthly.

After obtaining the data from CRSP, Compustat, and FR Y-9C, we merge them using the company identification (*PERMCO*) and calendar year. The correlation matrix reports no large correlation between the various independent variables. These low correlations suggest no evidence of multicollinearity in our following regressions, which is also confirmed by a low condition number of 2.67 (see Belsley et al. 1980). The first multivariate analysis is a logit regression:

$$\text{Crisis} = \beta_0 + \beta_1 STF + \beta_2 ADQ + \beta_3 \log(AT) + \beta_4 \text{Leverage} + \varepsilon \tag{64.1}$$

The second analysis is an OLS regression where the standard deviation of return on asset (*ROA*) serves as a proxy for bank risk:

$$\begin{aligned}\text{Std(ROA)} = \beta_0 + \beta_1 STF + \beta_2 ADQ + \beta_3 STF \times \text{Crisis} + \beta_4 ADQ \\ \times \text{Crisis} + \beta_5 \log(AT) + \beta_6 M2B + \beta_7 \text{Leverage} + \varepsilon\end{aligned} \tag{64.2}$$

To understand the effect of short-term financing on bank performance, we regress the bank's EPS on short-term financing activities and other control variables:

$$\begin{aligned}\text{EPS} = \beta_0 + \beta_1 STF + \beta_2 ADQ + \beta_3 STF \times \text{Crisis} + \beta_4 ADQ \times \text{Crisis} \\ + \beta_5 \log(AT) + \beta_6 M2B + \beta_7 \text{Leverage} + \varepsilon\end{aligned} \tag{64.3}$$

We also consider the stock returns as a measure of bank performance and conduct the following OLS regression:

$$\begin{aligned}\text{Return} = \beta_0 + \beta_1 STF + \beta_2 ADQ + \beta_3 STF \times \text{Crisis} + \beta_4 ADQ \times \text{Crisis} \\ + \beta_5 \log(AT) + \beta_6 M2B + \beta_7 \text{Leverage} + \varepsilon\end{aligned} \tag{64.4}$$

As a robustness check, we add the fixed effects to all regression models and redo our empirical analysis. Finally, we examine if the above results are driven by endogeneity concerns. Specifically, are significant omitted variable(s) correlated with both dependent and independent variables driving our results spuriously? In order to do so, we consider the bankruptcy filing of Lehman Brothers on September 15, 2011 (2008Q3) as an exogenous shock. We employ a difference-in-differences (DiD) approach (see Meyer 1995; Angrist and Krueger 1999 for detailed explanations of this methodology). We specifically analyze whether banks with different short-term financing and deteriorated assets react differently in terms of equity returns when they face the unexpected shock of Lehman Brothers bankruptcy. Accordingly, banks with more short-term financing and deteriorated assets are defined as the treatment group, and banks with less short-term financing and deteriorated assets are the control or non-treated group. We rank all commercial banks based on their short-term financing and deteriorated assets in the year 2007Q2 and Q3 (average over the two quarters). The dummy variable of post-Lehman bankruptcy is set to one if the date is 2008Q4 (the quarter after the bankruptcy filing of Lehman Brothers) and zero if the date is 2007Q4 (1 year before the bankruptcy filing of Lehman Brothers):

$$\begin{aligned}\text{Return} = \beta_0 + \beta_1\text{Post-Lehman} + \beta_2\text{Top-Quartile} + \beta_3\text{Top-Quartile} \\ \times \text{ Post-Lehman} + \beta_4\log(AT) + \beta_5 M2B + \beta_6\text{Leverage} + \varepsilon\end{aligned} \tag{64.5}$$

References

Acharya, V., & Richardson, M. (2010). Market failures and regulatory failures: Lessons from past and present crises. In M. Kawai & E. Prasad (Eds.), *Financial sector regulation and reforms in emerging markets*. Brookings Washington, DC: Institution Press (with Thomas Cooley, Matthew Richardson and Ingo Walter).

Adrian, T., & Shin, H. S. (2010). Liquidity and leverage. *Journal of Financial Intermediation, 19*, 418–437.

Angrist, J., & Krueger, A. (1999). Empirical strategies in labor economics. In: O. Ashenfelter & D. Card (Eds.), *Handbook of labor economics* (Vol. 3A, pp. 1277–1366).

Belsley, D., Kuh, E., & Welsch, R. (1980). *Regression diagnostics: Identifying influential data and sources of collinearity*. New York: Wiley.

Beltratti, A., & Stulz, R. M. (2011). *The credit crisis around the globe: Why did some banks perform better?* NBER working paper.

Benmelech, E., & Dvir, E. (2011). *Does short-term debt increase vulnerability to crisis? Evidence from the East Asian Financial Crisis* (NBER Working Papers 17468). National Bureau of Economic Research.

Bernanke, B. (1983). Nonmonetary effects of the financial crisis in the propagation of the great depression. *American Economic Review, 73*, 257–276.

Bernanke, B., & Gertler, M. (1989). Agency costs, net worth, and business fluctuations. *American Economic Review, 79*, 14–31.

Bhattacharyya, S., & Purnanandam, A. (2011). *Risk-taking by Banks: What did we know and when did we know it*? Working paper, University of Michigan.

Brunnermeier, M. K. (2009). Deciphering the liquidity and credit crunch 2007–2008. *Journal of Economic Perspectives, 23*, 77–100.

Brunnermeier, M. K., Gorton, G., & Krishnamurthy, A. (2011). *Liquidity mismatch*. Working paper, Princeton University.

Brunnermeier, M. K., & Oehmke, M. (2013). The maturity rat race. *Journal of Finance, 68*(2), 483–521.

Cheng, I., Hong, H., Scheinkman, J. A. (2010). *Yesterday's heroes: Compensation and creative risk taking*. NBER working paper 16176.

Diamond, D. W. (2004). Presidential address, committing to commit: Short-term debt when enforcement is costly. *Journal of Finance, 59*, 1447–1479.

Diamon, D. W., & Dybvig, P. H. (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy, 91*, 401–419.

Diamon, D. W., & Rajan, R. G. (2009). *The credit crisis: Conjectures about causes and remedies*. NBER working paper 14739.

Diamon, D. W., & Rajan, R. G. (2001). Banks, short-term debt and financial crises, policy implications and applications. *Carnegie-Rochester conference series on public policy, 54*, 37–71.

Eisenbach, T. M. (2010). *Rollover risk: Optimal but inefficient*. Working paper, Princeton University.

Ellul, A., & Yerramilli, V. (2010). *Stronger risk controls, lower risk: Evidence form U.S. Bank holding companies*. Working paper, Indiana University.

Fahlenbrach, R., & Stulz, R. M. (2011). Bank CEO incentives and the credit crisis. *Journal of Financial Economics, 99*, 11–26.

Gorton, G. (2009). *Haircuts*. NBER working paper 15273.

Gorton, G., & Metrick, A. (2012). Securitized banking and the run on repo. *Journal of Financial Economics, 104*, 425–451.

He, Z., & Xiong, W. (2009). *Dynamic debt runs* (NBER Working Papers 15482). National Bureau of Economic Research.

He, Z., & Xiong, W. (2012). Debt financing in asset markets. *American Economic Review, American Economic Association, 102*(3), 88–94.

Ivashina, V., & Scharfstein, D. S. (2009). *Bank lending during the financial crisis of 2008*. Working paper, Harvard University.

Kashayap, A. (2010). *Lessons from the financial crisis for risk management*. Working paper, University of Chicago.

Keys, B. J., Mukherjee, T., Seru, A., & Vig, V. (2010). Did securitization lead to lax screening? Evidence form subprime loans. *Quarterly Journal of Economics, 125*, 307–362.

Kiyotaki, N., & Moore, J. (1997). Credit cycles. *Journal of Political Economy, 105*, 211–248.

Meyer, B. (1995). Natural and quasi-experiments in economics. *Journal of Business and Economic Statistics, 13*, 151–161.

Rajan, R. G. (2005). *Has financial development made the world riskier?* NBER working papers 11728, National Bureau of Economics Research.

65 Accurate Formulas for Evaluating Barrier Options with Dividends Payout and the Application in Credit Risk Valuation

Tian-Shyr Dai and Chun-Yuan Chiu

Contents

Abstract

To price the stock options with discrete dividend payout reasonably and consistently, the stock price falls due to dividend payout must be faithfully modeled. However, this will significantly increase the mathematical difficulty since the post-dividend stock price process, the stock price process after the price falls due to dividend payout, no longer follows the lognormal diffusion process. Analytical pricing formulas are hard to be derived even for the simplest vanilla options.

The author was supported by NSC grant 100-2410-H-009-025

T.-S. Dai (✉)
National Chiao–Tung University, Taiwan, Republic of China
e-mail: d88006@csie.ntu.edu.tw; cameldai@mail.nctu.edu.tw; cameldai@gmail.com

C.-Y. Chiu
National Chiao–Tung University, Taiwan, Republic of China

Institute of Information Management, National Chiao Tung University, Taiwan, Republic of China
e-mail: r94922072@ntu.edu.tw

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_65,

This chapter approximates the discrete dividend payout by a stochastic continuous dividend yield, so the post-dividend stock price process can be approximated by another lognormally diffusive stock process with a stochastic continuous payout ratio up to the ex dividend date. Accurate approximation analytical pricing formulas for barrier options are derived by repeatedly applying the reflection principle. Besides, our formulas can be applied to extend the applicability of the first passage model — a branch of structural credit risk model. The stock price falls due to the dividend payout in the option pricing problem is analog to selling the firm's asset to finance the loan repayment or dividend payout in the first passage model. Thus, our formulas can evaluate vulnerable bonds or the equity values given that the firm's future loan/dividend payments are known.

Keywords

Barrier option • Option pricing • Stock option • Dividend • Reflection principle • Lognormal • Credit risk

65.1 Introduction

Black and Scholes (1973) arrive at their groundbreaking option pricing formula for non-dividend-paying stocks. Their option pricing model is extended to evaluate the credit risk of a defaultable firm by assuming that the firm defaults when its firm value fails to meet the debt obligation at maturity. Thus, both equities and the corporate debts can be viewed as contingent claims of the firm value, and their values can be evaluated by the aforementioned Black-Scholes option pricing formula (see Merton 1974). To deal with the dividend payout problem, Merton (1973) extends Black-Scholes formula by assuming that the stock pays a fixed continuous dividend yield. This assumption is used in the credit risk evaluation problem by allowing the firm to sell a fixed ratio of its asset continuously to finance the loan repayment or dividend payout (see Kim et al. 1993; Leland 1994). However, most dividends and coupon payments are paid discretely rather than continuously. Pricing stock options with discrete dividend payout seems to be first investigated in Black (1975). This discrete payout setting is analog to the setting that allows the firm to discretely sell its asset to finance the loan repayment or dividend payout under the credit risk evaluation problem. Although much financial literature alternatively assumes that the firm is restricted from selling its asset (see Leland 1994) or is allowed to sell its asset continuously at a fixed rate (see Kim et al. 1993; Leland 1994; Leland and Toft 1996), it is not the only — or even the typical — situation in the real-world financial markets. For example, British Petroleum Plc. sold its asset to finance the spill fund demanded by the US President Obama.[1] Recent news also report that many companies, like

[1]See "http://online.wsj.com/article/SB10001424-052748704862404575350830340543798.html" for the news entitled "BP Won't Issue New Equity to Cover Spill Costs."

Anglo American Plc. and Potash Corp. of Saskatchewan Inc., sold their asset to meet the required dividend payments.[2] Although this discrete payment setting might be more realistic, it incurs significant mathematical difficulty since the stock price process (or the firm's value process) becomes much more complicated (see Lando 2004).

Pricing stock options with discrete dividend payout has drawn a lot of attention in the literature. Frishling (2002) shows that the underlying stock price processes are usually modeled in three following different ways. **Model 1** suggests that the stock price minus the present value of future dividends over the life of the option follows the lognormal diffusion process (see Roll 1977; Geske 1979). **Model 2** suggests that the stock price plus the forward values of the dividends paid from today up to option maturity follows a lognormal diffusion process (see Heath and Jarrow (1988) and Musiela and Rutkowski (1997)). **Model 3** suggests that the stock price falls with the amount of dividend paid at the ex dividend date and follows the lognormal diffusion process between two ex dividend dates. Frishling (2002) argues that these three models are incompatible with each other and generate very different prices. In addition, Frishling (2002), Bender and Vorst (2001), and Bos and Vandermark (2002) argue that only model 3 can reflect the reality and generate consistent option prices. Except the aforementioned three models, Chiras and Manaster (1978) suggest that the discrete dividends can be transformed into a fixed continuous dividend yield. The stock option can then be analytically solved by Merton's formula (see Merton 1973). But Dai and Lyuu (2009) show that the pricing results of their approach can deviate significantly from those generated by **model 3**. The aforementioned observations suggest that the credit risk evaluation problem could be significantly mispriced if the aforementioned approaches (except **model 3**) are adopted.

On the other hand, pricing under **model 3** is mathematical intractable since the post-dividend stock price process, the stock price process after the price fall due to dividend payout, is no longer lognormally distributed. Bender and Vorst (2001), Bos and Vandermark (2002), Vellekoop and Nieuwenhuis (2006), Dai and Lyuu (2009), and Dai (2009) provide approximating analytical pricing formulas or efficient numerical methods for pricing vanilla options. But no announced papers derive analytical pricing formulas for pricing barrier stock options with discrete dividend payout.

A barrier option is a popular exotic option whose payoff depends on whether the path of the underlying stock has reached a certain predetermined price level called barrier. The study of pricing barrier options is of special interesting since this problem is dual to the problem of credit risk evaluation under the first passage model — a credit risk model that models the evolution of the firm value and forces the firm to default if its value is below a certain predefined default boundary.[3]

---

[2]See http://www.businessweek.com/news/2010-02-17/anglo-may-resume-dividend-after-asset-sales-analysts-say.html for the news entitled "Anglo May Resume Dividend After Asset Sales, Analysts Say" and http://fxnonstop.com/index.php/component/content/article/42555-myart26206 for the news entitled "Potash Weighs Asset Sales for Special Dividend."

[3]Note that the roles played by the stock price and the barrier in the barrier option pricing problem are analog to the roles played by the firm value and the default boundary in the first passage model.

Reiner and Rubinstein (1991) derive analytical pricing formula for the barrier option given the condition that the underlying stock pays no dividend or fixed continuous dividend yield. Thus, the process of the stock return can be expressed as a drifted Brownian motion, and the joint density of the extreme stock price over the option life and the stock price at the option maturity date can therefore be derived by taking advantages of the reflection principle and Girsanov's theorem. By using the risk-neutral variation technique, the pricing formulas can be derived with the aforementioned joint density function. Note that the Reiner and Rubinstein (1991) approach cannot be directly extended to price barrier options with discrete dividend under **model 3** or to evaluate the equity or the corporate debt values of a defaultable firm with discrete loan or dividend payout. In addition, deal with the discrete payout with the aforementioned models except **model 3** can produce unreasonable pricing results (see Frishling 2002). Besides, Gaudenzi and Zanette (2009) develop a tree model to address this pricing problem. The impact of discrete dividend is heuristically estimated by the linear interpolation method to avoid the combinatorial explosion problem due to non-recombining property of a bushy tree (see Dai 2009). However, it seems that their pricing results oscillate drastically due to nonlinearity error problem (see Figlewski and Gao 1999).

A dividend c_1 is paid at time t_1. The black solid curve denotes the stock price process prior to time t_1 (see Eq. 65.1). The stock price process after time t_1 is approximated by a stock price process that pays a continuous dividend q up to time t_1 (see Eq. 65.10). This approximation process is plotted in gray curve in time interval $[0, t_1]$ and is plotted in black dash curve after time t_1.

The major contribution of this chapter is to derive approximate analytical formulas for pricing barrier stock options with discrete dividend payout. As a by-product, our formula can be applied to evaluate the credit risk under the first passage model that allows the firm to sell its asset to finance the payout. Note that analytical pricing formula for the barrier option can be derived if the underlying stock pays no dividends (see Reiner and Rubinstein 1991). They show that the process of the stock return can be expressed as a Brownian motion with drift, so the joint density of the extreme stock price over a time interval and the stock price at the end of that time interval can be derived by taking advantages of the reflection principle for Brownian motion and Girsanov's theorem. Specifically speaking, assume that the stock price process $P(t)$ under the risk-neutral probability is given by

$$P(t) = P(0)e^{\lambda t+\sigma B(t)}, \tag{65.1}$$

where $\lambda \equiv r - 0.5\sigma^2$, r denotes the annual risk-free interest rate, σ denotes the volatility, and $B(t)$ denotes the standard Brownian motion. Then the process of the stock return can be expressed by a Brownian motion with drift: $\lambda t + \sigma B(t)$, as plotted in black solid curve in Fig. 65.1. In **model 3**, the stock price process is assumed to jump down with the dividend amount at an ex dividend date to avoid arbitrage. Specifically, assume that the i-th dividend c_i is paid at time t_i, where $t_1 < t_2 < t_3. \ldots$ In Fig. 65.1, we assume that the option maturity T is less than t_2 for simplicity. Thus, the stock price at time $t \in [t_1, t_2)$ must be expressed as

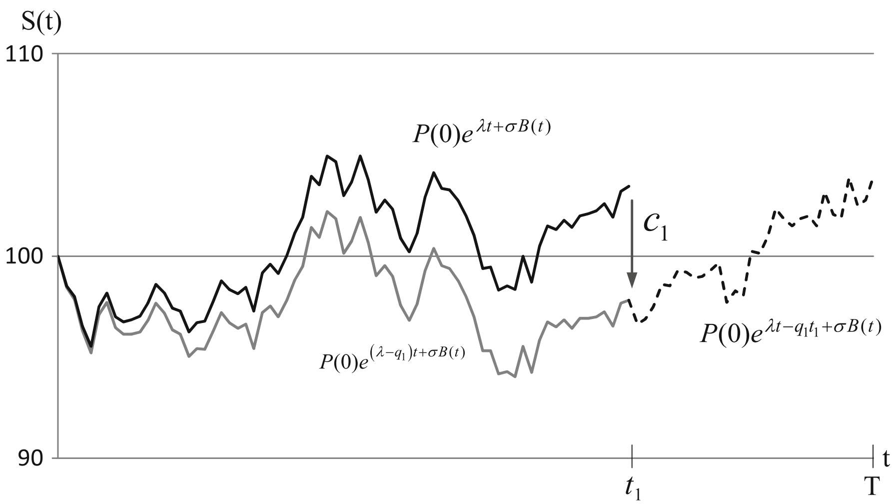


Fig. 65.1 Approximating the stock price process with discrete dividends

$$P(t) = \left(P(0)e^{\lambda t+\sigma B(t_1)} - c_1\right)e^{\lambda(t-t_1)+\sigma(B(t)-B(t_1))}, \tag{65.2}$$

which makes the stock return no longer a Brownian motion with drift. To address the problem, we extend Dai and Lyuu (2009) idea by setting a continuous dividend yield q to satisfy

$$P(0)e^{\lambda t+\sigma B(t_1)} - c_1 = P(0)e^{(\lambda-q)t+\sigma B(t_1)}.$$

q can be approximated solved by Taylor expansion to be a linear function of c_1 and $B(t_1)$. Due to the Markov property of the stock price process, the process of stock return after time t_1 can be approximated by another Brownian motion with drift: $\lambda t - qt_1 + \sigma B(t)$ as plotted in black dashed curve in Fig. 65.1. Since the processes of stock returns for the time interval $[0, t_1]$ and $[t_1, T]$ can be expressed by Brownian motions, the approximation analytical formulas can be derived by repeatedly applying the reflection principle. The numerical results suggest that our approximation pricing formulas provide accurate option pricing results. For evaluating the credit risk problem, our model can explicitly illustrate how the repayments by selling the firm asset influence the financial status of the firm and the credit qualities of other outstanding corporate debts.

The chapter is organized as follows. Section 65.2 introduces required background knowledge and preparations for formula derivation. In Sect. 65.3.1, we derive our approximation formula for the barrier option with single dividend. In Sect. 65.3.2, we extend our formulas to the multi-dividend case. Experimental results given in Sect. 65.4 verify the accuracy of our pricing formulas. Section 65.5 concludes the chapter.

65.2 Preliminaries

65.2.1 Pricing Formula for Barrier Option Without Dividend

Assume that a barrier stock option with strike K initiates at time 0 and matures at time T. The payoff of an up-and-out call option at maturity is as follows:

$$\text{payoff} = \begin{cases} (P(T)-K)^+ & \text{if } P_{\max} < B \\ 0 & \text{if } P_{\max} \geq B \end{cases},$$

where $(A)^+$ denotes max(A, 0), $P_{\max}$ denotes the maximum underlying stock price between time 0 and time T, and B denotes the barrier. Similarly, the payoff of a down-and-out call option at maturity is as follows:

$$\text{payoff} = \begin{cases} (P(T)-K)^+ & \text{if } P_{\min} < B \\ 0 & \text{if } P_{\min} \geq B \end{cases},$$

where $P_{\min}$ denotes the minimum stock price between time 0 and time T. For simplicity, our chapter will focus on up-and-out call option, and the extensions to other barrier options are straightforward.

To derive the pricing formula for up-and-out calls, we need to know whether the stock price process has ever hit the barrier. Note that the stock price process has ever hit the barrier during time interval [0, τ] if and only if the maximum stock price during time interval [0, τ] is greater than the barrier. The following theorem can be applied to derive the joint density of the stock price at time τ and the maximum stock price during the time interval [0, τ].

Theorem 65.1 *Let $\widetilde{W}(t) = \theta t + B(t)$ be a Brownian motion with a drift term θt and $\widetilde{M}(\tau) = \max_{0 \leq t \leq \tau} \widetilde{W}(t)$ be its maximum value over a certain time period* [0, τ]. *The joint density function of $\left(\widetilde{M}(\tau), \widetilde{W}(\tau)\right)$ is given by*

$$f_{\tilde{M}(\tau),\tilde{B}(\tau)}(m, w) = \begin{cases} \dfrac{2(2m-w)}{\tau\sqrt{2\pi\tau}} e^{\theta w - \frac{1}{2}\theta^2\tau - \frac{1}{2\tau}(2m-w)^2} & \text{if } m \geq w^+ \\ 0 & \text{otherwise.} \end{cases} \tag{65.3}$$

The support of this density function is illustrated in Fig . 65.2a.

This theorem can be derived by applying the reflection principle and Girsanov's theorem as discussed in Shreve (2007).

Reiner and Rubinstein (1991) derive analytical formulas for barrier options without discrete dividend payout by the aforementioned theorem. A detailed explanation of their derivation is given below since the derivation of our formula also takes advantage of their derivation. Define the stock return in Eq. 65.1 λt + $\sigma B(t)$ as $\sigma \hat{B}(t)$; that is, $\hat{B}(t)$ is a Brownian motion with drift term: $\hat{B}(t) \equiv \lambda t/\sigma + B(t)$. Define the maximum value of the Brownian motion

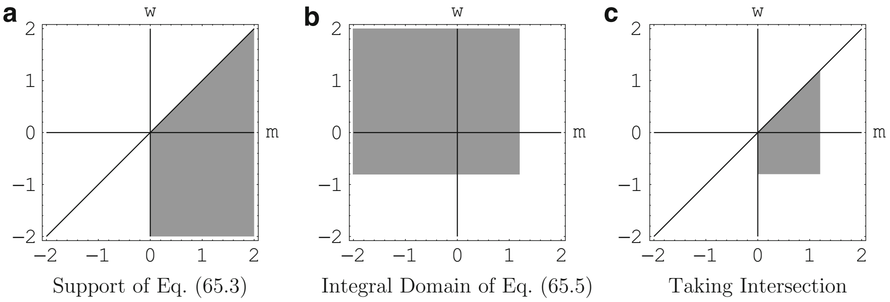


Fig. 65.2 Domain of double integral in Eq. 65.6

$\hat{M}(\tau)$ by $\hat{M}(\tau) = \max_{0\le t\le\tau}\widetilde{W}(t)$. Thus, the value of an up-and-out call option can be derived as follows:

$$C = e^{-rT}E\left\{\left(P(0)e^{\sigma\hat{W}(T)} - K\right)1_{\left\{\hat{W}(T)\ge k,\hat{M}(T)\le b\right\}}\right\}, \tag{65.4}$$

where k and b in Eq. 65.4 stand for $\frac{1}{\sigma}\log\frac{K}{P(0)}$ and $\frac{1}{\sigma}\log\frac{B}{P(0)}$, respectively. By substituting Eq. 65.3 into Eq. 65.4 with $\theta = \sigma/\lambda$, we have

$$C = \int_k^\infty\int_{-\infty}^b e^{-rT}\ \left(P(0)e^{\sigma w} - K\right)\ f_{\hat{M}}\ \ (T),\hat{W}(T)^{(m,\ \ w)dmdw} \tag{65.5}$$

$$= \int_k^b\int_{w^+}^b e^{-rT}\left(P(0)e^{\sigma w} - K\right)\frac{2(2m-w)}{T\sqrt{2\pi T}}e^{\theta w-\frac{1}{2}\theta^2T-\frac{1}{2T}(2m-w)^2}\ dmdw, \tag{65.6}$$

where the domain of integral in Eq. 65.5, i.e., $-\infty < m < b$ and $k < w < \infty$, is the support of the indicator function in Eq. 65.4 as illustrated in Fig. 65.2b. The domain of integral in Eq. 65.6 is the intersection of the support of the joint density function $f_{\hat{M}(T),\hat{W}(T)}(m,w)$ and the support of indicator function $1_{\left\{\hat{W}(T)\ge k,\hat{M}(T)\le b\right\}}$ as illustrated in Fig. 65.2c.

In the double integral formula Eq. 65.6, since only $f_{\hat{M}(T),\hat{W}(T)}(m,w)$ contains the variable m, $\int_{w^+}^b f_{\hat{M}(T),\hat{W}(T)}(m,w)\,dm$ can be evaluated first by the following lemma:

Lemma 65.2

$$\int_{v^+}^B\frac{2(2u-v)}{\beth\sqrt{2\pi\beth}}e^{\theta v-\frac{1}{2}\theta^2\beth-\frac{1}{2\beth}(2u-v)^2}du$$
$$= \frac{1}{\sqrt{2\pi\beth}}e^{\theta v-\frac{1}{2}\theta^2\beth-\frac{v^2}{2\beth}}\left(1 - e^{\frac{2B(v-B)}{\beth}}\right).$$

Panel (a) denotes the support of the density function of $f_{\tilde{M}(T),\tilde{W}(T)}$ in Eq. 65.3, i.e., a set of points $(m,\ w)$ that make $f_{\tilde{M}(T),\tilde{W}(T)}(m,w)$ nonzero.

Panel (b) denotes the domain of integral in Eq. 65.5, which is also the support of the indicator function of Eq. 65.4. Panel (c) denotes the intersection of shadow areas in Panel (a) and (b), which is the domain of integral in Eq. 65.6.

Proof

$$\begin{aligned}&\int_{v^+}^{B} \frac{2(2u-v)}{\text{ב}\sqrt{2\pi\text{ב}}} e^{\theta v - \frac{1}{2}\theta^2 \text{ב} - \frac{1}{2\text{ב}}(2u-v)^2} du\\ &= \frac{1}{\sqrt{2\pi\text{ב}}} e^{\theta v - \frac{1}{2}\theta^2\text{ב} - \frac{v^2}{2\text{ב}}}\left(1 - e^{\frac{2B(v-B)}{\text{ב}}}\right).\end{aligned}$$

By applying Lemma 65.2, Eq. 65.6 can be rewritten as

$$\begin{aligned}C = e^{-rT}\int_k^b & -\frac{K}{\sqrt{2\pi T}} e^{-\frac{w^2}{2T}+\theta w-\frac{T\theta^2}{2}} + \frac{K}{\sqrt{2\pi T}} e^{-\frac{w^2}{2T}+\theta w-\frac{T\theta^2}{2}+\frac{2b(w-b)}{T}}\\ &+ \frac{P'(0)}{\sqrt{2\pi T}} e^{-\frac{w^2}{2T}+\theta w+\sigma w-\frac{T\theta^2}{2}} - \frac{P'(0)}{\sqrt{2\pi T}} e^{-\frac{w^2}{2T}+\theta w+\sigma w-\frac{T\theta^2}{2}+\frac{2b(w-b)}{T}} dw.\end{aligned} \tag{65.7}$$

In Eq. 65.7, each term of the integrand is of the form $Le^{\phi_2 x^2+\phi_1 x+\phi_0}$, where ϕ_0, ϕ_1, ϕ_2, and L are all constants. The integral can be converted into the cumulative distribution function (CDF) of the standard normal distribution by the following identity:

$$\int_{-\infty}^{l} e^{\phi_2 x^2+\phi_1 x+\phi_0} dx = \sqrt{\frac{\pi}{-\phi_2}} e^{-\frac{\phi_1^2-4\phi_0\phi_2}{4\phi_2}} N\left(\frac{l-m}{s}\right),$$

where $\mathrm{m} = -\frac{\phi_1}{2\phi_2}$, $\mathrm{s} = \frac{1}{\sqrt{-2\phi_2}}$, and $N(\cdot)$ denotes the CDF of the standard normal distribution. Equation 65.31 in Appendix gives the derivation of this identity. Finally, we obtain the closed form pricing formula:

$$\begin{aligned}C = P(0)&\left\{N\left(\delta_+\left(T,\frac{P(0)}{K}\right)\right) - N\left(\delta_+\left(T,\frac{P(0)}{B}\right)\right)\right\}\\ &- Ke^{-rT}\left\{N\left(\delta_-\left(T,\frac{P(0)}{K}\right)\right) - N\left(\delta_-\left(T,\frac{P(0)}{B}\right)\right)\right\}\\ &- B\left(\frac{P(0)}{B}\right)^{-\frac{2r}{\sigma^2}}\left\{N\left(\delta_+\left(T,\frac{B^2}{KP(0)}\right)\right) - N\left(\delta_+\left(T,\frac{B}{P(0)}\right)\right)\right\}\\ &+ Ke^{-rT}\left(\frac{P(0)}{B}\right)^{-\frac{2r}{\sigma^2}+1}\left\{N\left(\delta-\left(T,\frac{B^2}{KP(0)}\right)\right) - N\left(\delta_-\left(T,\frac{B}{P(0)}\right)\right)\right\}\end{aligned} \tag{65.8}$$

where

$$\delta_\pm(\tau, s) \equiv \frac{1}{\sigma\sqrt{\tau}}\left[\log\ s + \left(r \pm \frac{1}{2}\sigma^2\right)\tau\right].$$

65.2.2 Approximate the Dividend-Paying Stock Price Process

Obviously, the stock price process before the first ex dividend date is as given in Eq. 65.1. Here we only focus on what happens after the first ex dividend date.

Extending the idea by Dai and Lyuu (2009), the stock price $P(t)$ between the first and second ex dividend date, defined as t_1 and t_2, respectively, can be approximated by a stock price process that pays a continuous dividend yield ζ_1 from time 0 to time t_1 as follows:

$$\begin{aligned} P(t) &= \left[P(0)e^{\lambda t_1+\sigma B(t_1)} - c_1\right] e^{\lambda(t-t_1)+\sigma(B(t)-B(t_1))} \\ &= P(0)e^{(\lambda-\zeta_1)t_1+\sigma B(t_1)+\lambda(t-t_1)+\sigma(B(t)-B(t_1))}, \end{aligned} \tag{65.9}$$

and hence,

$$P(0)e^{\lambda t_1+\sigma B(t_1)} - c_1 = P(0)e^{\lambda t_1+\sigma B(t_1)} \cdot e^{-\zeta_1 t_1}.$$

Since c_1 is usually small, ζ_1 is also small, and hence, $e^{-\zeta_1 t}$ can be approximated well by its first order Taylor expansion $1 - \zeta_1 t$. Replacing $e^{-\zeta_1 t_1}$ in the RHS of the above identity by $1 - \zeta_1 t_1$ gives

$$\begin{aligned} &P(0)e^{\lambda t_1+\sigma B(t_1)} - c_1 \approx P(0)e^{\lambda t_1+\sigma(B(t_1)-B(0))}(1-\zeta_1 t_1) \\ &\Rightarrow \zeta_1 \approx \frac{c_1 e^{-\lambda t_1}(1-\sigma(B(t_1)-B(0)))}{t_1 P(0)}, \end{aligned}$$

where $e^x \approx 1 + x$ is used again in the last approximation. By substituting $k_1 \equiv \frac{c_1 e^{-\lambda t_1}}{P(0)} + 1$, $\zeta_1 \equiv \frac{(k_1-1)(1-\sigma(B(t_1)-B(0)))}{t_1}$ into Eq. 65.9, we obtain the following approximating process for Eq. 65.9:

$$\begin{aligned} P(t) &\approx P(0)e^{\lambda t-(k_1-1)(1-\sigma(B(t_1)-B(0)))+\sigma(B(t_1)-B(0))+\sigma(B(t)-B(t_1))} \\ &= P(0)e^{(\lambda t-k_1+1)+k_1\sigma(B(t_1)-B(0))+\sigma(B(t)-B(t_1))} \end{aligned} \tag{65.10}$$

Note that the stock price after the second ex dividend date can be recursively defined by the aforementioned method. For example, the stock price $P(t)$ between the second ex dividend date t_2 and the third ex dividend date t_3 can be expressed as follows:

$$\begin{aligned} P(t) &= \left(P(t_1)e^{\lambda(t_2-t_1)+\sigma(B(t_2)-B(t_1))} - c_2\right)e^{\lambda(t-t_2)+\sigma(B(t)-B(t_2))} \\ &= P(0)e^{(\lambda-\zeta_1)t_1+\sigma(B(t_1)-B(0))}e^{(\lambda-\zeta_2)(t_2-t_1)+\sigma(B(t_2)-B(t_1))} \\ &\quad e^{\lambda(t-t_2)+\sigma(B(t)-B(t_2))} \\ &\approx P(0)e^{(\lambda t-k_1-k_2+2)+k_1k_2\sigma(B(t_1)-B(0))} \\ &\quad e^{+k_2\sigma(B(t_2)-B(t_1))+\sigma(B(t)-B(t_2))}, \end{aligned} \tag{65.11}$$

where

$$\zeta_2 \approx \frac{(k_2 - 1)[1 - k_1\sigma(B(t_1) - B(0)) - \sigma(B(t_2) - B(t_1))]}{t_2 - t_1}, \tag{65.12}$$

and

$$k_2 \equiv \frac{c_2 e^{-\lambda t_2 + k_1 + 1}}{P(0)} - 1.$$

65.3 Deriving Pricing Formulas

We will first derive the analytical approximating pricing formula in the single-discrete-dividend case in Sect. 65.3.1. Our approach used in Sect. 65.3.1 can be extended to obtain pricing formulas in general case. We will show that in Sect. 65.3.2 by the same approach as in Sect. 65.3.1 to derive the pricing formula in the multi-discrete-dividend case.

65.3.1 Single-Discrete-Dividend Case

Recall that combining Eq. 65.1 and 65.10, the stock price process can be approximated by

$$\dot{P}(t) \equiv \begin{cases} P(0)e^{\lambda t + \sigma(B(t) - B(0))} & 0 \le t < t_1 \\ P(0)e^{(\lambda t - k_1 + 1) + k_1\,\sigma(B(t_1) - B(0)) + \sigma(B(t) - B(t_1))} & t_1 \le t \le T, \end{cases} \tag{65.13}$$

where t_1 represents the only ex dividend date. Note that we denote the above approximating process for single-dividend case by $\dot{P}(t)$. Similarly, we will use $\ddot{P}(t)$ to denote the approximating process for two-dividend case later. Our goal is to compute the following expectation as an approximation to call value:

$$\dot{C} \equiv e^{-rT} E\left[\left(\dot{P}(T) - K\right) 1_{\{\Xi_1 \cap \Xi_2 \cap \Xi_3\}}\right], \tag{65.14}$$

where Ξ_1, Ξ_2 represent the events that stock price does not hit barrier B during time period $[0, t_1)$ and $[t_1, T]$, respectively, and Ξ_3 is the event that stock price is greater than strike price at maturity date. That is, the three events Ξ_1, Ξ_2, Ξ_3 are defined by

$$\begin{aligned} \Xi_1 &\equiv \{\dot{P}(t) < B| \ \ 0 \le t < t_1\}, \\ \Xi_2 &\equiv \{\dot{P}(t) < B| \ \ t_1 \le t \le T\}, \\ \Xi_3 &\equiv \{\dot{P}(T) > K\}. \end{aligned}$$

Similarly to Eq. 65.4, first we rewrite the process $\dot{P}(t)$ for $t \in [t_1, T]$ as follows:

$$\begin{aligned} \dot{P}(t) \ &= \ P(0)e^{(\lambda t - k_1 + 1) + k_1\,\sigma(B(t_1) - B(0)) + \sigma(B(t) - B(t_1))} \\ &= P'(0)e^{k_1\,\sigma\hat{B}}(t_1) + \sigma\hat{W}_1(t - t_1) \end{aligned}, \tag{65.15}$$

where we introduce a process $\hat{W}_1(t-t_1) \equiv \frac{\lambda}{\sigma}(t-t_1) + (B(t) - B(t_1))$ $\forall t \in (t_1, T)$, and $P'(0)$ stands for $P(0)e^{(1-k_1)(1+\lambda t_1)}$ for simplicity. Putting Eq. 65.13 and the above identity together, with $\hat{B}(t) \equiv \frac{\lambda t}{\sigma} + B(t)$ $\forall t \in [0, t_1]$ defined as aforementioned, we obtain

$$\dot{P}(t) = \begin{cases} P(0)e^{\sigma\hat{B}(t)} & 0 \le t < t_1 \\ P'(0)e^{k_1\sigma\hat{B}(t_1)+\sigma\hat{W}_1(t-t_1)} & t_1 \le t \le T. \end{cases} \quad (65.16)$$

Note that the two processes $\hat{B}(t)$ for $t \in [0,\ t_1]$ and $\hat{W}_1(t)$ for $t \in [t_1,\ T]$ are independent due to the Markov property of Brownian motion. Let $\hat{M}_1\eta_1 \equiv \max_{t_1 \le t \le T} \hat{W}_1(t-t_1)$ be the maximum value of $\hat{W}_1(t)$ over time period $[t_1, T]$, where η_1 denotes the abbreviation of $T - t_1$. Theorem 65.1 says that the joint density functions $f_{\hat{M}(t_1),\hat{B}(t_1)}$ and $f_{\hat{M}_1\eta_1,\hat{W}_1\eta_1}$ are given by

$$f_{\hat{M}(t_1),\hat{B}(t_1)}(m,w) = \begin{cases} \dfrac{2(2m-w)}{t_1\sqrt{2\pi t_1}}e^{\varepsilon}1 & \text{if } m \ge w^+ \\ 0 & \text{otherwise;} \end{cases} \quad (65.17)$$

$$f_{\hat{M}_1\eta_1,\hat{W}_1\eta_1}(m_1,w_1) = \begin{cases} \dfrac{2(2m_1-w_1)}{\eta_1\sqrt{2\pi\eta_1}}e^{\varepsilon}2 & \text{if } m_1 \ge w_1{}^+ \\ 0 & \text{otherwise,} \end{cases} \quad (65.18)$$

where $\varepsilon_1 = \theta w - \frac{1}{2}\theta^2 t_1 - \frac{1}{2t_1}(2m-w)^2$, $\varepsilon_2 = \theta w_1 - \frac{1}{2}\theta^2\eta_1 - \frac{1}{2\eta_1}(2m_1-w_1)^2$ and $\theta = \lambda/\sigma$. For convenience, from now on we will use the symbols f_0 and f_1 to represent $f_{\hat{M}(t_1),\hat{B}(t_1)}$ and $f_{\hat{M}_1\eta_1,\hat{W}_1\eta_1}$, respectively. Substituting Eq. 65.16 into Eq. 65.14, with the above joint density functions, we can compute · analytically. Note that Ξ_1, Ξ_2, Ξ_3 can be rewritten as

$$\begin{aligned} \Xi_1 &= \{\hat{M}(t_1) < b\}, \\ \Xi_2 &= \{\hat{M}_1\eta_1 < b' - k_1\hat{B}(t_1)\}, \\ \Xi_3 &= \{\hat{W}_1\eta_1 > k' - k_1\hat{B}(t_1)\}, \end{aligned}$$

where b, b', and k' represent $\frac{1}{\sigma}\log\frac{B}{P(0)}$, $\frac{1}{\sigma}\log\frac{B}{P'(0)}$ and $\frac{1}{\sigma}\log\frac{K}{P'(0)}$ for simplicity. Thus, the analytical pricing formula can be derived by law of iterated expectation as follows:

$$\begin{aligned} \dot{C} &= e^{-rT}E\left[E\left[\left(\dot{P}(T)-K\right)1_{\{\Xi_1\cap\Xi_2\cap\Xi_3\}}|\hat{B}(t_1),\hat{M}(t_1)\right]\right] \\ & e^{-rT}\int_{-\infty}^{b}\int_{w^+}^{b}\int_{k\prime-k_1w}^{b\prime-k_1w}\int_{w_1{}^+}^{b\prime\ -k_1w}\left(P'(0)e^{k_1\sigma w+\sigma w_1}-K\right) \\ & \cdot\frac{2(2m_1-w_1)}{\eta_1\sqrt{2\pi\eta_1}}e^{\varepsilon}2 \\ & \cdot\frac{2(2m-w)}{t_1\sqrt{2\pi t_1}}e^{\varepsilon}1_{dm_1\,dw_1\,dmdw.} \end{aligned} \quad (65.19)$$

To simplify the above multiple integral, first recall that Eq. 65.6 can be derived as Eq. 65.8 by first evaluating $\int_{w^+}^{b} f_{\hat{B}(t),\hat{M}(t)}(m,w)\,dm$ with lemma 2.2. Similarly, to simplify Eq. 65.19, first we evaluate $\int_{w^+}^{b} f_0(m,w)\,dm$ and $\int_{w_1^+}^{b\prime-k_1 w} f_1(m_1,w_1)$, and then obtain a double integral similar to Eq. 65.7. We will rewrite this double integral in terms of a bivariate normal CDF, just like the way to rewrite Eq. 65.7 in terms of the standard normal CDF. Like single-variate case, the CDF of a bivariate normal distribution can be approximated efficiently and accurately by a number of numerical schemes.[4] However, literatures only give the approximation for

$$\int_{-\infty}^{d}\int_{-\infty}^{c} f_{X,Y}(x,y)dxdy,$$

while the double integral we deal with is of the form

$$\int_{-\infty}^{d}\int_{-\infty}^{h(x)} f_{X,Y}(x,y)dxdy,$$

where $f_{X,\,Y}$ stands for the joint density function of a bivariate normal distribution, c and d are constants, and the value of h depends on x. To tackle this problem, we use a change of variable to adapt the double integral. Specifically, Eq. 65.19 can be simplified as follows:

$$\begin{aligned}\dot{C} = e^{-rT}&\int_{-\infty}^{b}\int_{k\prime-k_1 w}^{b\prime-k_1 w}\left(P'(0)e^{k_1\sigma w+\sigma w_1}-K\right)\\ &\cdot\left(\int_{w_1^+}^{b\prime-k_1 w}\frac{2(2m_1-w_1)}{\eta_1\sqrt{2\pi\eta_1}}e^{\varepsilon}2dm_1\right)\end{aligned}\qquad(65.20)$$

$$\cdot\left(\int_{w^+}^{b}\frac{2(2m-w)}{t_1\sqrt{2\pi t_1}}e^{\varepsilon}1\,dm\right)dw_1\,dw.\qquad(65.21)$$

By applying lemma 65.2 twice on Eqs. 65.20 and 65.21 and the change of variable

$$\begin{cases}x = w_1 + k_1 w\\ y = w,\end{cases}$$

[4]See, for example, Hull (2003).

Table 65.1 Coefficients of the exponential terms of $Q(i)$

$Q(1) = -\frac{K}{2\pi\sqrt{\eta_1 t_1}} e^{-\frac{y^2}{2t_1} + \theta y + \theta(x - yk_1) - \frac{1}{2}\theta^2\eta_1 - \frac{\theta^2 t_1}{2} - \frac{(x-yk_1)^2}{2\eta_1}}$,
$Q(2) = \frac{K}{2\pi\sqrt{\eta_1 t_1}} e^{-\frac{y^2}{2t_1} + \theta y + \theta(x - yk_1) - \frac{1}{2}\theta^2\eta_1 - \frac{\theta^2 t_1}{2} - \frac{(x-yk_1)^2}{2\eta_1} + \frac{2(x-b\prime)(b\prime-yk_1)}{\eta_1}}$,
$Q(3) = \frac{K}{2\pi\sqrt{\eta_1 t_1}} e^{-\frac{y^2}{2t_1} + \theta y + \theta(x - yk_1) - \frac{1}{2}\theta^2\eta_1 - \frac{\theta^2 t_1}{2} - \frac{(x-yk_1)^2}{2\eta_1} + \frac{2b(y-b)}{t_1}}$,
$Q(4) = -\frac{K}{2\pi\sqrt{\eta_1 t_1}} e^{-\frac{y^2}{2t_1} + \theta y + \theta(x - yk_1) - \frac{1}{2}\theta^2\eta 1 - \frac{\theta^2 t_1}{2} - \frac{(x-yk_1)^2}{2\eta_1} + \frac{2(x-b\prime)(b\prime-yk_1)}{\eta_1} + \frac{2b(y-b)}{t_1}}$,
$Q(5) = \frac{P'(0)}{2\pi\sqrt{\eta_1 t_1}} e^{-\frac{y^2}{2t_1} + \theta y + x\sigma + \theta(x - yk_1) - \frac{1}{2}\theta^2\eta 1 - \frac{\theta^2 t_1}{2} - \frac{(x-yk_1)^2}{2\eta_1}}$,
$Q(6) = -\frac{P'(0)}{2\pi\sqrt{\eta_1 t_1}} e^{-\frac{y^2}{2t_1} + \theta y + x\sigma + \theta(x - yk_1) - \frac{1}{2}\theta^2\eta_1 - \frac{\theta^2 t_1}{2} - \frac{(x-yk_1)^2}{2\eta_1} + \frac{2(x-b\prime)(b\prime-yk_1)}{\eta_1}}$,
$Q(7) = -\frac{P'(0)}{2\pi\sqrt{\eta_1 t_1}} e^{-\frac{y^2}{2t_1} + \theta y + x\sigma + \theta(x - yk_1) - \frac{1}{2}\theta^2\eta_1 - \frac{\theta^2 t_1}{2} - \frac{(x-yk_1)^2}{2\eta_1} + \frac{2b(y-b)}{t_1}}$,
$Q(8) = \frac{P'(0)}{2\pi\sqrt{\eta_1 t_1}} e^{-\frac{y^2}{2t_1} + \theta y + x\sigma + \theta(x - yk_1) - \frac{1}{2}\theta^2\eta_1 - \frac{\theta^2 t_1}{2} - \frac{(x-yk_1)^2}{2\eta_1} + \frac{2(x-b\prime)(b\prime-yk_1)}{\eta_1} + \frac{2b(y-b)}{t_1}}$.

we have[5]

$$
\begin{aligned}
\dot{C} = \; & e^{-rT} \int_{-\infty}^{b} \int_{k\prime}^{b\prime} \left(P'(0)e^{\sigma x} - K\right) \\
& \cdot \frac{1}{\sqrt{2\pi\eta_1}} e^{\theta\left(x - k_1 y\right) - \frac{1}{2}\theta^2 \eta_1 - \frac{(x-k_1 y)^2}{2\eta_1}} \\
& \left(1 - e^{\dfrac{2(b' - k_1 y)\left(x - k_1 y - (b' - k_1 y)\right)}{\eta_1}}\right) \\
& \cdot \frac{1}{\sqrt{2\pi t_1}} e^{\theta y - \frac{1}{2}\theta^2 t_1 - \frac{y2}{2t_1}} \left(1 - e^{\frac{2b(y-b)}{t_1}}\right) \mathrm{dxdy}. \\
= \; & e^{-rT} \int_{-\infty}^{b} \int_{k\prime}^{b\prime} \sum_{i=0}^{8} Q(i)\, dxdy,
\end{aligned}
\tag{65.22}
$$

where the integrands $Q(1)$, $Q(2)$, ···, $Q(8)$, by immediately expanding whole the integrand, are given in Table 65.1.

Since each of the integrands $Q(1)$, $Q(2)$, ···, $Q(8)$ is a quadratic form of x, y taking exponential and multiplied by a constant, we rewrite the double integral of each integrand in terms of a bivariate normal CDF. Thus, all the double integrals of $Q(1)$, $Q(2)$, ···, $Q(8)$ can be evaluated. To simplify these double integrals, we introduce the following formula:

[5]The Jacobian determinant $\frac{\partial(w_1,w)}{\partial(x,y)} = 1$.

Lemma 65.3 *Let* $F_{X_1,X_2}(a,b,\Omega)$ *be the CDF of a bivariate normal distributed random vector* (X_1, X_2). *That is,*

$$\begin{aligned} F_{X_1,X_2}(a,b,\boldsymbol{\Omega}) &\equiv \int_{-\infty}^{a}\int_{-\infty}^{b} f_{X_1,X_2}(x_1,x_2)dx_1dx_2 \\ &= \int_{-\infty}^{a}\int_{-\infty}^{b} \frac{1}{2\pi\sqrt{|\Omega|}}e^{-\frac{1}{2}\mathrm{X}^T\Omega\mathrm{X}}dx_1dx_2 \end{aligned}$$

where $\mathbf{x}$ *and* Ω *denote* $(x_1, x_2)^T$ *and the covariance matrix, respectively, and both* X_1 *and* X_2 *have mean* 0 *and variance* 1. *The following double integral can be expressed in terms of* $F_{X1,\ X2}$ *as*

$$\begin{aligned} &G(p,q,\phi_1,\phi_2,\phi_3,\phi_4,\phi_5,\phi_6) \\ &\equiv \int_{-\infty}^{q}\int_{-\infty}^{p} e^{\phi_1x^2+\phi_2xy+\phi_3y^2+\phi_4x+\phi_5y+\phi_6}dxdy \\ &= \frac{2\pi}{\sqrt{\beth}}\exp\left(\phi_6+\frac{\phi_2\phi_4\phi_5-\phi_3\phi_4^2-\phi_1\phi_5^2}{\beth}\right) \\ &\quad F_{\Upsilon_1,\Upsilon_2}\left(\frac{\sqrt{\beth}p+\dfrac{2\phi_3\phi_4-\phi_2\phi_5}{\sqrt{\beth}}}{\sqrt{-2\phi_3}},\frac{\sqrt{\beth}q+\dfrac{2\phi_1\phi_5-\phi_2\phi_4}{\sqrt{\beth}}}{\sqrt{-2\phi_1}},\Omega\right), \end{aligned}$$

where

$$\beth \equiv 4\phi_1\phi_3-\phi_2^2, \tag{65.23}$$

$$\Omega \equiv \begin{pmatrix} 1 & \dfrac{\phi_2}{2\sqrt{\phi_1\phi_3}} \\ \dfrac{\phi 2}{2\sqrt{\phi_1\phi_3}} & 1 \end{pmatrix}. \tag{65.24}$$

Proof To make $\phi_1x^2 + \phi_2xy + \phi_3y^2 + \phi_4x + \phi_5y + \phi_6$ equal $\mathbf{z}^T\vartheta\mathbf{z} + \mathbf{z}^T\beta + \delta$, we set

$$\vartheta \equiv \begin{pmatrix} \phi_1 & \dfrac{\phi_2}{2} \\ \dfrac{\phi_2}{2} & \phi_3 \end{pmatrix}, \beta \equiv \begin{pmatrix} \phi_4 \\ \phi_5 \end{pmatrix}, \delta \equiv \phi_6.$$

By Eqs. 65.38, 65.39, 65.42, and 65.43, we have

$$
\begin{aligned}
\mathrm{m} &= -\frac{1}{4\phi_1\phi_3-\phi_2^2}\begin{pmatrix}2\phi_3\phi_4-\phi_2\phi_5\\ 2\phi_1\phi_5-\phi_2\phi_4\end{pmatrix},\\
\delta' &= \phi_6+\frac{\phi_2\phi_4\phi_5-\phi_3\phi_4^2-\phi_6\phi_5^2}{4\phi_1\phi_3-\phi_2^2},\\
\Phi_{1,1} &= \sqrt{\frac{-2\phi_3}{4\phi_1\phi_3-\phi_2^2}},\\
\Phi_{2,2} &= \sqrt{\frac{-2\phi_1}{4\phi_1\phi_3-\phi_2^2}},\\
\Omega &= \begin{pmatrix}1 & \frac{\phi_2}{2\sqrt{\phi_1\phi_3}}\\ \frac{\phi_2}{2\sqrt{\phi_1\phi_3}} & 1\end{pmatrix},\\
|-\vartheta| &= \frac{4\phi_1\phi_3-\phi_2^2}{4}.
\end{aligned}
$$

By substituting all above equations into Eq. 65.45, we obtain

$$
\begin{aligned}
&\int_{-\infty}^{q}\int_{-\infty}^{p} e^{\phi_1x^2+\phi_2xy+\phi_3y^2+\phi_4x+\phi_5y+\phi_6}dxdy\\
&=\frac{2\pi}{\sqrt{\beth}}\exp\left(\phi_6+\frac{\phi_2\phi_4\phi_5-\phi_3\phi_4^2-\phi_1\phi_5^2}{\beth}\right)\\
&F_{\Upsilon_1\Upsilon_2}\left(\frac{\sqrt{\beth}p+\frac{2\phi_3\phi_4-\phi_2\phi_5}{\sqrt{\beth}}}{\sqrt{-2\phi_3}},\frac{\sqrt{\beth}q+\frac{2\phi_1\phi_5-\phi_2\phi_4}{\sqrt{\beth}}}{\sqrt{-2\phi_1}},\Omega\right)
\end{aligned}
$$

where $\beth = 4\phi_1\phi_3-\phi_2^2$. Thus, the proof is completed.

With lemma 65.3, the double integrals of $Q(1)$, $Q(2)$, ⋯, $Q(8)$ can be evaluated immediately. For example, the integral of $Q(1)$ is given by

$$
\begin{aligned}
&\int_{-\infty}^{b}\int_{k\prime}^{b\prime}Q(1)dxdy=\int_{-\infty}^{b}\int_{k\prime}^{b\prime}-\frac{K}{2\pi\sqrt{\eta_1t_1}}\\
&e^{-\frac{y^2}{2t_1}+\theta y+\theta(x-yk_1)-\frac{1}{2}\theta^2\eta_1-\frac{\theta^2t_1}{2}-\frac{(x-yk_1)^2}{2\eta 1}}dxdy\\
&=-\frac{K}{2\pi\sqrt{\eta_1t_1}}\left(\int_{-\infty}^{b}\int_{-\infty}^{b\prime}e^{\varepsilon_3}dxdy\right.\\
&\quad\left.-\int_{-\infty}^{b}\int_{-\infty}^{k\prime}e^{\varepsilon_4}dxdy\right),
\end{aligned}
\tag{65.25}
$$

where $\varepsilon_3 = -\frac{1}{2\eta_1}x^2 + \frac{k_1}{\eta_1}xy - \left(\frac{k_1^2}{2\eta_1} + \frac{1}{2t_1}\right)y^2 + \theta x + (\theta - \theta k_1)y - \frac{T\theta^2}{2}$, and $\varepsilon_4 = -\frac{1}{2\eta_1}x^2 + \frac{k_1}{\eta_1}xy - \left(\frac{k_1^2}{2\eta_1} + \frac{1}{2t_1}\right)y^2 + \theta x + (\theta - \theta k_1)y - \frac{T\theta^2}{2}$. Note that the integrand is a quadratic form of x, y taking exponential. Let $\phi_1(1)$, $\phi_2(1)$, $\phi_3(1)$, $\phi_4(1)$, $\phi_5(1)$, $\phi_6(1)$ be the coefficients of x^2, xy, y^2, x, y and the constant term of the quadratic form, respectively. That is,

$$
\begin{aligned}
\phi_1(1) &\equiv -\frac{1}{2\eta_1}, \\
\phi_2(1) &\equiv \frac{k_1}{\eta_1}, \\
\phi_3(1) &\equiv -\left(\frac{k_1^2}{2\eta_1} + \frac{1}{2t_1}\right), \\
\phi_4(1) &\equiv \theta, \\
\phi_5(1) &\equiv (\theta - \theta k_1), \\
\phi_6(1) &\equiv -\frac{T\theta^2}{2}.
\end{aligned}
$$

By Lemma 3.1, Eq. 65.25 can be rewritten in terms of bivariate normal CDF as

$$
\equiv D(1)\begin{bmatrix} G(b', b, \phi_1(1), \phi_2(1), \phi_3(1), \phi_4(1), \phi_5(1), \phi_6(1)) \\ -G(k', b, \phi_1(1), \phi_2(1), \phi_3(1), \phi_4(1)\phi_5(1), \phi_6(1)) \end{bmatrix},
$$

where $D(1) \equiv -\frac{K}{2\pi\sqrt{\eta_1 t_1}}$, and $\daleth(1)$, $\Omega(1)$ are obtained by substituting $a(1)$, $a(2)$, ⋯, $a(6)$ into Eqs. 65.23, 65.24, respectively. That is,

$$
\begin{aligned}
\daleth(1) &\equiv 4\phi_1(1)\phi_3(1) - \phi_2(1)^2, \\
\Omega(1) &\equiv \begin{pmatrix} 1 & \dfrac{\phi_2(1)}{2\sqrt{\phi_1(1)\phi_3(1)}} \\ \dfrac{\phi_2(1)}{2\sqrt{\phi_1(1)\phi_3(1)}} & 1 \end{pmatrix}.
\end{aligned}
$$

Similarly, we have

$$
\begin{aligned}
&\int_{-\infty}^{b}\int_{k\prime}^{b\prime} Q(i)dxdy \\
&\quad = D(i)[G(b', b, \phi_1(i), \phi_2(i), \phi_3(i), \phi_4(i), \phi_5(i), \phi_6(i)) \\
&\qquad - G(k', b, \phi_1(i), \phi_2(i), \phi_3(i), \phi_4(i), \phi_5(i), \phi_6(i))], \\
&\qquad \forall i = 2, 3, \ldots 8,
\end{aligned}
$$

where we define $\phi_1(i)$, $\phi_2(i)$, $\phi_3(i)$, $\phi_4(i)$, $\phi_5(i)$, $\phi_6(i)$ as the coefficients of x^2, xy, y^2, x, y and the constant term of the quadratic form in the exponential part of $Q(i)$ for $i = 2, 3, \cdots, 8$. Specifically, the parameters are given by $\phi_1(i) = \phi_1(1)$, $\phi_3(i) = \phi_3(1)$, $\phi_2(i) = (-1)^{i+1}\phi_2(1)$ $\forall i = 2, 3, \cdots, 8$, and $D(i)$, $\phi_4(i)$, $\phi_5(i)$, $\phi_6(i)$ are given by Table 65.2.

Table 65.2 Coefficients of $D(i)$, $\phi_4(i)$, $\phi_5(i)$, and $\phi_6(i)$

i	D(i)	$\phi_4(i)$	$\phi_5(i)$	$\phi_6(i)$
2	$\frac{K}{2\pi\sqrt{\eta_1 t_1}}$	$\frac{2b'}{\eta_1}+\theta$	$\theta+k_1\left(\frac{2b'}{\eta_1}-\theta\right)$	$-\frac{4b'^2+T^2\theta^2-T\theta^2 t_1}{2T-2t_1}$
3	$\frac{K}{2\pi\sqrt{\eta_1 t_1}}$	θ	$\frac{2b}{t_1}+\theta-\theta k_1$	$-\frac{2b^2}{t_1}-\frac{T\theta^2}{2}$
4	$-\frac{K}{2\pi\sqrt{\eta_1 t_1}}$	$\frac{2b'}{\eta_1}+\theta$	$\frac{2b}{t_1}+\theta+k_1\left(\frac{2b'}{\eta_1}-\theta\right)$	$-\frac{2b^2}{t_1}-\frac{T\theta^2}{2}-\frac{2b'^2}{\eta_1}$
5	$\frac{P'(0)}{2\pi\sqrt{\eta_1 t_1}}$	$\theta+\sigma$	$\theta-\theta k_1$	$-\frac{T\theta^2}{2}$
6	$-\frac{P'(0)}{2\pi\sqrt{\eta_1 t_1}}$	$\frac{2b'}{\eta_1}+\theta+\sigma$	$\theta+k_1\left(\frac{2b'}{\eta_1}-\theta\right)$	$-\frac{4b'^2+T^2\theta^2-T\theta^2 t_1}{2T-2t_1}$
7	$-\frac{P'(0)}{2\pi\sqrt{\eta_1 t_1}}$	$\theta+\sigma$	$\frac{2b}{t_1}+\theta-\theta k_1$	$-\frac{2b^2}{t_1}-\frac{T\theta^2}{2}$
8	$\frac{P'(0)}{2\pi\sqrt{\eta_1 t_1}}$	$\frac{2b'}{\eta_1}+\theta+\sigma$	$\frac{2b}{t_1}+\theta+k_1\left(\frac{2b'}{\eta_1}-\theta\right)$	$-\frac{2b^2}{t_1}-\frac{T\theta^2}{2}-\frac{2b'^2}{\eta_1}$

Thus, we obtain the pricing formula

$$\begin{aligned}\dot{C} &= e^{-rT}\int_{-\infty}^{b}\int_{k\prime}^{b\prime} Q(1)+Q(2)+\ldots+Q(8)dxdy \\ &= e^{-rT}\sum_{i=1}^{8}[D(i)[G(b',b,\phi_1(i),\phi_2(i),\phi_3(i),\phi_4(i),\phi_5(i),\phi_6(i)) \\ &\quad -G(k',b,\phi_1(i),\phi_2(i),\phi_3(i),\phi_4(i),\phi_5(i),\phi_6(i))]].\end{aligned}$$

65.3.2 Multi-Discrete-Dividend Case

The above approach can be repeatedly applied to derive approximated pricing formulas for barrier stock options with multiple-discrete-dividend payout. For simplicity, we derive the pricing formula for the two-dividend case in this section. The extensions for three or more dividends cases are straightforward. Note that $t_1 < t_2 < T < t_3$ in the two-dividend case.

To evaluate the option, we need to derive the joint density function of the maximum stock prices over the time intervals [0, t_1), [t_1, t_2), and [t_2, T] and the stock price at time T. Let $\hat{M}_1(t_2-t_1) \equiv max_{t_1\le t<t_2}\hat{W}_1(t-t_1)$ be the maximum value of $\hat{W}_1(t)$ over the time interval [t_1, t_2) and $\hat{M}_2\eta_2 \equiv max_{t_2\le t\le T}\hat{W}_2(t-t_2)$ be the maximum value of $\hat{W}_2(t)$ over the time interval [t_2, T], where η_2 denotes the abbreviation of $T-t_2$.

The joint density function of $\hat{M}_1(t_2-t_1)$ and $\hat{W}_1(t_2-t_1)$ and the density function of $\hat{M}_2\eta_2$ and $\hat{W}_2\eta_2$ can be derived by applying Theorem 65.1 as follows:

$$f_{\hat{M}_1}(t_2-t_1),\hat{W}_1(t_2-t_1)(m_1,w_1) = \begin{cases} \dfrac{2(2m_1-w_1)}{(t_2-t_1)\sqrt{2\pi(t_2-t_1)}}e^{\theta w_1-\frac{1}{2}\theta^2(t_2-t_1)-\frac{1}{2(t_2-t_1)}\left(2m_1-w_1\right)^2} & \text{if } m_1 \ge {w_1}^+, \\ 0 & \text{otherwise,} \end{cases} \tag{65.26}$$

$$f_{\hat{M}_2}\eta_2, \hat{W}_2\eta_2\,(m_2, w_2) = \begin{cases} \dfrac{2(2m_2 - w_2)}{\eta_2\sqrt{2\pi\eta_2}} e^{\theta w_2 - \frac{1}{2}\theta^2\eta_2 - \frac{1}{2\eta_2}\left(2m_2 - w_2\right)^2} & \text{if } m_2 \geq {w_2}^+, \\ 0 & \text{otherwise.} \end{cases} \tag{65.27}$$

For simplicity, we use $\ddot{f}_0$, $\ddot{f}_1$, and $\ddot{f}_2$ to represent the density functions $f_{\hat{M}}(t_1), \hat{B}(t_1)$ (see Eq. 65.17), $f_{\hat{M}_1}(t_2 - t_1), _{\hat{W}_1}(t_2 - t_1)$, and $f_{\hat{M}_2}\eta_2, \hat{W}_2\eta_2$, respectively. Note that the drifted Brownian motions $\hat{B}(t)$ for $t \in [0, t_1)$, $\hat{W}_1(t - t_1)$ for $t \in [t_1, t_2)$, and $\hat{W}_2(t - t_2)$ for $t \in [t_2, t_3]$ are independent due to the Markov property of Brownian motion; the joint density function of maximum stock prices over $[0, t_1)$, $[t_1, t_2)$, and $[t_2, T]$ and the stock prices at time t_1, t_2, and T can be calculated by directly multiplying $\ddot{f}_0$ with $\ddot{f}_1$ and $\ddot{f}_2$

The option value can be evaluated by the risk-neutral variation method as follows:

$$\ddot{C} \equiv e^{-rT} E\left[\left(\hat{P}(T) - K\right) 1_{\left\{\ddot{\Xi}_1 \cap \ddot{\Xi}_2 \cap \ddot{\Xi}_3 \cap \ddot{\Xi}_4\right\}}\right], \tag{65.28}$$

where $\ddot{\Xi}_1, \ddot{\Xi}_2, \ddot{\Xi}_3$ represent the events that the stock price process does not hit the barrier B during the time interval $[0, t_1)$, $[t_1, t_2)$, and $[t_2, T]$, respectively, and $\ddot{\Xi}_4$ denotes the event that the stock price at maturity is greater than the strike price. Specifically, $\ddot{\Xi}_1, \ddot{\Xi}_2, \ddot{\Xi}_3$, and $\ddot{\Xi}_4$ are defined as

$$\begin{aligned} &\ddot{\Xi}_1 \equiv \left\{\hat{M}(t_1) < b\right\}, \\ &\ddot{\Xi}_2 \equiv \left\{\hat{M}_1(t_2 - t_1) < b' - k_1\hat{B}(t_1)\right\}, \\ &\ddot{\Xi}_3 \equiv \left\{\hat{M}_{2\eta 2} < b'' - k_1k_2\hat{B}(t_1) - k_2\hat{W}(t_2 - t_1)\right\}, \\ &\ddot{\Xi}_4 \equiv \left\{\hat{W}_{2\eta 2} < k'' - k_1k_2\hat{B}(t_1) - k_2\hat{W}_1(t_2 - t_1)\right\}, \end{aligned}$$

where $k'' \equiv \frac{1}{\sigma}\log\frac{K}{S''(0)}$, and $b'' \equiv \frac{1}{\sigma}\log\frac{B}{S''(0)}$, respectively. Thus, we can compute the pricing formula in Eq. 65.28 by applying the law of iterated expectation as follows:

$$\begin{aligned} \ddot{C} &= e^{-rT} E\Big[E\Big[E\Big[\left(\hat{P}(T) - K\right) 1_{\left\{\ddot{\Xi}_1 \cap \ddot{\Xi}_2 \cap \ddot{\Xi}_3 \cap \ddot{\Xi}_4\right\}} \\ &\qquad \left|\hat{B}(t_1), \hat{M}(t_1), \hat{W}_1(t_2 - t_1), \hat{M}_1(t_2 - t_1)\right]\left|\hat{B}(t_1), \hat{M}(t_1)\right]\Big] \\ &= e^{-rT} \int_{-\infty}^{b} \int_{w^+}^{b} \int_{-\infty}^{b\prime - k_1 w} \int_{w_1^+}^{b\prime - k_1 w} \int_{k'' - k_1k_2w - k_2w_1}^{b'' - k_1k_2w - k_2w_1} \int_{w_2^+}^{b''| - k_1k_2w - k_2w_1} \\ &\qquad \left(S''(0) e^{k_1k_2\sigma w + k_2\sigma w_1 + \sigma w_2} - K\right) \\ &\qquad \cdot \ddot{f}_2(m_2, w_2) \cdot \ddot{f}_1(m_1, w_1) \cdot \ddot{f}_0(m, w) dm_2 dw_2 dm_1 dw_1 dm dw, \end{aligned} \tag{65.29}$$

where the domain of integral in Eq. 65.29 is obtained by taking the intersection of the supports of $\ddot{f}_2(m_2, w_2)$,$\ddot{f}_1(m_1, w_1)$,$\ddot{f}_0(m, w)$ Since only $\ddot{f}_2(m_2, w_2)$ contains m_2, $\ddot{f}_1(m_1, w_1)$ contains m_1, and $\ddot{f}_0(m, w)$ contains m in the integrand in Eq. 65.29, $\int \ddot{f}_0(m, w)dm$, $\int \ddot{f}_1(m_1, w_1)dm_1$ and $\int \ddot{f}_2(m_2, w_2)\, dm_2$ can be simplified as follows:

$$
\begin{aligned}
\ddot{C} = e^{-rT} & \int_{-\infty}^{b} \int_{-\infty}^{b\prime - k_1 w} \int_{k'' - k_1 k_2 w - k_2 w_1}^{b'' - k_1 k_2 w - k_2 w_I} \\
& \left(S''(0) e^{k_1 k_2 \sigma w + k_2 \sigma w_1 + \sigma w_2} - K \right) \\
& \cdot \left(\int_{w_2^+}^{b'' - k_1 k_2 w - k_2 w_1} \frac{2(2m_2 - w_2)}{\eta_2 \sqrt{2\pi \eta_2}} \right. \\
& \left. \times e^{\theta w_2 - \frac{1}{2}\theta^2 \eta_2 - \frac{1}{2\eta_2}(2\mathrm{m}_2 - w_2)^2} dm_2 \right) \\
& \cdot \left(\int_{w_1^+}^{b' - k_1 w} \frac{2(2m_1 - w_1)}{(t_2 - t_1)\sqrt{2\pi (t_2 - t_1)}} \right. \\
& \left. \times e^{\theta w_1 - \frac{1}{2}\theta^2 (t_2 - t_1) - \frac{1}{2(t_2 - t_1)}(2m_1 - w_1)^2} dm_1 \right) \\
& \cdot \left(\int_{w^+}^{b} \frac{2(2m - w)}{t_1 \sqrt{2\pi t_1}} e^{\theta w - \frac{1}{2}\theta^2 t_1 - \frac{1}{2t_1}(2m - w)^2} dm \right) dw_2 dw_1 dw.
\end{aligned}
$$

To eliminate the variables in the lower and the upper limits for the integrals on w_1 and w_2, we substitute

$$
\begin{cases}
x = w_2 + k_2 w_1 + k_1 k_2 w, \\
y = w_1 + k_1 w, \\
z = w
\end{cases},
$$

into the aforementioned formula to get[6]

$$
\ddot{C} = e^{-rT} \int_{-\infty}^{b} \int_{-\infty}^{b\prime} \int_{k''}^{b''} \sum_{i=1}^{16} O(i) dx dy dz, \tag{65.30}
$$

where $O(1)$, $O(2)$,⋯, $O(16)$ are defined in Table 65.3

Again, each of the integrands $O(1)$, $O(2)$,⋯, $O(16)$ is a quadratic form taking exponential. We reformulate the integrals of them in terms of the CDF of the multivariate normal distribution by the lemma shown below. Appendix gives the proof of this lemma.

[6]The Jacobian determinant $\frac{\partial(w_2, w_1, w)}{\partial(x, y, z)}$ is 1.

Table 65.3 The definitions of $O(1), \ldots, O(16)$

$$O(1) = -\frac{K}{\sqrt{8\pi^3\eta_2(t_2-t_1)t_1}}e^{-\frac{z^2}{2t_1}+\theta z+\theta(y-zk_1)+\theta(x-yk_2)-\frac{\theta^2 t_1}{2}-\frac{1}{2}\theta^2\eta_2-\frac{1}{2}\theta^2(t_2-t_1)-\frac{(x-yk_2)^2}{2\eta_2}-\frac{(y-zk_1)^2}{2(t_2-t_1)}},$$

$$O(2) = \frac{K}{\sqrt{8\pi^3\eta_2(t_2-t_1)t_1}}e^{-\frac{z^2}{2t_1}+\theta z+\theta(y-zk_1)+\theta(x-yk_2)-\frac{\theta^2 t_1}{2}-\frac{1}{2}\theta^2\eta_2-\frac{1}{2}\theta^2(t_2-t_1)+\frac{2b(z-b)}{t_1}-\frac{(x-yk_2)^2}{2\eta_2}-\frac{(y-zk_1)^2}{2(t_2-t_1)}},$$

$$O(3) = \frac{K}{\sqrt{8\pi^3\eta_2(t_2-t_1)t_1}}e^{-\frac{z^2}{2t_1}+\theta z+\theta(y-zk_1)+\theta(x-yk_2)-\frac{\theta^2 t_1}{2}-\frac{1}{2}\theta^2\eta_2-\frac{1}{2}\theta^2(t_2-t_1)-\frac{(x-yk_2)^2}{2\eta_2}+\frac{2\left(x-b''\right)\left(b''-yk_2\right)}{\eta_2}-\frac{(y-zk_1)^2}{2(t_2-t_1)}},$$

$$O(4) = -\frac{K}{\sqrt{8\pi^3\eta_2(t_2-t_1)t_1}}e^{-\frac{z^2}{2t_1}+\theta z+\theta(y-zk_1)+\theta(x-yk_2)-\frac{\theta^2 t_1}{2}-\frac{1}{2}\theta^2\eta_2-\frac{1}{2}\theta^2(t_2-t_1)+\frac{2b(z-b)}{t_1}-\frac{(x-yk_2)^2}{2\eta_2}+\frac{2\left(x-b''\right)\left(b''-yk_2\right)}{\eta_2}-\frac{(y-zk_1)^2}{2(t_2-t_1)}},$$

$$O(5) = \frac{K}{\sqrt{8\pi^3\eta_2(t_2-t_1)t_1}}e^{-\frac{z^2}{2t_1}+\theta z+\theta(y-zk_1)+\theta(x-yk_2)-\frac{\theta^2 t_1}{2}-\frac{1}{2}\theta^2\eta_2-\frac{1}{2}\theta^2(t_2-t_1)-\frac{(x-yk_2)^2}{2\eta_2}-\frac{(y-zk_1)^2}{2(t_2-t_1)}+\frac{2(y-b')(b'-zk_1)}{t_2-t_1}},$$

$$O(6) = -\frac{K}{\sqrt{8\pi^3\eta_2(t_2-t_1)t_1}}e^{-\frac{z^2}{2t_1}+\theta z+\theta(y-zk_1)+\theta(x-yk_2)-\frac{\theta^2 t_1}{2}-\frac{1}{2}\theta^2\eta_2-\frac{1}{2}\theta^2(t_2-t_1)+\frac{2b(z-b)}{t_1}-\frac{(x-yk_2)^2}{2\eta_2}-\frac{(y-zk_1)^2}{2(t_2-t_1)}+\frac{2(y-b')(b'-zk_1)}{t_2-t_1}},$$

$$O(7) = -\frac{K}{\sqrt{8\pi^3\eta_2(t_2-t_1)t_1}}e^{-\frac{z^2}{2t_1}+\theta z+\theta(y-zk_1)+\theta(x-yk_2)-\frac{\theta^2 t_1}{2}-\frac{1}{2}\theta^2\eta_2-\frac{1}{2}\theta^2(t_2-t_1)-\frac{(x-yk_2)^2}{2\eta_2}+\frac{2\left(x-b''\right)\left(b''-yk_2\right)}{\eta_2}-\frac{(y-zk_1)^2}{2(t_2-t_1)}+\frac{2(y-b')(b'-zk_1)}{t_2-t_1}},$$

$$O(8) = \frac{K}{\sqrt{8\pi^3\eta_2(t_2-t_1)t_1}}e^{-\frac{z^2}{2t_1}+\theta z+\theta(y-zk_1)+\theta(x-yk_2)-\frac{\theta^2 t_1}{2}-\frac{1}{2}\theta^2\eta_2-\frac{1}{2}\theta^2(t_2-t_1)+\frac{2b(z-b)}{t_1}-\frac{(x-yk_2)^2}{2\eta_2}+\frac{2\left(x-b''\right)\left(b''-yk_2\right)}{\eta_2}-\frac{(y-zk_1)^2}{2(t_2-t_1)}+\frac{2(y-b')(b'-zk_1)}{t_2-t_1}},$$

$$O(9) = \frac{S''(0)}{\sqrt{8\pi^3\eta_2(t_2-t_1)t_1}} e^{-\frac{z^2}{2t_1}+\theta z+x\sigma+\theta(y-zk_1)+\theta(x-yk_2)-\frac{\theta^2 t_1}{2}-\frac{1}{2}\theta^2\eta_2-\frac{1}{2}\theta^2(t_2-t_1)-\frac{(x-yk_2)^2}{2\eta_2}-\frac{(y-zk_1)^2}{2(t_2-t_1)}},$$

$$O(10) = -\frac{S''(0)}{\sqrt{8\pi^3\eta_2(t_2-t_1)t_1}} e^{-\frac{z^2}{2t_1}+\theta z+x\sigma+\theta(y-zk_1)+\theta(x-yk_2)-\frac{\theta^2 t_1}{2}-\frac{1}{2}\theta^2\eta_2-\frac{1}{2}\theta^2(t_2-t_1)+\frac{2b(z-b)}{t_1}-\frac{(x-yk_2)^2}{2\eta_2}-\frac{(y-zk_1)^2}{2(t_2-t_1)}},$$

$$O(11) = -\frac{S''(0)}{\sqrt{8\pi^3\eta_2(t_2-t_1)t_1}} e^{-\frac{z^2}{2t_1}+\theta z+x\sigma+\theta(y-zk_1)+\theta(x-yk_2)-\frac{\theta^2 t_1}{2}-\frac{1}{2}\theta^2\eta_2-\frac{1}{2}\theta^2(t_2-t_1)-\frac{(x-yk_2)^2}{2\eta_2}+\frac{2(x-b'')(b''-yk_2)}{\eta_2}-\frac{(y-zk_1)^2}{2(t_2-t_1)}},$$

$$O(12) = \frac{S''(0)}{\sqrt{8\pi^3\eta_2(t_2-t_1)t_1}} e^{-\frac{z^2}{2t_1}+\theta z+x\sigma+\theta(y-zk_1)+\theta(x-yk_2)-\frac{\theta^2 t_1}{2}-\frac{1}{2}\theta^2\eta_2-\frac{1}{2}\theta^2(t_2-t_1)+\frac{2b(z-b)}{t_1}-\frac{(x-yk_2)^2}{2\eta_2}+\frac{2(x-b'')(b''-yk_2)}{\eta_2}-\frac{(y-zk_1)^2}{2(t_2-t_1)}},$$

$$O(13) = -\frac{S''(0)}{\sqrt{8\pi^3\eta_2(t_2-t_1)t_1}} e^{-\frac{z^2}{2t_1}+\theta z+x\sigma+\theta(y-zk_1)+\theta(x-yk_2)-\frac{\theta^2 t_1}{2}-\frac{1}{2}\theta^2\eta_2-\frac{1}{2}\theta^2(t_2-t_1)-\frac{(x-yk_2)^2}{2\eta_2}-\frac{(y-zk_1)^2}{2(t_2-t_1)}+\frac{2(y-b')(b'-zk_1)}{t_2-t_1}},$$

$$O(14) = \frac{S''(0)}{\sqrt{8\pi^3\eta_2(t_2-t_1)t_1}} e^{-\frac{z^2}{2t_1}+\theta z+x\sigma+\theta(y-zk_1)+\theta(x-yk_2)-\frac{\theta^2 t_1}{2}-\frac{1}{2}\theta^2\eta_2-\frac{1}{2}\theta^2(t_2-t_1)+\frac{2b(z-b)}{t_1}-\frac{(x-yk_2)^2}{2\eta_2}-\frac{(y-zk_1)^2}{2(t_2-t_1)}+\frac{2(y-b')(b'-zk_1)}{t_2-t_1}},$$

$$O(15) = \frac{S''(0)}{\sqrt{8\pi^3\eta_2(t_2-t_1)t_1}} e^{-\frac{z^2}{2t_1}+\theta z+x\sigma+\theta(y-zk_1)+\theta(x-yk_2)-\frac{\theta^2 t_1}{2}-\frac{1}{2}\theta^2\eta_2-\frac{1}{2}\theta^2(t_2-t_1)-\frac{(x-yk_2)^2}{2\eta_2}+\frac{2(x-b'')(b''-yk_2)}{\eta_2}-\frac{(y-zk_1)^2}{2(t_2-t_1)}+\frac{2(y-b')(b'-zk_1)}{t_2-t_1}},$$

$$O(16) = -\frac{S''(0)}{\sqrt{8\pi^3\eta_2(t_2-t_1)t_1}} e^{-\frac{z^2}{2t_1}+\theta z+x\sigma+\theta(y-zk_1)+\theta(x-yk_2)-\frac{\theta^2 t_1}{2}-\frac{1}{2}\theta^2\eta_2-\frac{1}{2}\theta^2(t_2-t_1)+\frac{2b(z-b)}{t_1}-\frac{(x-yk_2)^2}{2\eta_2}+\frac{2(x-b'')(b''-yk_2)}{\eta_2}-\frac{(y-zk_1)^2}{2(t_2-t_1)}+\frac{2(y-b')(b'-zk_1)}{t_2-t_1}}.$$

Lemma 65.4 *Let $F_{X_1,X_2,X_3}(a,b,c,\Omega)$ be the CDF of a multivariate normal distributed random vector (X_1, X_2, X_3). That is,*

$$\begin{aligned}
&F_{X_1,X_2,X_3}(a,b,c,\Omega) \\
&\equiv \int_{-\infty}^{a}\int_{-\infty}^{b}\int_{-\infty}^{c} f_{X_1,X_2,X_3}(x_1,x_2,x_3)dx_1dx_2dx_3 \\
&= \int_{-\infty}^{a}\int_{-\infty}^{b}\int_{-\infty}^{c} \frac{1}{\sqrt{8\pi^3|\Omega|}} e^{-\frac{1}{2}x^T\Omega \mathrm{x}} dx_1dx_2dx_3
\end{aligned}$$

where $\mathbf{x}$ and Ω denote $(x_1, x_2, x_3)^T$ and the covariance matrix, respectively, and each of X_1, $X_{2,}$ and X_3 has mean 0 and variance 1. The following triple integral can be expressed in terms of F_{X_1,X_2,X_3} as

$$\begin{aligned}
&H(p,q,r,A,B,C) \equiv \int_{-\infty}^{r}\int_{-\infty}^{q}\int_{-\infty}^{p} \\
&e^{\phi_1x^2+\phi_2y^2+\phi_3z^2+\phi_4xy+\phi_5yz+\phi_6xz+\phi_7x+\phi_8y+\phi_9z+\phi_{10}}dxdydz \\
&= e^{C\prime}\sqrt{\frac{\pi^3}{|-A|}F_{\Upsilon_1,\,\Upsilon_2,\,\Upsilon_3}\left(\frac{\chi_1-\mathrm{m}_1}{\Phi_{1,1}},\frac{\chi_2-\mathrm{m}_2}{\Phi_{2,2}},\frac{\chi_3-\mathrm{m}_3}{\Phi_{3,3}},\Omega\right)},
\end{aligned}$$

where

$$\begin{aligned}
A &= \begin{pmatrix} \phi_1 & \frac{\phi_4}{2} & \frac{\phi_6}{2} \\ \frac{\phi_4}{2} & \phi_2 & \frac{\phi_5}{2} \\ \frac{\phi_6}{2} & \frac{\phi_5}{2} & \phi_3 \end{pmatrix}, B = \begin{pmatrix} \phi_7 \\ \phi_8 \\ \phi_9 \end{pmatrix}, C = \phi_{10}, \\
\Phi &= \begin{pmatrix} \Phi_{1,1} & 0 & 0 \\ 0 & \Phi_{2,2} & 0 \\ 0 & 0 & \Phi_{3,3} \end{pmatrix}, \\
\Phi_{j,j} &= \sqrt{\left((-2A)^{-1}\right)_{j,j}} \forall j = 1,2,3, \\
\mathrm{m} &= -\frac{1}{2}A^{-1}B, \\
C' &= C - \frac{1}{4}B^TA^{-1}B \\
\Omega &= (-2\mathrm{S}A\Psi)^{-1}.
\end{aligned}$$

Let $A(i)$, $B(i)$, $C(i)$ be

$$A(i) = \begin{pmatrix} \phi_1(i) & \frac{\phi_4(i)}{2} & \frac{\phi_6(i)}{2} \\ \frac{\phi_4(i)}{2} & \phi_2(i) & \frac{\phi_5(i)}{2} \\ \frac{\phi_6(i)}{2} & \frac{\phi_5(i)}{2} & \phi_3(i) \end{pmatrix},$$

$$B(i) = \begin{pmatrix} \phi_7(i) \\ \phi_8(i) \\ \phi_9(i) \end{pmatrix}, C(i) = \phi_{10}(i),$$

respectively. Thus, we have

$$\begin{aligned}&\int_{-\infty}^{b}\int_{-\infty}^{b\prime}\int_{k''}^{b''} O(i)dxdydz\\&= L(i)\left[H\left(b'', b', b, A(i), B(i), C(i)\right)\right.\\&\quad \left.-H\left(k'', b', b, A(i), B(i), C(i)\right)\right], \forall i = 2, 3, \cdots, 8,\end{aligned}$$

where

$$\begin{aligned}
\phi_1(i) &= -\frac{1}{2\eta_2},\\
\phi_2(i) &= -\left(\frac{k_2^2}{2\eta_2} + \frac{1}{2(t_2 - t_1)}\right),\\
\phi_3(i) &= -\left(\frac{k_1^2}{2(t_2 - t_1)} + \frac{1}{2t_1}\right),\\
\phi_6(i) &= 0,\\
\phi_4(i) &= \begin{cases} \dfrac{k_2}{\eta_2} & \text{if}\, i = 4l + 1\,\text{or}\;\; i = 4l + 2\\ -\dfrac{k_2}{\eta_2} & \text{otherwise}\end{cases}\\
\phi 5(i) &= \begin{cases} \dfrac{k_2}{t_2 - t_1} & \text{if}\;\; i = 8l + 1, 8l + 2, 8l + 3, \text{or}\, 8l + 4\\ -\dfrac{k_2}{t_2 - t_1} & \text{otherwise}\end{cases}, \forall i = 1, 2, \cdots, 16,\\
L(i) &= \frac{K}{\sqrt{8\pi^3\eta_2(t_2 - t_1)t_1}}, \quad i = 2, 3, 5, 8,\\
L(i) &= -L(2), \quad i = 1, 4, 6, 7\\
L(i) &= \frac{S''(0)}{\sqrt{8\pi^3\eta_2(t_2 - t_1)t_1}}, \quad i = 9, 12, 14, 15,\\
L(i) &= -L(9), \quad i = 10, 11, 13, 16,
\end{aligned}$$

and the rest parameters can be easily derived.

Finally, we obtain the following pricing formula in two-dividend case:

$$\begin{aligned}\ddot{C} &= e^{-rT}\int_{-\infty}^{b}\int_{-\infty}^{b\prime}\int_{k''}^{b''}\\&\quad O(1) + O(2) + \cdots + O(16)dxdydz\\&= e^{-rT}\sum_{i=1}^{16}\left[L(i)\left[H\left(b'', b', b, A(i),\;\; B(i),\;\; C(i)\right)\right.\right.\\&\quad \left.\left.- H(k'', b', b,\;\; A(i),\;\; B(i),\;\; C(i))\right]\right].\end{aligned}$$

Table 65.4 Comparing our pricing results and the exact values in single-dividend case

$P(0)$	44	46	48	50	52
RR	1.0063271	1.25184	1.45995	1.60138	1.65384
ours	1.0063272	1.25183	1.45994	1.60137	1.65384
diff	−1.44794E-07	−4.6E-06	1.61E-06	1.81E-06	9.21E-07

Table 65.5 Comparing the pricing results by different models in single-dividend case

$P(0)$	B	ours	M2	M1
48	1.3456	1.3427	1.3317	1.3641
52	1.5829	1.5796	1.5767	1.6401
56	1.4389	1.431	1.4395	1.5423
60	0.9164	0.9106	0.9266	1.07
64	0.1932	0.1868	0.1932	0.3697
MAE		0.0089	0.0141	0.1765
RSE		0.0054	0.0093	0.1109

65.4 Numerical Results

Option values in different initial underlying stock prices with $r = 0.03$, $\sigma = 0.2$, $K = 50$, $B = 65$, $T = 1$, $t_1 = 0.5$, and $c_1 = 0$. "Ours" and "RR" stand for the pricing results by Eq. 65.8 and our closed form approximation for single-dividend case, respectively. "diff" gives their difference.

To show that our approximating formula prices accurately, in this section we give the numerical results of our formula, compared with other pricing schemes.

First we show that our pricing formula can exactly price a barrier option in zero-dividend case. Equation 65.8 given in Reiner and Rubinstein (1991) is our special case when all discrete dividend c_i are zero. Table 65.4 gives the comparison of the pricing results by our formula in single-dividend case and Eq. 65.8, with settings $r = 0.03$, $\sigma = 0.2$, $K = 50$, $B = 65$, $T = 1$, $t_1 = 0.5$, and $c_1 = 0$. As shown in this table, Eq. 65.8 and our pricing formula give almost the same result. Notice that the approximations for CDF $N(x)$ and $F_{\Upsilon_1,\Upsilon_2}(a, b, \Omega)$ could cause insignificant errors.

Option values approximated by model 1 (denoted as "M1"), 2 (denoted as "M2"), and our formula in different initial underlying stock prices with $r = 0.03$, $\sigma = 0.2$, $K = 50$, $B = 65$, $T = 1$, $t_1 = 0.5$, and $c_1 = 1$. The columns "error" list the difference between their left column and the benchmark (denoted as "B"). MAE denotes the maximum absolute error, and RSE stands for the root-mean-square error.

As mentioned in Frishling (2002), only model 3 can reflect the real-world phenomenon. Though models 1 and 2 suggested by Roll (1977) and Heath and Jarrow (1988) try to approximate model 3, Table 65.5 shows that their pricing results are less accurate than ours. Table 65.5 gives the comparison of the pricing

Table 65.6 Comparing the pricing results by different models in single-dividend case. All settings are the same with table 65.5, but now we fix $P(0) = 50$ and compare the pricing results for different dividend c_1

c_1	B	ours	M2	M1
0.3	1.57589	1.57301	1.57046	1.58565
0.9	1.52022	1.51291	1.50619	1.54856
1.5	1.44776	1.44926	1.439	1.50439
2.1	1.3843	1.38283	1.36935	1.45376
2.7	1.30605	1.31449	1.29772	1.39734
MAE		0.01004	0.01495	0.09737
RMSE		0.00528	0.01067	0.06299

Table 65.7 Comparing the pricing results in two-dividend case

$P(0)$	B	ours	M2	M1
48	1.003312	1.002807	0.996404	1.061938
52	1.048434	1.043798	1.047928	1.151891
56	0.877129	0.873651	0.887206	1.031609
60	0.536375	0.531522	0.54854	0.728746
64	0.110358	0.107166	0.113098	0.319177
MAE		0.004919	0.015314	0.208819
RSE		0.003754	0.009409	0.146997

results by models 1 and 2 and our formula in single-dividend case, where we use the results by Monte Carlo simulation with 1,000,000 paths as a benchmark. Table 65.6 gives the same comparison as Table 65.5, while Table 65.6 compares the pricing results for several different c_1 and one fixed $P(0)$, but Table 65.5 compares the results for different $P(0)$ fixing c_1. As shown in Table 65.6, our pricing formula produces poorer results as c_1 increases, but on average, our pricing results are still more accurate than others. Furthermore, typically dividend c_1 is quite small, and hence, our pricing formula approximates option values accurately. In two-dividend case, the pricing results comparison is shown in Table 65.7. As shown in these tables, our formula also produces more accurate results the model 1 and 2. Thus, we can conclude that our formula is better in all cases.

65.5 Conclusions

There are several different ways to model the stock price process with discrete dividend. As suggested by Frishling (2002), only model 3 can reflect real-world phenomenon. However, it is hard to price a barrier option under model 3. Our chapter suggests a way to derive analytically approximating pricing formulas for a barrier option under model 3. Though the resulting analytical formulas involve multiple integral, as shown above, all of them can be reformulated in terms of the

CDF of a multivariate normal distribution. As a result, our pricing formula prices efficiently. Furthermore, numerical results show that our pricing formulas produce accurate result.

Option values in two-dividend case approximated by models 1 and 2 and our formula in different initial underlying stock prices with $r = 0.03$, $\sigma = 0.2$, $K = 50$, $B = 65$, $T = 1.5$, $t_1 = 0.5$, and $t_2 = t_1 + 0.5$.

Appendix: Solve the Integration of Exponential Functions by the CDF of Multivariate Normal Distribution

If the price of the underlying asset is assumed to follow the lognormal diffusion process, most option pricing formulas, including the pricing formulas in this chapter, can be expressed in terms of multiple integrations of an exponential function, where the exponent term is a quadratic function of integrators $z_1, z_2, \cdots$. The integration problem can be numerically solved by reexpressing the formulas in terms of CDF of multivariate normal distribution, which can be efficiently solved by accurate numerical approximation methods (see Hull 2003). These numerical methods are provided by mathematical softwares, like Matlab and Mathematica. Take the simplest case – the single integral, for example. Under the premise $\phi_2 < 0$, the integral $\int_{-\infty}^{l} e^{\phi 2x^2+\phi_1 x+\phi_0}\, dx$ can be rewritten as

$$\sqrt{\frac{\pi}{-\phi_2}} e^{-\frac{\phi_1^2-4\phi_0\phi_2}{4\phi_2}} N\left(\frac{l-m}{s}\right), \tag{65.31}$$

where

$$\mathrm{m} = -\frac{\phi_1}{2\phi_2}$$
$$\mathrm{s} = \frac{1}{\sqrt{-2\phi_2}},$$

and $N(\cdot)$ denotes the CDF of a univariate standard normal distribution. Note that the above identity is used to derive the pricing formula in no-dividend case.

However, the integration for multivariate case is not straightforward. To address this problem, we derive a general formula for the multivariate integration with n integrators: $z_1, z_2, \cdots, z_n$.

Some matrix and vector calculations are employed to simplify derivation. For simplicity, for any matrix $\beth$, we use $|\beth|$, $\beth^T$, and $\beth^{-1}$ to denote the determinant, the transpose, and the inverse matrix of $\beth$. $\beth_{i,j}$ stands for the element located at the i-th row and j-th column of $\beth$. For any vector v, we use v_i to denote the i-th element of v. We further assume that $\mathbf{z} = (z_1, z_2, \cdots, z_n)^T$ is a column vector with n variables, β is an $n \times 1$ constant vector, δ is a constant, and ϑ is an $n \times n$ symmetric invertible negative-definite constant matrix. Then the general integral formula is derived in the following theorem:

Theorem 65.5 *For any general quadratic formula* $\mathbf{z}^T\ \vartheta\mathbf{z} + \beta^T\mathbf{z} + \delta$, *the n-variate integral for* $e^{\mathrm{z}^T\vartheta\mathrm{z}+\beta^T\mathrm{z}+\delta}$

$$\int_{-\infty}^{\chi_n}\int_{-\infty}^{\chi_{n-1}}\cdots\int_{-\infty}^{\chi_1} e^{\mathrm{z}^T\vartheta\mathrm{z}+\beta^T\mathrm{z}+\delta}dz \tag{65.32}$$

can be expressed in terms of a CDF of an n-dimensional standard normal distribution

$$F_{\Upsilon_1,\Upsilon_2,\cdots,\Upsilon_n}(l_1,l_2,\cdots,l_n,\Omega)\equiv$$
$$\int_{-\infty}^{l_n}\int_{-\infty}^{l_{n-1}}\cdots\int_{-\infty}^{l_1}\frac{1}{\sqrt{|\Omega|(2\pi)^n}}e^{-\frac{1}{2}y^T\Omega^{-1}y}dy,$$

where Ω *denotes the covariance matrix of a n-variate standard normal random vector* $(\Upsilon_1, \Upsilon_2,\cdots, \Upsilon_n)$.

Proof To express the integral in Eq. 65.32 in terms of a CDF of a standard normal distribution, the exponent term $\mathbf{z}^T\ \vartheta\mathbf{z} + \beta^T\mathbf{z} + \delta$ should be expressed in terms of the exponent term of a standard normal distribution. That is,

$$\mathbf{z}^T\vartheta\mathbf{z}+\beta^T\mathbf{z}+\delta=-\frac{1}{2}\mathbf{y}^T\boldsymbol{\Omega}^{-1}\mathbf{y}+\delta' \tag{65.33}$$

for some constant δ'. This can be achieved by first deriving a proper constant vector **m** and a proper diagonal matrix **S** and then substituting

$$\mathrm{y}=\mathrm{S}^{-1}\mathrm{n} \tag{65.34}$$

into the left-hand side of Eq. 65.33, where **n** is the abbreviation of $\mathbf{z} - \mathbf{m}$. The following lemma derives a proper **m** by completing the square identity for $\mathbf{z}^T\ \vartheta\mathbf{z} + \beta^T\mathbf{z}$.

Lemma 65.6 *Under the premises that* ϑ *is a symmetric invertible* $n \times n$ *matrix, and* $\mathbf{z}$, β *are both* $n \times 1$ *vectors, we have*

$$\mathbf{z}^T\vartheta\mathbf{z}+\beta^T\mathbf{z}=\left(\mathbf{z}+\frac{1}{2}\vartheta^{-1}\beta\right)^T\vartheta\left(\mathbf{z}+\frac{1}{2}\vartheta^{-1}\beta\right)-\frac{1}{4}\beta^T\vartheta^{-1}\beta. \tag{65.35}$$

Proof By expanding the right-hand side of Eq. 65.35, we have

$$\begin{aligned}&\left(\mathbf{z}+\tfrac{1}{2}\vartheta^{-1}\beta\right)^T\vartheta\left(\mathbf{z}+\frac{1}{2}\vartheta^{-1}\beta\right)-\frac{1}{4}\beta^T\vartheta^{-1}\beta\\&=\mathbf{z}^T\vartheta\mathbf{z}+\beta^T\mathbf{z}=\text{the left-hand side of Eq.65.35.}\end{aligned} \tag{65.36}$$

With lemma 65.6, we obtain

$$\mathbf{z}^T\vartheta\mathbf{z}+\mathbf{z}^T\beta+\delta=\mathbf{n}^T\vartheta\mathbf{n}+\delta', \tag{65.37}$$

where

$$\mathbf{m} \equiv -\frac{1}{2}\vartheta^{-1}\beta, \tag{65.38}$$

$$\delta' \equiv \delta - \frac{1}{4}\beta^T\vartheta^{-1}\beta. \tag{65.39}$$

The diagonal matrix Φ can be derived by equating the right-hand sides of Eq. 65.33 and Eq. 65.37 to get

$$\mathbf{n}^T\vartheta\mathbf{n} + \delta' = -\frac{1}{2}\mathbf{y}^T\boldsymbol{\Omega}^{-1}\mathbf{y} + \delta'.$$

Subtracting δ' from both sides of above equation yields

$$\mathbf{n}^T\vartheta\mathbf{n} = -\frac{1}{2}\mathbf{y}^T\Omega^{-1}\mathbf{y} \tag{65.40}$$

$$-\frac{1}{2}\mathbf{n}^T(\Phi\Omega\Phi)^{-1}\mathrm{n}, \tag{65.41}$$

By comparing the right-hand side of Eqs. 65.40 and 65.41, we have $-\frac{1}{2}(\Phi\Omega\Phi)^{-1} = \vartheta$, which is rewritten as $\Phi\Omega\Phi = (-2\vartheta)^{-1}$. Recall that Φ is a diagonal matrix. All diagonal elements of Ω are all 1 since Ω is a covariance matrix of multivariate standard normal random variables. Thus, we have $(\Phi\Omega\Phi)i$, $i = \Phi_{i,i}^2$, which leads us to obtain

$$\Phi_{i,j} \equiv \begin{cases} \sqrt{\left((-2\vartheta)^{-1}\right)_{i,i}} & \text{if } i{=}j \\ 0 & \text{otherwise} \end{cases} \tag{65.42}$$

and

$$\Omega \equiv (-2\Phi\vartheta\Phi)^{-1}. \tag{65.43}$$

Now we can evaluate Eq. 65.32 with δ', $\mathbf{m}$, Φ, and Ω defined above. By applying the change of variable Eq. 65.34, Eq. 65.32 can be rewritten as

$$\begin{aligned} &\int_{z_n=-\infty}^{z_n=\chi_n}\int_{z_{n-1}=-\infty}^{z_{n-1}=\chi_{n-1}}\cdots\int_{z_1=-\infty}^{z_1=\chi_1} e^{\mathbf{z}^T\vartheta\mathbf{z}+\beta^T\mathbf{z}+\delta}\mathbf{dz} \\ &= \int_{z_n=-\infty}^{z_n=\chi_n}\int_{z_{n-1}=-\infty}^{z_{n-1}=\chi_{n-1}}\cdots\int_{z_1=-\infty}^{z_1=\chi_1} e^{-\frac{1}{2}\mathbf{y}^T\boldsymbol{\Omega}^{-1}\mathbf{y}+\delta\prime} \\ &\left|\left|\frac{\partial\mathbf{x}}{\partial\mathbf{y}}\right|\right|dy. \end{aligned} \tag{65.44}$$

Since the elements in vector $\mathbf{y}$ can be represented as $\left(\frac{z_1-\mathbf{m}_1}{\Phi_{1,1}}, \frac{z_2-\mathbf{m}_2}{\Phi_{2,2}}, \cdots, \frac{z_n-\mathbf{m}_n}{\Phi_{n,n}}\right)^T$, the Jacobian determinant can be straightforwardly computed to get $\left|\frac{\partial \mathrm{x}}{\partial \mathrm{y}}\right| = \prod_{i=1}^{n} \Phi_{i,i} = |\Phi|$. Thus, Eq. 65.44 can be further rewritten as the following closed form formula:

$$e^{\delta\prime}\sqrt{\frac{\pi^n}{|-\vartheta|}}F_{\Upsilon_1,\Upsilon_2,\cdots,\Upsilon_n}\left(\frac{\chi_1-\mathbf{m}_1}{\Phi_{1,1}}, \frac{\chi_2-\mathbf{m}_2}{\Phi_{2,2}}, \cdots, \frac{\chi_n-\mathbf{m}_n}{\Phi_{n,n}}, \Omega\right), \tag{65.45}$$

where $|\Phi|\sqrt{|\Omega|} = \sqrt{|\Phi\Omega\Phi|} = \sqrt{|-2\vartheta|^{-1}}$ and $\mathrm{I} - 2\vartheta| = 2^n\mathrm{I} - \vartheta|$ are substituted into Eq. 65.45. This integration formula is similar to the single integral case given. Q.E.D.

References

Bender, R. & Vorst T. (2001). Options on dividends paying stocks. In *Proceeding of the 2001 International Conference on Mathematical Finance*, Shanghai.

Black, F. (1975). Fact and fantasy in the use of options. *Financial Analysts Journal, 31*(36–41), 61–72.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy, 81*, 637.

Bos, M., & Vandermark, S. (2002). Finessing fixed dividends. *Risk, 15*, 157–158.

Chiras, D., & Manaster, S. (1978). The informational content of option prices and a test of market efficiency. *Journal of Financial Economics, 6*, 213–234.

Dai, T.-S. (2009). Efficient option pricing on stocks paying discrete or path-dependent dividends with the stair tree. *Quantitative Finance, 9*, 827–838.

Dai, T.-S., & Lyuu, Y. D. (2009). Accurate approximation formulas for stock options with discrete dividends. *Applied Economics Letters, 16*, 1657–1663.

Figlewski, S., & Gao, B. (1999). The adaptive mesh model: A new approach to efficient option pricing. *Journal of Financial Economics, 53*, 313–351.

Frishling, V. (2002). A discrete question. *Risk, 15*, 115–116.

Gaudenzi, M., & Zanette, A. (2009). Pricing American barrier options with discrete dividends by binomial trees. *Decisions in Economics and Finance, 32*, 129–148.

Geske, R. (1979). A note on an analytical valuation formula for unprotected American call options on stocks with known dividends. *Journal of Financial Economics, 7*(4), 375–380.

Heath, D., & Jarrow, R. (1988). Ex-dividend stock price behavior and arbitrage opportunities. *Journal of Business, 61*, 95–108.

Hull, J. (2003). Options, futures, and other derivatives. Prentice Hall.

Kim, I., Ramaswamy, K., & Sundaresan, S. (1993). Does default risk in coupons affect the valuation of corporate bonds? A contingent claims model. *Financial Management, 22*, 117–131.

Lando, D. (2004). Credit risk modeling: Theory and applications. Princeton, NJ: Princeton University Press.

Leland, H. E. (1994). Corporate debt value, bond covenants, and optimal capital structure. *Journal of Finance, 49*, 157–196.

Leland, H. E., & Toft, K. (1996). Optimal capital structure, endogenous bankruptcy, and the term structure of credit spreads. *Journal of Finance, 51*, 987–1019.

Merton, R. (1973). Theory of rational option pricing. *Journal of Economics and Management Science, 4*, 141–183.

Merton, R. (1974). On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance, 29*, 449–470.
Musiela, M., & Rutkowski, M. (1997). *Martingale methods in financial modeling*. Sydney: Springer.
Reiner, E., & Rubinstein, M. (1991). Breaking down the barriers. *Risk, 4*, 28–35.
Roll, R. (1977). An analytic valuation formula for unprotected American call options on stocks with known dividends. *Journal of Financial Economics, 5*, 251–258.
Shreve, E. (2007). Stochastic calculus for finance II: Continuous-time models. New York: Springer Finance.
Vellekoop, M., & Nieuwenhuis, J. (2006). Efficient pricing of derivatives on assets with discrete dividends. *Applied Mathematical Finance, 13*, 265–284.

Pension Funds: Financial Econometrics on the Herding Phenomenon in Spain and the United Kingdom

66

Mercedes Alda García and Luis Ferruz

Contents

Abstract

This work reflects the impact of the Spanish and UK pension funds investment on the market efficiency; specifically, we analyze if manager's behavior enhances the existence of herding phenomena.

To implement this study, we apply a less common methodology: the estimated cross-sectional standard deviations of betas. We also estimate the betas

M. Alda García (✉) • L. Ferruz
Facultad de Economía y Empresa, Departamento de Contabilidad y Finanzas, Universidad de Zaragoza, Zaragoza, Spain
e-mail: malda@unizar.es; lferruz@unizar.es

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_66,

with an econometric technique less applied in the financial literature: state-space models and the Kalman filter. Additionally, in order to obtain a robust estimation, we apply the Huber estimator.

Finally, we apply several models and study the existence of herding towards the market, size, book-to-market, and momentum factors.

The results are similar for the two countries and style factors, revealing the existence of herding. Nonetheless, this is smaller on size, book-to-market, and momentum factors.

Keywords

Herding • Pension funds • State-space models • Kalman filter • Huber estimation • Imitation • Behavioral finance • Estimated cross-sectional standard deviations of betas • Herding towards the market • Herding towards size factor • Herding towards book-to-market factor • Herding towards momentum factor

66.1 Introduction

For some time, Western countries have been undergoing a range of demographic and social changes: increased life expectancy, ageing population, shorter active phases, and longer retirement periods. All of these changes, together with increasing concern over the viability of public pension systems, have led to a greater role for complementary pension systems.

Together with all these factors, society has become increasingly aware of the need to save for a higher retirement income; as Sanz (1999) observes, the state pension incomes are lower than those received during working life. Thus, together with tax incentives, pension plans are one of the key financial products for saving in developed economies.

The need to find a backup to the state pension has caused more and more professional and nonprofessional investors to take out pension plans by investing in pension funds.

As a consequence, these products are beginning to take a leading role in the industry of collective investment; according to INVERCO (Spanish Association of Investment and Pension Funds), the investment in such products worldwide in 2010 exceeded 12½ billion euros, of which more than 3 billion euros came from Europe, highlighting the United Kingdom, which represents a third of the total European investment.

The origin of pension funds and pension plans can be traced to the foundation of the welfare states, when various governments began to expand public spending, in particular on social welfare: education, health, pensions, housing, and unemployment benefit.

Among the different parts of the welfare state, pensions are included under the heading of the Social Welfare Systems, and the European pension systems are organized around the *Livonia* focus which is divided into three pillars, as González and García (2002) notes:

1. The first pillar consists of the Social Security System. It is integrated by the public system, which is mandatory, defined benefit pensions, and pay-as-you-go system, which guarantees a minimum level of pension. The pay-as-you-go system means that active workers pay, with their contributions, the pensions of retirees at that time; although each worker's pension depends on the contributions made during working life.
2. The second pillar consists of private and complementary occupational plans, within the scope of companies and workers associations. This pillar may be either voluntary or compulsory and may replace or supplement the first pillar, so it may be either private or public.
3. The third pillar consists of individual savings decisions, so it is private and voluntary; an example is the personal pension plans.

Therefore, the first pillar is the Public Social Welfare System, while the second and third comprise the Complementary Social Welfare System.

The evolution of each of these pillars depends on the evolution of the state system, so if the latter has expanded, providing more generous state pensions, the other two have been developed to a lesser extent, and vice versa. For instance, Fernández (2011) notes that the third pillar is weak in Spain because public pensions are close to the final salary.

Therefore, complementary pension systems have varying characteristics across Europe depending on the private pension's development (second and third pillars), and we can divide them into three groups:

– Countries in which private pension systems have reached a high stage of development: the United Kingdom, Ireland, the Netherlands, and Sweden.
– Countries in which private pension systems have reached a considerable degree of development, but are still evolving: Spain, Portugal, Italy, and Germany.
– Countries in which private pension systems are less developed: France, Belgium, and some East European countries.

With regard to legislation in Europe on these products, it must be noted that each country has its own, so that across Europe we can observe significant differences.

Among the different European countries, we focus on Spain and the United Kingdom because they represent two different systems and their industries have different size; as a result, we want to study if these characteristics provide different results on the topic studied: the herding phenomenon.

Respect the pension systems, these two markets are the most representative of two different pension systems:

Spain belongs to the "Mediterranean model," which is characterized by generous public pensions. In contrast, the United Kingdom belongs to the "Anglo-Saxon model," with less generous public pensions, so private pensions are more developed.

Equally, the differences in size between the two markets allow us to observe whether they enhance or not the existence of herding.

With regard to the United Kingdom, according to the OECD (2010), it is the leading country in Europe (1.1 billion euros) and the second in worldwide terms, just behind the United States.

For its part, Spain occupies seventh place in Europe with 84 milliard of euros invested in 2010, and given that these products did not come on the market here until 1988, this country has experienced a significant growth over the past years; as a result, this trend could have implications for the existence of herding.

We focus on the study of herding because traditionally, most of the financial studies examine the pension funds' performance, analyzing the manager behavior; however, the market efficiency can also be affected by the pension funds' investment, for example, through the herding phenomenon, an aspect that we study in this paper.

We apply a less common methodology to detect herding, based on the state-space models, as in the work of Christie and Huang (1995). Nonetheless, this model suffers from several inadequacies: it supposes that the instability of the market means that the whole market should demonstrate negative or positive returns and it introduces dummy variables arbitrarily; therefore, we also focus on the work of Hwang and Salmon (2004), who calculate herding using a single-factor model. In particular, they base on the market return from the CAPM model, using the betas dispersion of all market stocks.

The rest of the work is organized as follows: in the second section we describe the herding phenomenon and the third section carries out a literature review on the topic. The fourth section develops the methodology. The fifth section compiles the data and presents the empirical results. Finally, we show the main conclusions.

66.2 The Herding Phenomenon

This topic is part of the behavioral finance, which is focused on the study of the rationality of investment decisions and the implications of the cognitive processes in the make decisions (Fromlet 2001).

Investors' preference, such as the avoidance of loss, may produce some irrational reactions and affect the market efficiency (Kahnemann and Tversky 1979; Tversky and Kahnemann 1986). This behavior may imply price fluctuations, not necessarily related to the arrival of new market information, but rather by the emergence of collective phenomena, like herding (Thaler 1991; Shefrin 2000), affecting the efficiency and stability of the market.

In financial literature, herding arises when investors decide to imitate the decisions of other participants in the market or market movements; that it is to say, they imitate market agents who are thought to be better informed, rather than follow their own beliefs and information.

We assume the manager as investor, given that even though individual investors make the investments in pension plans, the managers are responsible for buying and selling in the market; likewise, they also vary the composition of the pension fund portfolio. Therefore, managers are the final investors and they carry out the investment in the market, albeit according to the guidelines established by the members of the pension plans.

In financial language, the herding phenomenon is one of the most widely discussed because, in the field of the asset pricing, it helps to explain market anomalies; however, the difficulty in its measurement and calculation has limited their research.

It is generally accepted that herding can lead to a situation in which market prices cannot reflect all of the information, then the market becomes unstable and moves towards the inefficiency. For that reason, market regulators show an interest in reducing this type of phenomena.

Theoretical and empirical studies have focused on finding causes and implications related to herding. The majority agree that this may due to both rational and irrational investor behavior.

According to Devenow and Welch (1996), irrational viewpoint focuses on the psychology of the investor, where the investor follows others blindly.

On the other hand, rational herding may appear due to diverse causes:

The first of them is the existence of imperfect information, that it is to say, when it is believed that other market participants are better informed. Banerjee (1992), Bikhchandani et al. (1992), Hirshleifer et al. (1994), Calvo and Mendoza (2000), Avery and Zemsky (1998), Chari and Kehoe (2004), Gompers and Metrick (2001), Puckett and Yan (2007), and Sahut et al. (2011) show evidence of such imperfect information.

The second aspect that causes rational herding is the costs of reputation. Scharfstein and Stein (1990), Trueman (1994), Rajan (1994), or Maug and Naik (1996), focusing on agency theory, show evidence of this. These studies prove that mutual fund managers imitate others in order to obtain bonuses as set out in their compensation-reputation scheme rewarding.

Compensation schemes also cause rational herding, as an investor will be rewarded based on their performance against the others; therefore, the deviations with respect to the market consensus could lead to an undesirable cost. Studies such as those of Roll (1992), Brennan (1993), Rajan (1994), or Maug and Naik (1996) demonstrate this.

In addition to these explanations, some authors have considered other factors, like the degree of institutional participation, the spread of opinions, derivatives markets and their sophistication, or uninformed investors. Among these are Patterson and Sharma (2006), Demirer and Kutan (2006), Henker et al. (2006), and Puckett and Yan (2007).

Despite the uncertainty surrounding the causes of this behavior, the study of herding in financial markets has followed two lines of investigation. The first one analyzes the tendency of individuals (individual investors). Among them, we highlight the works of Lakonishok et al. (1992a); Grinblatt et al. (1995); Wermers (1999), and Uchida and Nakagawa (2007).

The second trend focuses on market herding as a whole, that is to say, as a collective behavior of all participants buying or selling a certain asset at the same time. The most representative studies within this line of investigation are those of Christie and Huang (1995), Chang et al. (2000), Hwang and Salmon (2004), Patterson and Sharma (2006), and Wang (2008).

In this paper we focus on this second approach; to this purpose, we use the observing deviations from the equilibrium expressed in CAPM prices, a focus less used in financial literature.

Based on the works of Christie and Huang (1995) and Hwang and Salmon (2004), we capture herding with the use of observed returns data, instead of measuring it in the same way as Lakonishok et al. (1992a), by detailed records of individual trading activities which may not be available in many cases. For this reason, in order to detect herding, we use the cross-sectional dispersion of the betas.

Nonetheless, as this model suffers from some deficiencies as it supposes that market betas are statics, we build a time-varying distribution of the cross-sectional dispersion of betas. Likewise, as its betas do not take into account outliers, we apply a robust estimation.

The estimation method applied is the state-space model, using the Kalman filter. This methodology is more innovative in financial literature, and as far as we know, it has not been applied to pensions, so we will obtain new empirical evidence.

Likewise, as this technique has not been applied to pension funds, it will allow us to detect, for the first time, if pension fund managers produce herding behavior in the markets.

66.3 Literature Review

As we mentioned, we focus on the study of herding considering the market as a whole. In this trend, we find various types of models that detect herding: models of returns' dispersion and state-space models.

In the first ones we can distinguish between linear and nonlinear models. In the linear models, the most common measurement is the cross-sectional standard deviation of returns (henceforth CSSD), while in the nonlinear ones is the cross-sectional absolute deviation of returns (henceforth CSAD).

Studies based on these models show mixed evidence of this behavior. Christie and Huang (1995) find evidence in American stocks, but not during the market crises. Chang et al. (2000) find this in Taiwan, South Korea, and Japan, but not in Hong Kong and the United States. Lin and Swanson (2003) do not find this in international securities. Gleason et al. (2004) show this for ETFs funds. Bowe and Domuta (2004) also find positive results in the Jakarta stock market. Weiner (2006) finds scarce evidence for herding in the oil market. Demirer and Kutan (2006), together with Tan et al. (2008), study the Chinese market, but they do not find evidence of herding. In the Polish market, Goodfellow et al. (2009) find evidence of individual herding in bear markets, while they do not find evidence for institutional herding. Bohl et al. (2011) observe that restrictions on short-term positions lead to adverse herding in the United States, the United Kingdom, Germany, France, Australia, and South Korea. Economou et al. (2011) detect the presence of herding in Greek, Italian, and Portuguese markets.

With respect to the Spanish market, there are different studies that detect herding: Blasco and Ferreruela (2007, 2008) Lillo et al. (2008), and Blasco et al. (2011).

On the other hand, the state-space model (used in this work) also provides mixed evidence: Gleason et al. (2003) use it on the European futures markets, but the results display absence of herding. Hwang and Salmon (2004) find it in the Korean and American markets. In addition, they clearly observe it when the markets are stable and the investors are sure of the futures markets' direction. The authors conclude that financial crises stimulate a return towards efficiency.

Wang (2008) applies this to various markets (developed and emerging) and concludes that herding in emerging markets is greater than in developed ones.

Demirer et al. (2010) apply some models to the Taiwanese market, obtaining different results depending on the method. These authors find a lack of herding with the linear method (CSSD), but with the nonlinear model (CSAD) and the state-space model, they find strong evidence.

Shapour et al. (2010) also discover such evidence in the Tehran stock market, but no evidence of herding when they study towards size and book-to-market factors.

With regard to studies of herding in collective investment instruments (mutual and pension funds), we observe different analyses:

Oehler and Goeth-Chi (2000) examine German mutual funds that invest on bond market. The results show herding, but to a lesser degree than in the stock market. Kim and Wei (2002) also get positive evidence in domestic and international Korean mutual funds. Voronkova and Bohl (2005) do not detect an influence of Polish pension funds on the stock market. Wylie (2005) examines this behavior in the portfolio holdings of UK equity mutual fund managers, revealing some modest results. Walter and Weber (2006) analyze whether the German mutual fund managers demonstrate this behavior, and their results confirm it. Lobao et al. (2007) obtain evidence in Portuguese mutual funds, detecting a stronger tendency to herd among medium-cap funds, and a decrease when the stock market corrects itself or is more volatile.

Ferruz et al. (2008a, b) notice evidence of this phenomenon in value, growth, and cash stocks in Spanish equity mutual funds. Hsieh et al. (2010) show that mutual funds on 13 emerging Asian countries influence on the existence of herding, and this phenomenon is more pronounced during and after crises; for this reason, they suggest that mutual funds' behavior may have contributed to the crises. Fong et al. (2011) document the existence of herding due to information cascades, in a sample of US equity mutual funds. Jame (2011) examines the magnitude and effects of herding on pension fund price in the United States. Jame uses the measurement proposed by Lakonishok et al. (1992b) and confirms that pension funds are involved on the development of herding.

Given the small number of studies that analyze this topic in pension funds, this paper contributes to the financial literature by studying the influence of Spanish and British pension fund managers in their respective markets.

66.4 Methodology

66.4.1 CAPM and Herding

Herding leads to mispricing so rational decisions may be disturbed through the use of biased beliefs and views of expected returns and risks. More specifically, in the CAPM model, herding produces biased betas, and they deviate from equilibrium.

In order to observe empirically how this phenomenon affects the betas, we take as a starting point the CAPM model in equilibrium:

$$E_t(r_{it}) = \beta_{imt} E_t(r_{mt}) \tag{66.1}$$

where r_{it} and r_{mt} are the excess return of fund i and the excess market return over the risk-free asset during the period t, respectively, β_{imt} is the systematic risk measure, and $E_t(\cdot)$ is the conditional expectation at time t in order to price the fund i.

The CAPM model assumes that β_{imt} does not change over time, despite considerable empirical evidence demonstrating that betas are not constant, among those, Harvey (1989), Ferson and Harvey (1991, 1993), and Ferson and Korajczyk (1995).

Nonetheless, following Hwang and Salmon (2004), the evidence shows that the betas do not change over time in equilibrium, which means that the variation of betas can be interpreted as behavioral anomalies, such as herding, rather than from fundamental changes in the beta or the equilibrium relationship between $E_t(r_{mt})$ and $E_t(r_{it})$.

In this way, the individual cross-sectional dispersion of betas is lower than in equilibrium; given that if all returns were expected to be equal to the market return, all betas would be equal to one and the cross-sectional variance would be zero.

In addition, if we assume that $E_t(r_{mt})$ represents the market as a whole and investor first forms a view of the market as a whole and then considers the value of the individual asset, subsequently the investor's behavior is conditional on $E_t(r_{mt})$ and the observed beta (β_{imt}) will be biased, at least in the short-term, given $E_t(r_{mt})$.

In this way, the biased betas appear because the beliefs of the investors change, they follow the market more than they should in equilibrium, and they ignore the equilibrium relation, trying to match the individual asset returns with the market return. In this case takes place the so-called herding towards the market.

The opposite behavior is also possible, producing adverse herding. This appears when high betas (larger than one) become higher and low betas (smaller than one) become lower. On this occasion, individual return becomes more sensitive for large beta stocks and less sensitive for low beta stocks. This leads to a reversion in the long-term equilibrium of β_{imt}. In fact, adverse herding should exist if herding exists, since there must be some systematic adjustment back to the CAPM equilibrium.

66.4.2 Measuring Herding

When there is herding in the market portfolio, the CAPM equilibrium does not occur, and both beta and expected return are biased. Therefore, instead of the equilibrium in (1), Hwang and Salmon (2004) assume that the following relation is produced in presence of herding towards the market:

$$\frac{E_t^b(r_{it})}{E_t(r_{mt})} = \beta_{imt}^b = \beta_{imt} - h_{mt}(\beta_{imt} - 1) \tag{66.2}$$

where $E_t^b(r_{it})$ is the market's biased short run conditional expectation on the excess returns of fund i, β_{imt}^b is the market beta at time t in the presence of herding, and h_{mt} is a latent herding parameter that changes over time, less than one ($h_{mt} \leq 1$) and conditional on market fundaments.

When $h_{mt} = 0$, then $\beta_{imt}^b = \beta_{imt}$, and herding does not exist, producing the CAPM equilibrium. However, when $h_{mt} = 1$, then $\beta_{imt}^b = 1$, it is perfect herding towards the market portfolio in the sense that all individual funds move in the same direction and magnitude as the market portfolio.

In general, when the herding parameter (h_{mt}) is between zero and one ($0 < h_{mt} < 1$), a certain degree of herding exists in the market, determined by the magnitude of the herding coefficient.

Considering the situation described in the previous section, the relationship between the real and biased expected excess fund returns and its beta can be explained.

Therefore, for a fund with $\beta_{imt} > 1$, then $E_r(r_{it}) > E_t(r_{mt})$, and the fund presents herding towards the market, so that $E_t^b(r_{it})$ moves towards $E_t(r_{mt})$, and $E_r(r_{it}) > E_t^b(r_{it}) > E_t(r_{mt})$. As a result, the fund seems less risky than it should be, suggesting that $\beta_{imt}^b < \beta_{imt}$.

On the other hand, for a fund with $\beta_{imt} < 1$, it gives $E_r(r_{it}) < E_t(r_{mt})$, and the fund presents herding towards the market when $E_t^b(r_{it})$ moves towards $E_t(r_{mt})$, which is why $E_r(r_{it}) < E_t^b(r_{it}) < E_t(r_{mt})$. The fund seems riskier than it should be, suggesting that $\beta_{imt}^b > \beta_{imt}$.

Finally, for a fund with a beta equal to one, $\beta_{imt} = 1$, the fund is neutral to herding.

As we have already mentioned, the existence of herding implies adverse herding, allowing $h_{mt} < 0$; therefore, for a fund with $\beta_{imt} > 1$, then $E_t^b(r_{it}) > E_r(r_{it}) > E_t(r_{mt})$, while a fund with $\beta_{imt} < 1$ will produce the following: $E_t^b(r_{it}) < E_r(r_{it}) < E_t(r_{mt})$.

66.4.3 Models for Measuring Herding

Herding of a market portfolio can be captured with the parameter h_{mt} of the expression (66.2); however, neither the beta nor the herding parameter is observed.

For this reason, we use state-space models in order to extract those parameters. As we aim to measure herding in terms of the market as a whole, we assume that the Eq. 66.2 captures all of the market assets, so we can calculate herding using all of the assets and not only one, eliminating the effects of the idiosyncratic movements of individual betas (β_{imt}^b).

Since the cross-sectional mean of the betas (β_{imt}^b or β_{imt}) is always one, Hwang and Salmon (2004) show that

$$
\begin{aligned}
Std_c\left(\beta_{imt}^b\right) &= \sqrt{E_c\left((\beta_{imt} - h_{mt}(\beta_{imt} - 1) - 1)^2\right)} \\
&= \sqrt{E_c\left((\beta_{imt} - 1)^2\right)}(1 - h_{mt}) = Std_c(\beta_{imt})(1 - h_{mt}) \qquad (66.3)
\end{aligned}
$$

where $E_c(\cdot)$ represents the cross-sectional expectation, $Std_c(\beta^b_{imt})$ is the cross-sectional standard deviation of the beta in equilibrium, and $\sqrt{E_c\left((\beta_{imt} - h_{mt}(\beta_{imt} - 1) - 1)^2\right)}$ is a direct function of the herding parameter.

In order to minimize the impact of the idiosyncratic changes in β_{imt}, a great number of assets are used in the calculation of $Std_c(\beta_{imt})$, so $Std_c(\beta_{imt})$ will be stochastic in order to observe the movements in the equilibrium beta.

Nonetheless, as it is expected that the market as a whole $Std_c(\beta_{imt})$ does not change significantly in the short term, unless the structure of companies changes suddenly, it is assumed that $Std_c(\beta_{imt})$ does not exhibit any systematic movement and that the changes in $Std_c(\beta^b_{imt})$ in the short term are due to the changes in h_{mt}, that is to say, due to the presence of herding.

66.4.4 The State-Space Models

In the previous section, we remark that the herding parameter is not observed, so we apply state-space models. Those models may be estimated by using the Kalman filter, which is an algorithm to perform filtering on the state-space model.

In order to extract herding from model (66.3), we follow the procedure used by Hwang and Salmon (2004). First, taking logarithms of (66.3)

$$\log\left[Std_c\left(\beta^b_{imt}\right)\right] = \log[Std_c(\beta_{imt})] + \log(1 - h_{mt}) \tag{66.4}$$

and considering the assumptions carried out for $Std_c(\beta_{imt})$, it is rewritten like this:

$$\log\left[Std_c\left(\beta^b_{imt}\right)\right] = \mu_m + \upsilon_{mt} \tag{66.5}$$

where $\mu_m = E[\log[Std_c(\beta_{imt})]]$ and $\upsilon_{mt} \sim iid(0, \sigma^2_{m\upsilon})$, then

$$\log\left[Std_c\left(\beta^b_{imt}\right)\right] = \mu_m + H_{mt} + \upsilon_{mt} \tag{66.6}$$

where $H_{mt} = \log(1 - h_{mt})$.

Therefore, we suppose that the herding (H_{mt}) evolves over time and follows a dynamic process; for example, assuming a mean zero AR(1), we obtain the model (66.7):

$$\log\left[Std_c\left(\beta^b_{imt}\right)\right] = \mu_m + H_{mt} + \upsilon_{mt} \tag{66.7}$$

$$H_{mt} = \phi_m H_{mt-1} + \eta_{mt}$$

where $\eta_{mt} \sim iid(0, \sigma^2_{m\eta})$.

As a result, we have a standard state-space model, similar to those used in stochastic volatility modeling which is estimated using the Kalman filter. Furthermore, we focus on the movements of the latent variable (H_{mt}).

It should be observed that when $\sigma^2_{m\eta} = 0$, the model (66.7) becomes

$$\log\left[Std_c\left(\beta^b_{imt}\right)\right] = \mu_m + \upsilon_{mt} \tag{66.8}$$

This means that herding does not exist, so $H_{mt} = 0$ for all t.

A significant value of $\sigma^2_{m\eta}$ can be interpreted as the existence of herding, and a significant value of ϕ supports the autoregressive structure considered.

A restriction is that the herding process (H_{mt}) should be stationary, and as we do not expect herding to be an explosive process towards the market portfolio, it must be $|\phi_m| \leq 1$.

66.4.5 Herding with Market and Macroeconomic Variables

We explain above that $Std_c(\beta^b_{imt})$ changes over time depending on the level of herding in the market. However, it is interesting to study whether this behavior, extracted from $Std_c(\beta^b_{imt})$, is robust in the presence of variables that reflect the state of the market: the degree of volatility, the market return, or potential variables that reflect macroeconomic fundamentals.

Therefore, if the herding parameter becomes insignificant when these variables are included, then the changes in $Std_c(\beta^b_{imt})$ could be explained by changes in the fundamentals rather than herding.

In order to consider the influence of market volatility and market return, Hwang and Salmon (2004) include them as independent variables in the model (66.7), obtaining the model (66.9):

$$\log\left[Std_c\left(\beta^b_{imt}\right)\right] = \mu_m + H_{mt} + c_{m1}\log\ \sigma_{mt} + c_{m2}r_{mt} + \upsilon_{mt} \tag{66.9}$$

$$H_{mt} = \phi_m H_{mt-1} + \eta_{mt}$$

where $\log\sigma_{mt}$ and r_{mt} are the market log volatility and the market return at time t, respectively.

In a second step, Hwang and Salmon (2004) include the factors of the Fama and French (1993) in the model (66.9). Nevertheless, we also include the four factors of Carhart's (1997) model, considering the size (SMB), book-to-market (HLM), and momentum (PR1YR) factors, as model (66.10) exhibits:

$$\begin{aligned}\log\left[Std_c\left(\beta^b_{imt}\right)\right] = {} & \mu_m + H_{mt} + c_{m1}\ \log\ \sigma_{mt} + c_{m2}r_{mt} + c_{m3}SMB_t + c_{m2}HML_t \\ & + c_{m3}PR1YR + \upsilon_{mt}\end{aligned}$$

$$H_{mt} = \phi_m H_{mt-1} + \eta_{mt} \tag{66.10}$$

Lastly, we add three macroeconomic variables to the model (66.9) – dividend yield (DY), time spread (TS), and the short-term interest rate (STIR), in order to consider information variables representative of the economic cycle:

$$\begin{aligned}\log\left[Std_c\left(\beta^b_{imt}\right)\right] &= \mu_m + H_{mt} + c_{m1}\log\sigma_{mt} + c_{m2}r_{mt} + c_{m6}DY_t + c_{m7}TS_t \\ &\quad + c_{m8}STIR_t + \upsilon_{mt} \\ H_{mt} &= \phi_m H_{mt-1} + \eta_{mt} \end{aligned} \tag{66.11}$$

66.4.6 Generalized Herding Measurement in Linear Factor Models

In the previous section, we observe that we can measure herding in any factor, employing linear factor models. Therefore, supposing that the excess return of the fund i (r_{it}) follows this linear model:

$$r_{it} = \alpha^b_{it} + \sum_{k=1}^{K} \beta^b_{ikt\ f kt} + \varepsilon_{it}, i = 1, \ldots, N \text{ y } t = 1, \ldots, \mathrm{T} \tag{66.12}$$

where α^b_{it} is the intercept that changes over time, β^b_{ikt} are the coefficients on factor k at time t, f_{kt} is the realized value of factor k at time t, and ε_{it} has a mean zero with a variance σ^2_ε.

The factors in model (66.12) may be risk-specific factors or factors to detect an anomaly. One factor included is the excess market return, as in the conventional linear factor models.[1]

The superscript b on the betas indicates that they are biased betas under herding; therefore, the herding towards the factor k at time t can be captured by (66.13):

$$\beta^b_{ikt} = \beta_{ikt} - h_{kt}(\beta_{ikt} - E_c[\beta_{ikt}]) \tag{66.13}$$

where $E_c[\beta_{ikt}]$ is the cross-sectional expected beta for factor k at time t.

66.4.7 Robust Estimate of the Betas

The first step to calculate the different models considered is to estimate the market betas from the CAPM model (66.14) and from the four-factor Carhart model (66.15):

$$r_{it} = \alpha^b_{it} + \beta^b_{imt} r_{mt} + \varepsilon_{it} \tag{66.14}$$

[1]It should be clear that the linear factor model used does not require that the market is in equilibrium or efficient.

$$r_{it} = \alpha_{it}^{b} + \beta_{imt}^{b} r_{mt} + \beta_{iSt}^{b} SMB_t + \beta_{iHt}^{b} HML_{Ht} + \beta_{iPt}^{b} PR1YR_{Mt} + \varepsilon_{it} \tag{66.15}$$

With these betas we obtain the estimated cross-sectional standard deviation of the betas, and these are used in the state-space model.

Although the ordinary least squares (OLS) estimation is the most common technique to estimate the beta, this has some drawbacks. Firstly, they behave badly when the errors are not from a normal i.i.d. distribution, particularly when the data is heavily tailed, which are very frequent in return data. Furthermore, the existence of outliers may also influence on the OLS beta, thus leading to a distorted perspective on the relationship between asset returns and index returns.

In order to overcome these disadvantages and provide a better fit, Martín and Simin (2002) indicate that a robust estimation of beta should be implemented. One robust regression is the M-estimation method thorough the Huber estimation.

With the different betas estimated, we obtain the cross-sectional standard deviation of the betas on the market portfolio as

$$Std_c\left(\hat{\beta}_{imt}^{b}\right) = \sqrt{\frac{\sum_{i=1}^{N_t}\left(\hat{\beta}_{imt}^{b} - \overline{\hat{\beta}_{imt}^{b}}\right)^2}{N_t}} \tag{66.16}$$

where $\overline{\hat{\beta}_{imt}^{b}} = \frac{1}{N_t}\sum_{i=1}^{N_t}\hat{\beta}_{imt}^{b}$ and N_t is the number of funds in the month t.

Finally, we estimate the four state-space models considered: (66.7), (66.9), (66.10), and (66.11), with the standard deviation of the betas.

66.5 Data and Empirical Results

66.5.1 Data

The database was provided by Thomson Reuters. The data comprises the monthly returns obtained by all private pension funds with European equity investment vocation registered for sale in Spain (84 pension funds) and in the United Kingdom (690 pension funds).

The time period analyzed is from January 1999 to September 2010. We require that the pension funds present data for at least 24 months to ensure the consistency of the analyses. In this way, our database is free of the so-called survivorship bias.

The market benchmark used is the MSCI Europe index, given the European equity vocation of the pension funds, and it is necessary to use European benchmark portfolio to assess performance on an appropriate basis. The representative variable for risk-free asset is the 1-month Euribor rate.

The macroeconomic variables used are as follows:

- *Dividend yield*[2] is the ratio between the dividends paid out by the MSCI Europe in the previous 12 months and the current index price.
- *Time spread* is the annualized difference between the return on the EMU 10-year bond[3] and the 3-month Euribor rate.
- *Short-term interest rate* is the 3-month Euribor rate.

In the investment style analysis, we consider the four factors of the Carhart (1997) model: excess market return, size (SMB), book-to-market (HML), and momentum (PR1YR).

We follow the instructions by Fama and French (1993) to build the size and book-to-market factors. In regard to the size factor (SMB_t), we build the mimicking portfolio as the difference between the portfolio made up of the MSCI Europe small value price, MSCI Europe small core price, and MSCI Europe small growth price indices and the portfolio made up of the MSCI Europe large value price, MSCI Europe large core price, and MSCI Europe large growth price indices.

In relation with the book-to-market factor (HML_t), with the monthly returns obtained by the indices, the mimicking portfolio is the difference between the portfolio made up of the MSCI Europe small value price and MSCI Europe large value price indices and the portfolio made up of the MSCI Europe small growth price and MSCI Europe large growth price indices.

Finally, the 1-year momentum factor ($PR1YR_t$) is approached following the Carhart instructions, in our case, with the monthly returns obtained by a group of market indices representative of the geographic universe studied. As we analyze European equity funds, we use the 16 MSCI indices of the countries integrated in the MSCI Europe index: Austria, Belgium, Denmark, Finland, France, Germany, Greece, Ireland, Italy, the Netherlands, Norway, Portugal, Spain, Sweden, Switzerland, and the United Kingdom. All of the indices have been obtained from the official MSCI website.

Based on these 16 market indices, we build the equal-weight average of indices with the highest 30 % (five) 11-month returns, lagged 1 month, minus the equal-weight average of indices with the lowest 30 % (five) 11-month returns, lagged 1 month.

The descriptive statistics of the different risk factors are displayed on Table 66.1.

The Table 66.1 reveals that the excess market return is the factor with least mean return, presenting the minimum negative value. Nonetheless, the maximum value is also found in this factor; hence the standard deviation is somewhat greater than the rest.

[2]For its calculation we apply the difference between the monthly return obtained by the corresponding MSCI gross and the MSCI price; then we obtain the total of the 12 previous values for a determined month. Information obtained from MSCI: http://www.msci.com/

[3]Data obtained from the Bank of Spain: www.bde.es

Table 66.1 Properties of the risk factors

Variable	Average	Standard deviation	Minimum	Maximum	Kurtosis	Skewness
Market excess return	−0.0005	0.0488	−0.1350	0.1323	3.5650	−0.3032
SMB	0.0051	0.0269	−0.0813	0.0858	3.8340	−0.3169
HML	0.0017	0.0254	−0.0896	0.0972	6.2798	0.3345
PR1YR	0.0042	0.0392	−0.1146	0.1106	3.4888	−0.1526

This table includes the main statistics (average, standard deviation, minimum, maximum, kurtosis, and skewness) for the four risk factors calculated: market excess return, size (SMB), book-to-market (HML), and momentum (PR1YR)

With respect to kurtosis, it is high in all factors (greater than three), indicating leptokurtosis, and therefore, they are not Gaussian. Furthermore, it is also remarkable the negative skewness on all factors.

66.5.2 Empirical Results

Firstly, we estimate the betas of the models (66.14) and (66.15), and next we calculate the cross-sectional deviations, as we describe in Sect. 66.4.7.

The main characteristics of the standard deviations of the estimated betas are reflected in Table 66.2. This table differentiates between the market betas from the CAPM model and the market betas from the Carhart four-factor model. All of them are estimated with the Huber robust technique.

The first two columns of Table 66.2 show all cross-sectional standard deviations of the market betas: $Std_c\left(\hat{\beta}^b_{imt}\right)$. These are significantly different from zero in all cases. However, the betas from the CAPM model present negative skewness, while the skewness is positive in the market betas from the four-factor model, which is common in series with volatility. We also observe a high kurtosis, revealing non-normality. This is confirmed with the Jarque-Bera test, as we reject the null hypothesis of normality; therefore, the standard deviation of the betas is not Gaussian.

The correlation between these two market betas is high; then, if there is herding, this may reveal similar herding between models.

Finally, the properties of the cross-sectional deviations of the betas of the SMB, HML, and PR1YR factors also exhibit similar properties: they are not normal and negative skewness.

66.5.2.1 Results of Herding Towards the Market Factor

In this section we study the herding towards the market factor; that is to say, we start from the market betas of the CAPM and the four-factor models; after that, we calculate the standard deviation of the betas: $Std_c(\beta^b_{imt})$, and with these, we estimate the models (66.7), (66.9), (66.10), and (66.11).

The results of these models are displayed on Tables 66.3 and 66.4 for the Spanish and British pension funds, respectively.

Table 66.2 Properties of the cross-sectional standard deviation of estimated betas

	CAPM model	Four-factor Carhart (1997) model			
	Market return beta (A)	Market return beta (B)	Beta of SMB factor	Beta of HML factor	Beta of PR1YR
Panel A: Pension funds with European equity investment vocation in Spain					
Average	0.7592	0.0163	0.1361	−0.0423	0.8030
Standard deviation	0.1663	0.1408	0.2520	0.0887	0.1775
Minimum	0.0073	−0.2947	−0.5478	−0.1919	−0.0037
Maximum	1.0822	0.3548	0.8198	0.3626	1.1754
Kurtosis	10.5915	2.5177	3.0209	6.9308	10.2584
Skewness	−2.0215	0.1710	−0.1230	1.1898	−1.8888
Jarque − Bera test (p-value)	(0.0000)***	(0.0000)***	(0.0000)***	(0.0000)***	(0.0000)***
Correlation A–B	0.9507				
Panel B: Pension funds with European equity investment vocation in the United Kingdom					
Average	0.8116	0.2207	0.1275	0.0193	0.8725
Standard deviation	0.1418	0.2129	0.4075	0.1367	0.1346
Minimum	0.2079	−0.3876	−1.8963	−0.8442	0.2027
Maximum	1.1433	1.1677	1.4798	0.8482	1.2735
Kurtosis	6.8994	6.9736	4.5372	9.3433	9.2515
Skewness	−1.3375	1.5837	−0.6795	−0.1274	−1.6848
Jarque-Bera test (p-value)	(0.0000)***	(0.0000)***	(0.0000)***	(0.0000)***	(0.0000)***
Correlation A–B	0.8935				

This table is divided into two panels (A and B), corresponding to the Spanish and British pension funds. Each panel includes the main statistics (average, standard deviation, minimum, maximum, kurtosis, skewness, and Jarque-Bera normality test) of the cross-sectional standard deviation for the estimated market betas in the CAPM model and the four betas of Carhart (1997) model: market, size (SMB), book-to-market (HML), and momentum (PR1YR) betas. *** represents significance at 1 % level

We interpret the significance of the market variables included in the models (66.9), (66.10), and (66.11) as adjustment in the mean level (μ_m) of $\log[Std_c(\beta^b_{imt})]$ on the equation without herding, so we examine the degree of herding given the state of the market.

The Spanish results (Table 66.3) display that the model (66.7), both in the case of the market betas of the CAPM model as in Carhart model, presents evidence of quite persistent herding, as the coefficient $\hat{\phi}_m$ is large and significant.

Likewise, the standard deviation of n_{mt} (σ_{mn}) is significant; therefore, given the level of market volatility and return, Spanish pension funds lead to herding towards the market portfolio.

The coefficients of the model (66.9) show that the herding is still significant when the two market variables are included: volatility and return, which suggests that the changes in the volatility of the sensitivity factor $[Std_c(\beta^b_{imt})]$ can be explained by herding rather than by changes in fundamentals.

Table 66.3 Estimates of state-space models for herding towards the market factor in Spain

	Model (66.7), calculated with the market return beta of the CAPM model	Herding models with the market return beta of the four-factor Carhart (1997) model			
		Model (66.7), no exogenous variables	Model (66.9), with the factors: excess market return and volatility	Model (66.10), with the factors: excess market return, volatility, SMB, HML, and PR1YR	Model (66.11), with the factors: excess market return, volatility, DY, TS, and STIR
μ	0.9100(0.049)**	0.9780(0.06)*	0.9723 (0.102)	0.9833(0.201)	0.9162(0.100)
ϕ_m	0.8890(0.000)***	0.9201(0.000)***	0.6250(0.000)***	0.9755(0.000)***	0.4488(0.000)***
σ_{mu}	0.1931(0.0000)***	0.1926(0.135)	0.1550(0.050)**	0.1331(0.046)***	0.1510(0.040)**
σ_{mn}	0.2837(0.000)***	0.2487(0.002)***	0.2009(0.000)***	0.2770(0.001)***	0.2609(0.099)*
$\log\sigma_{mt}$			−0.0120(0.000)***	−0.0086(0.000)***	−0.0124(0.000)***
r_{mt}			0.0004(0.000)***	0.0034(0.000)***	0.0108(0.000)***
SMB				0.0260(0.211)	
HML				0.0025(0.319)	
PR1YR				−0.0024(0.196)	
DY					−0.0546(0.778)
TS					−0.0518(0.248)
STIR					−0.1869(0.541)

This table displays the estimates of the state-space models for herding towards the market factor in the Spanish pension funds for the period from January 1999 to September 2010. The first column exhibits model (66.7) results, calculated with the market return beta from the CAPM model. The second, third, fourth, and fifth columns show the results for the models (66.7), (66.9), (66.10), and (66.11) estimated with the market return beta from the four-factor Carhart (1997) model. SMB represents the size factor, HML is the book-to-market factor, PR1YR is the momentum factor, DY is the dividend yield, TS is the time spread, and STIR is the short-term interest rate

*, **, and *** represent significance at 10 %, 5 %, and 1 % level, respectively

Table 66.4 Estimates of state-space model for herding towards the market factor in the United Kingdom

	Model (66.7), calculated with the market return beta of the CAPM model	Herding models with the market return beta of the four-factor Carhart (1997) model: Model (66.7), no exogenous variables	Model (66.9), with the factors: excess market return and volatility	Model (66.10), with the factors: excess market return, volatility, SMB, HML, and PR1YR	Model (66.11), with the factors: excess market return, volatility, DY, TS, and STIR
μ	0.9407(0.073)*	0.7093(0.181)	0.9794(0.119)	0.6569(0.201)	0.2851(0.099)*
ϕ_m	0.9048(0.000)***	0.5399(0.000)***	0.8295(0.000)***	0.9011(0.000)***	0.9736(0.000)***
σ_{mu}	0.2084(0.201)	0.2080(0.444)	0.1095(0.100)	0.1043(0.000)***	0.1418(0.000)***
σ_{mn}	0.1540(0.075)*	0.0845(0.108)	0.1110(0.000)***	0.1031(0.000)***	0.2050(0.000)***
$\log \sigma_{mt}$			0.0061(0.005)***	0.0007(0.009)***	0.0005(0.091)*
r_{mt}			−0.0955(0.090)*	−0.0004(0.000)***	−0.0136(0.000)***
SMB				0.0183(0.170)	
HML				−0.0008(0.206)	
PR1YR				0.0011(0.217)	
DY					−0.0345(0.561)
TS					−0.0188(0.100)
STIR					0.0101(0.190)

This table displays the estimates of state-space models for herding towards the market factor in the UK pension funds for the time period from January 1999 to September 2010. The first column exhibits model (66.7) results, calculated with the market return beta from the CAPM model. The second, third, fourth, and fifth columns show the results for the models (66.7), (66.9), (66.10), and (66.11), estimated with the market return beta from the four-factor Carhart (1997) model. SMB represents the size factor, HML is the book-to-market factor, PR1YR is the momentum factor, DY is the dividend yield, TS is the time spread, and STIR is the short-term interest rate

*, **, and *** represent significance at 10 %, 5 %, and 1 % level, respectively

Therefore, the betas' deviation decreases when the market volatility rises, but increases with the level of market return, as the logarithm of market volatility and the market return have significant negative and positive coefficients, respectively.

As a result, when the market becomes riskier and is falling, $Std_c(\beta^b_{imt})$ decreases, while it increases when the market becomes less risky and rises. Therefore, a reduction in the standard deviation due to the herding process suggests that herd behavior is significant and exists independently of a particular state of the market.

The model (66.10) includes the SMB, HML, and PR1YR factors as explanatory variables, but none of them is significant. Moreover, the herding ($\hat{\phi}_m$) increases, so the results are very similar to those of the previous models.

Model (66.11) includes three macroeconomic variables, but none of them is significantly different to zero. Nonetheless, the herding is still persistent because the coefficient $\hat{\phi}_m$ is significant.

On the other hand, the UK results are reported in Table 66.4, where the coefficients $\hat{\phi}_m$ and σ_{mn} are significant and persistent. Equally, the additional variables of the models (66.10) and (66.11) are not significant.

Nonetheless, the market variables (logarithm of the volatility and market return) present signs contrary to the previous results, that is to say, positive and negative, respectively. Nevertheless, these results are consistent with previous studies, which find that herding arises most likely during market instability, in other words, periods of high volatility.

As a consequence, we obtain evidence of herding in the two countries analyzed. Nonetheless, we do not observe significant differences between herding models, as the factors of the investment styles or macroeconomic variables are not significant.

Therefore, herding is not influenced by the size of the company, the relation between book equity and market equity, or the momentum strategy. However, herding varies with the model applied; so the inclusion of additional variables is useful, but does not provide more information.

66.5.2.2 Results of Herding Towards Size, Book-to-Market, and Momentum Factors

Starting with the betas estimated for the different factors (size, book-to-market, and momentum) of the four-factor Carhart model (66.15), we also calculate the standard deviation of these betas: $Std_c(\beta^b_{iSt})$, $Std_c(\beta^b_{iHt})$, and $Std_c(\beta^b_{iPR1YRt})$.

After that, we repeat the above analysis, but, in this case, we study the herding towards the size, book-to-market, and momentum factors.

The results of these analyses are very similar to the previous one; consequently we do not display them,[4] but we show a summary in Table 66.5, indicating if the variables of the different models are significant. Nonetheless, next we discuss the different results.

[4]These tables are available upon request.

Table 66.5 Summary of the state-space models for herding towards style factors in Spain and the United Kingdom

	Herding towards SMB factor		Herding towards HML factor		Herding towards PR1YR factor	
	Spain	United Kingdom	Spain	United Kingdom	Spain	United Kingdom
Existence of herding	Yes, + and significant	Yes, + and significant	Yes, + and significant	Yes, + and significant	Yes, + and significant	Yes, + and significant
Market volatility	Signif.	Signif.	Signif.	Signif.	Signif.	Signif.
Market return	Signif.	Signif.	Signif.	Signif.	Signif.	Signif.
SMB	No signif.	No signif.	No signif.	No signif.	No signif.	No signif.
HML	No signif.	No signif.	No signif.	No signif.	No signif.	No signif.
PR1YR	No signif.	No signif.	No signif.	No signif.	No signif.	No signif.
DY	No signif.	No signif.	No signif.	No signif.	No signif.	No signif.
TS	No signif.	No signif.	No signif.	No signif.	No signif.	No signif.
STIR	No signif.	No signif.	No signif.	No signif.	No signif.	No signif.

This table displays a summary of the state-space models for herding towards the style factors (size, book-to-market, and momentum) in Spain and the United Kingdom, indicating the existence of herding and if the variables of the different models (66.7), (66.9), (66.10), and (66.11) are significant (signif.) or not significant (no signif.)

We find the same behavior in all factors: existence of herding and the herding coefficients ($\hat{\phi}_S$, $\hat{\phi}_H$, $\hat{\phi}_{PR1YR}$) are greater than 0.7, on average. Likewise, all standard deviations are highly significant and persistent ($\sigma_{s\eta}$, $\sigma_{H\eta}$ $\sigma_{PR1YR\eta}$).

However, we notice less degree of herding towards the book-to-market in the Spanish pension funds, presenting a parameter $\hat{\phi}_H$ around 0.3. This behavior is also perceived in the case of the herding towards the momentum factor in Spain and the United Kingdom, where the coefficient $\hat{\phi}_{PR1YR}$ is 0.6, on average.

Although the herding is less intense in these cases, the phenomenon is still significant towards the size (H_{St}), book-to-market (H_{Ht}), and momentum (H_{PR1YRt}) factors, and the results are similar to the herding towards the market (H_{mt}).

Furthermore, we also observe that the additional variables of the four-factor model (66.10) and the macroeconomic variables of the model (66.11) are not significant.

Lastly, the logarithms of market volatility and the market return, especially the latter, explain less the cross-sectional standard deviation of the SMB, HML, and PR1YR betas, as the market return is less significant or not significant.

66.5.2.3 Relationship Between the Herding Towards Factors and Countries

As in the previous section we find similar herding towards factors and between countries, we study the herding patterns in greater detail in Table 66.6, displaying the correlations between the different herding coefficients (coefficients of herding towards the market, size, book-to-market, and momentum factors).

Table 66.6 Relation between the herding of different pension funds in each country

	Spain market	Spain SMB	Spain HML	Spain PR1YR	UK market	UK SMB	UK HML	UK PR1YR
Spain market	1							
Spain SMB	-0.255^{***}	1						
Spain HML	-0.156^{***}	-0.005^{***}	1					
Spain PR1YR	-0.134^{***}	0.111^{***}	0.093^{***}	1				
UK market	0.053^{***}	0.075^{***}	0.022^{**}	0.286^{***}	1			
UK SMB	-0.012^{***}	0.009	-0.068^{***}	-0.051^{***}	0.099^{***}	1		
UK HML	-0.045^{***}	-0.057^{***}	0.041^{***}	0.082^{***}	-0.009^{***}	0.118^{***}	1	
UK PR1YR	0.034^{***}	0.050^{***}	-0.230^{***}	-0.096^{***}	0.129^{***}	0.078^{***}	0.128^{***}	1

This table represents the correlation coefficients of herding measures towards different factors: market return (represented as market), size (SMB), book-to-market (HML), and momentum (PR1YR), from model (66.9) in the different factors and pension funds in Spain and the United Kingdom

*, **, and *** represent significance at 10 %, 5 %, and 1 % level, respectively

Although we do not display the herding towards the last three factors in the previous section, we calculate the correlation between their coefficients, considering the herding measures obtained in the model (66.9), as they are more significant, in general.

This table exhibits the correlation between herding towards factors (h_{mt}, h_{St}, h_{Ht}, and h_{PR1YRt}) at 5 % significance level.

The results show that among the 36 pairwise correlation coefficients, 15 are significantly positive, 12 are significantly negative, and 4 are not significant. We also detect positive negative correlations between common factors in both countries, except in the momentum factor, so we notice international herding.

Overall, we notice more positive correlation between the UK factors and between Spanish and the UK factors than between Spanish factors.

In conclusion, we observe a relationship between the herding coefficients for the same factor in Spain and the United Kingdom.

66.6 Conclusions

The pension fund industry has acquired great significance in recent years, especially in Europe, due to doubts about the future viability of public pensions. For this reason, more works focus on their study, especially in the analysis of their performance and the manager behavior.

However, it is also interesting to study the consequences of their investment on the market, as it can affect the market efficiency. Specifically, we study the existence of herding, a phenomenon that arises when investors decide to imitate the observed decisions of others, instead of following their own beliefs and information. This phenomenon can lead to a situation in which market prices do not reflect all of the information and the market moves towards inefficiency.

We analyze whether the behavior of the pension fund managers studied arises the herding behavior in the equity markets of Spain and the United Kingdom, applying a focus less used in financial literature: the estimated cross-sectional of standard deviations of market betas.

The estimation technique is also less common because we use the state-space models, employing the Kalman filter. Furthermore, in order to carry out a robust estimation of the beta from the CAPM model and from the four-factor model of Carhart (1997), we implement the robust M-estimation with Huber estimator.

We apply different models in order to analyze the existence of herding. Firstly, we only detect the existence and persistence of the phenomenon. After that, we add two market variables (volatility and return), then three style factors (size, book-to-market, momentum), and finally, three macroeconomic variables (dividend yield, time spread, and short-term interest rate), in order to examine their influence on herding.

In addition, we check the existence of herding towards the size, book-to-market, and momentum factors; for this purpose, we use the estimated betas of these factors in Carhart's model. The analysis of the momentum factor is innovative, as we do not find previous studies that analyze this aspect.

The results obtained are similar for the two countries studied, as well as for the different types of investment. These reveal the existence of herding towards the market and the style factors, showing significant movements and persistence independently from and given market conditions. Nonetheless, this effect is smaller in the case of herding towards the different style factors.

Additionally, we do not find relevant influence of the macroeconomic variables, nor the style factors with any of the different models, so these variables do not significantly influence herding. As a consequence, we note that herding arises regardless of the company size, its situation (growing or not), as well as the implemented strategies.

Hence, as we only study equity pension funds, we confirm that the Spanish and British pension fund managers influence the degree of herding in the stock markets.

Moreover, we find the existence of herding in all models and markets, so pension fund managers encourage the existence of herding towards the market, size, book-to-market, and momentum factors. As a consequence, managers follow the performance of the market and the different styles more than they should in equilibrium, so they move towards matching the return on individual assets with that of the market and styles. Additionally, we discover a relationship between the herding coefficients for the same factor in Spain and the United Kingdom.

As a result, the behavior of Spanish and British pension fund managers influences the market and, consequently, the performance of pension funds.

Appendix 1: The State-Space Models

We remark that the herding parameter is not observed, so in order to extract herding parameter we apply state-space models.

A state-space model is defined by two equations:

$$Y_t = c + SX_t + e_t \tag{66.17}$$

$$X_t = d + HX_{t-1} + z_t \tag{66.18}$$

where:

X_t is the hidden vector at time *t*.

Y_t is the observation vector at time *t*.

c and *d* are vectors with constants.

e is the error.

z is the state error.

e and *z* are both multivariate normally distributed, with mean zero and covariance matrices of R and Q, respectively.

Those models can be estimated by using the Kalman filter, which is an algorithm to perform filtering on the state-space model.

The estimate of the state equation by the Kalman filter algorithm also offers a smoothing time series, by performing fixed interval smoothing, i.e., computing $Y_{t|t} = P[Y_t|Y_1, \ldots, Y_{t-1}]$ for $t \leq T$.

The objective is, in the formula (66.17), to minimize the difference between the observation Y_t, and the prediction based on the previous observations, ($Y_{t|t} = P[Y_t|Y_1, \ldots, Y_{t-1}]$) by recursive maximum likelihood estimation.

The Kalman filter can be considered as an online estimation procedure, which is used to estimate the parameters online when new observations are entered after they have already been estimated. On the other hand, the smoothed Kalman filter is a method only used when the total series are observed.

The Kalman filter results are close to the maximum likelihood estimates, while the smoother results are exact to the maximum likelihood estimates.

Appendix 2: Robust Estimate of the Betas

In order to calculate the different betas from the CAPM model and from the four-factor Carhart model, the ordinary least squares (OLS) estimation is the most common technique for estimating betas; however, this has some drawbacks.

Firstly, they behave badly when the errors are not from a normal i.i.d. (independent and identically distributed) distribution, particularly when the data is heavily tailed, which are very frequent in return data.

Furthermore, the existence of outliers may also influence on the OLS beta, thus leading to a distorted perspective on the relationship between asset returns and index returns.

In order to overcome these disadvantages and provide a better fit, Martín and Simin (2002) indicate that a robust estimation of beta should be implemented. One of the most commonly applied methods of robust regression is the M-estimation method, a generalization of maximum likelihood estimation.

In order to explain this estimate method, we considered a linear model as a starting point:

$$y_i = X_i\beta + \varepsilon_i \tag{66.19}$$

where $i = 1,.., n$

Thus, the fitted model is

$$y_t = X_t b + e_t \tag{66.20}$$

The M-estimate principle is to minimize the objective function:

$$\sum_{i=1}^{n} \rho(e_i) = \sum_{i=1}^{n} \rho(y_i - X_i b) \tag{66.21}$$

where the function $\rho(.)$ gives the contribution of each residual to the objective function.

Table 66.7 Objective functions and weight functions for the ordinary least squares estimation and the Huber estimation

Estimation method	Objective function (ρ)	Weight function (w_i)
Ordinary least square (OLS)	e^2	1
Huber estimation	$e^2/2$ when $\|e\| \le k$	1 when $\|e\| \le k$
	$k\|e\| - k^2/2$ when $\|e\| > k$	$k/\|e\|$ when $\|e\| > k$

This table compares the objective functions and weight functions for the ordinary least squares estimator and the Huber estimator

If we define $\psi = \rho'$, as the first order derivative of $\rho(.)$, by differentiating the objective function with respect to b and setting the partial derivatives to zero, we obtain a system of estimating equations:

$$\sum_{i=1}^{n} \psi(y_i - X_i b) X_i' = 0 \tag{66.22}$$

If the weight function is $w(e) = \psi(e)/e$ and $w_i = w(e_i)$, the estimating equations become

$$\sum_{i=1}^{n} w_i e_i X_i' = 0 \tag{66.23}$$

These equations can be solved as a weighted least squares problem, with the objective of minimizing:

$$\sum_{i=1}^{n} w_i^2 e_i^2 \tag{66.24}$$

The weights depend on the residuals, the residuals depend on the estimated coefficients, and the estimated coefficients depend on the weights, so an iteration procedure is needed in order to solve the problem.

To solve this iterative procedure, we apply the Huber estimation, given that this allows us to determine the weighted, the residuals, and the estimated coefficients.

In order to compare the OLS estimator with the robust Huber estimator, Table 66.7 distinguishes the objective functions and weighted functions for each one of the methods.

In Table 66.7 we observe that both functions increase without bound, as the residuals departs from zero; nonetheless, the Huber objective function increases more slowly.

In fact, the least squares assigns equal weight to each observation, but the weights of the Huber estimator decline for $|e| > k$, where e is the residual term and k is called a tuning constant for the Huber estimation.

In the OLS estimation, a smaller k parameter provides more resistance to outliers; however, it offers a lower efficiency when the errors are normally distributed. In contrast, with Huber estimation, k has a general value of $k = 1.345\ \sigma$

(where σ is the conventional standard deviation), producing 95 % efficiency when the errors are normal, and it also offers protection against outliers; therefore, this estimation is better than the OLS estimation.

References

Avery, C., & Zemsky, P. (1998). Multidimensional uncertainty and herd behavior in financial markets. *American Economic Review, 88*, 724–748.

Banerjee, A. (1992). A simple model of herd behavior. *Quarterly Journal of Economics, 107*, 797–818.

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A theory of fads, fashion, custom, and cultural change as informational cascades. *Journal of Political Economy, 100*, 992–1026.

Blasco, N., & Ferreruela, S. (2007). Comportamiento Imitador en el Mercado Bursátil Español: Evidencia Intradiaria. *Revista de Economía Financiera, 13*, 56–75.

Blasco, N., & Ferreruela, S. (2008). Testing intentional herding in familiar stocks: An experiment in an international context. *The Journal of Behavioral Finance, 9*(2), 72–84.

Blasco, N., Corredor, P., & Ferreruela, S. (2011). Detecting intentional herding: What lies beneath intraday data in the Spanish stock market. *Journal of the Operational Research Society, 62*, 1056–1066.

Bohl, S., Martin, T., Klein, A. (2011). *Short selling bans and institutional investors herding behavior: Evidence from the global financial crisis*. Working paper. http://ssrn.com. Accessed 27 Dec 2011.

Bowe, M., & Domuta, D. (2004). Investor herding during financial crisis: A clinical study of the Jakarta stock exchange. *Pacific-Basin Finance Journal, 12*(4), 387–418.

Brennan, M. (1993). *Agency and asset prices*. Finance working paper no. 6–93, UCLA.

Calvo, G., & Mendoza, E. (2000). Rational contagion and the globalization of securities markets. *Journal of International Economics, 51*(1), 79–113.

Carhart, M. (1997). On persistence in mutual fund performance. *The Journal of Finance, 52*(1), 57–82.

Chang, E. C., Cheng, J. W., & Khorana, A. (2000). An examination of herd behavior in equity markets: An international perspective. *Journal of Banking and Finance, 24*, 1651–1679.

Chari, V. V., & Kehoe, P. (2004). Financial crises as herds: Overturning the critiques. *Journal of Economic Theory, 119*(1), 128–150.

Christie, W. G., & Huang, R. D. (1995). Following the pied piper: Do individual returns herd around the market? *Financial Analyst Journal, 51*(4), 31–37.

Demirer, R., & Kutan, A. (2006). Does herding behavior exist in Chinese stock markets? *Journal of International Financial Markets, Institutions and Money, 16*(2), 123–142.

Demirer, R., Kutanand, A., & Chen, C. (2010). Do investors herd in emerging stock markets?: Evidence from the Taiwanese market. *Journal of Economic Behavior & Organization, 76*(2), 283–295.

Devenow, A., & Welch, I. (1996). Rational herding in financial economics. *European Economic Review, 40*, 603–615.

Economou, F., Kostakis, A., & Philippas, N. (2011). Cross-country effects in herding behavior: Evidence from four south European markets. *Journal of International Financial Markets Institutions and Money, 21*(3), 443–460.

Fama, E. F., & French, K. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1), 3–56.

Fernández, V. (2011). Los fondos de pensiones Como activo financiero familiar. Comparativa europea y perspectivas en España. *Ahorro familiar en España, 17*, 319–332.

Ferruz, L., Sarto, J. L., & Vicente, L. (2008a). Herding behavior in Spanish equity funds. *Applied Economics Letters, 15*(7), 573–576.

Ferruz, L., Sarto, J. L., & Vicente, L. (2008b). Convergencia estratégica en la industria española de fondos de inversión. *El Trimestre Económico, 75*(4), 1043–1060.

Ferson, W., & Harvey, C. R. (1991). The variation of economic risk premiums. *Journal of Political Economy, 99*, 285–315.

Ferson, W., & Harvey, C. R. (1993). The risk and predictability of international equity returns. *Review of Financial Studies, 6*, 527–566.

Ferson, W., & Korajczyk, R. A. (1995). Do arbitrage pricing models explain the predictability of stock returns? *Journal of Business, 68*, 309–349.

Fong, K., Gallagher, D. R., Gardner, P., & Swan, P. L. (2011). Follow the leader: Fund managers trading in signal-strength sequence. *Accounting & Finance, 51*(3), 684–710.

Fromlet, H. (2001). Behavioral finance – theory and practical application. *Business Economics, 36*, 63–69.

Gleason, K. C., Lee, C. I., & Mathur, I. (2003). Herding behavior in European futures markets. *Finance Letters, 1*, 5–8.

Gleason, K. C., Mathur, I., & Peterson, M. A. (2004). Analysis of intraday herding behavior among the sector ETFs. *Journal of Empirical Finance, 11*, 681–694.

Gompers, P. A., & Metrick, A. (2001). Institutional investors and equity prices. *Quarterly Journal of Economics, 116*(1), 229–259.

González, P., & García, J. C. (2002). Regulación del sector de seguros y planes de pensiones. *Revista ICE, Sistema Financiero: Novedades y Tendencias, 801*, 51–67.

Goodfellow, C., Bohl, M. T., & Gebka, B. (2009). Together we invest? Individual and institutional investors' trading behavior in Poland. *International Review of Financial Analysis, 18*(4), 212–221.

Grinblatt, M., Titman, S., & Wermers, R. (1995). Momentum investment strategies, portfolio performance and herding: A study of mutual fund behavior. *American Economic Review, 85*, 1088–1105.

Harvey, C. R. (1989). Time-varying conditional covariances in tests of asset pricing models. *Journal of Financial Economics, 24*, 289–317.

Henker, J., Henker, T., & Mitsios, A. (2006). Do investors herd intraday in Australian equities? *International Journal of Managerial Finance, 2*(3), 196.

Hirshleifer, D., Subrahmanyam, A., & Titman, S. (1994). Security analysis and trading patterns when some investors receive information before others. *Journal of Finance, 49*(5), 1665–1698.

Hsieh, S., Yang, T.-Y., & Tai, Y. (2010). Positive trading effects and herding behavior in Asian markets: Evidence from mutual funds. *The International Journal of Business and Finance Research, 4*(2), 177–188.

Hwang, S., & Salmon, M. (2004). Market stress and herding. *Journal of Empirical Finance, 11*, 585–616.

Jame, R. E. (2011). *Pension fund herding and stock return*. Working paper. http://ssrn.com. Accessed 21 Dec 2011.

Kahnemann, D., & Tversky, A. (1979). Prospect theory; an analysis of decisions under risk. *Econometrica, 47*, 262–291.

Kim, K., & Wei, S. (2002). Offshore investment funds: Monsters in emerging markets? *Journal of Development Economics, 68*(1), 205–224.

Lakonishok, J., Shleifer, A., & Vishny, R. (1992a). The impact of institutional trading on stock prices. *Journal of Financial Economics, 32*, 23–43.

Lakonishok, J., Shleifer, A. Vishny, R. (1992b). The structure and performance of the money management industry. *Brookings Papers: Microeconomics, 1992*, 339–391.

Lillo, F., Moro, E., Vaglica, G., & Mantegna, R. (2008). Specialization of strategies and herding behavior of trading firms in a financial market. *New Journal of Physics, 10*, 043019 (15 pp).

Lin, A. Y., & Swanson, P. E. (2003). The behavior and performance foreign investors in emerging equity markets: Evidence from Taiwan. *International Review of Finance, 4*(3–4), 189–210.

Lobao, J., Serra, A., & Serra, P. (2007). Herding behavior evidence from Portuguese mutual funds. In G. Gregoriou (Ed.), *Diversification and portfolio management of mutual funds* (pp. 167–197). New York: Palgrave MacMillan.
Martín, R., & Simin, T. (2002). Robust beta mining. In U. Fayyad, G. Grinstein, & A. Wierse (Eds.), *Information visualization in data mining and knowledge discovery* (pp. 309–312). San Francisco: Morgan Kaufmann.
Maug, E., & Naik, N. (1996). *Herding and delegated portfolio management: The impact of relative performance evaluation on asset allocation*. IFA working paper 223. http://ssrn.com. Accessed 27 Dec 2011.
OECD. (2010). *Pension markets in focus*, 7. http://www.oecd.org. Accessed 22 Jul 2011.
Oehler, A., & Goeth-Chi, G. (2000). *Institutional herding in bond markets*. Working paper, Digital archive for economics and business studies. http://ssrn.com. Accessed 27 Dec 2011.
Patterson, D. M., & Sharma, V. (2006). *Do traders follow each other at the NYSE?* Working paper, University of Michigan-Dearborn.
Puckett, A., & Yan, X. (2007). The determinants and impact of short-term institutional herding. Working paper. http://ssrn.com. Accessed 27 Dec 2011.
Rajan, R. G. (1994). Why credit policies fluctuate: A theory and some evidence. *Quarterly Journal of Economics, 436*, 399–442.
Roll, R. (1992). A mean/variance analysis of tracking error. *Journal of Portfolio Management, 18*(4), 13–22.
Sahut, J. M., Gharbi, S., & Gharbi, H. O. (2011). Stock volatility, institutional ownership and analyst coverage. *Bankers Markets & Investors* 110 (Janvier-Febrier).
Sanz, J. (1999). Los planes de pensiones y jubilación Como sistema complementario de las pensiones. *Acciones e investigaciones sociales, 9*, 157–166.
Scharfstein, D. S., & Stein, J. C. (1990). Herd behavior and investment. *American Economic Review, 80*, 465–479.
Shapour, M., Raeei, R., Ghalibaf, H., & Golarzi, H. (2010). Analysis of herd behavior of investors in Tehran stock exchange using with state-space model. *Journal of Financial Accounting Research Summer, 2*(4), 49–60.
Shefrin, H. (2000). *Beyond greed and fear; understanding behavioral finance and the psychology of investing*. Cambridge: HBS Press.
Tan, L., Chiang, T. C., Mason, J. R., & Nelling, E. (2008). Herding behavior in Chinese stock markets: An examination of A and B shares. *Pacific-Basin Finance Journal, 16*, 61–77.
Thaler, R. (1991). *Quasi-rational economics*. New York: Russel Sage Foundation.
Trueman, B. (1994). Analyst forecasts and herding behavior. *Review of Financial Studies, 7*, 97–124.
Tversky, A., & Kahnemann, D. (1986). Rational choice and the framing of decisions. *Journal of Business, 59*, 251–278.
Uchida, H., & Nakagawa, R. (2007). Herd behavior in the Japanese loan market: Evidence from bank panel data. *Journal of Financial Intermediation, 16*, 555–583.
Voronkova, S., & Bohl, M. (2005). Institutional traders' behavior in an emerging stock market: Empirical evidence on polish pension fund investors. *Journal of Business Finance and Accounting, 32*(7–8), 1537–1560.
Walter, A., & Weber, F. (2006). Herding in the German mutual fund industry. *European Financial Management, 12*(3), 375–406.
Wang, D. (2008). *Herd behavior towards the market index: Evidence from 21 financial markets*. IESE business school working paper. http://ssrn.com. Accessed 27 Dec 2011.
Weiner, R. J. (2006). *Do birds of a feather flock together? Speculator herding in the world oil market*. Discussion paper. Resources for the Futures, RFF DP 06-31. http://ideas.repec.org/p/rff/dpaper/dp-06-31.html. Accessed 27 Dec 2011.
Wermers, R. (1999). Mutual fund herding and the impact on stock prices. *Journal of Finance, 54*, 581–622.
Wylie, S. (2005). Fund manager herding: A test of the accuracy of empirical results using U.K. Data. *The Journal of Business, 78*(1), 381–403.

Estimating the Correlation of Asset Returns: A Quantile Dependence Perspective

67

Nicholas Sim

Contents

Abstract

In the practice of risk management, an important consideration in the portfolio choice problem is the correlation structure across assets. However, the correlation is an extremely challenging parameter to estimate as it is known to vary substantially over the business cycle and respond to changing market conditions. Focusing on international stock markets, I consider a new approach of estimating correlation that utilizes the idea that the condition of a stock market is related to its return performance, particularly to the conditional quantile of its return, as the lower return quantiles reflect a weak market while the upper quantiles reflect a bullish one.

Combining the techniques of quantile regression and copula modeling, I propose the *copula quantile-on-quantile regression* (C-QQR) approach to construct the correlation between the conditional *quantiles* of stock returns. The C-QQR approach uses the copula to generate a regression function for modeling the dependence between the conditional quantiles of the stock returns under consideration. It is estimated using a two-step quantile regression

N. Sim
School of Economics, University of Adelaide, Adelaide, SA, Australia
e-mail: nicholas.sim@adelaide.edu.au

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_67,

procedure, where in principle, the first step is implemented to model the conditional quantile of one stock return, which is then related in the second step to the conditional quantile of another return. The C-QQR approach is then applied to study how the US stock market is correlated with the stock markets of Australia, Hong Kong, Japan, and Singapore.

Keywords
Stock markets • Copula • Correlation • Quantile regression • Quantile dependence • Business cycle • Dynamics • Risk management • Investment • Tail risk • Extreme events • Market uncertainties

67.1 Introduction

In the practice of risk management, an important consideration in the portfolio choice problem is the correlation structure across assets. The correlation is especially crucial for conveying the level of portfolio risk as it is the parameter that underscores the extent of how well diversified a given portfolio is. Nevertheless in practice, estimating the actual level of correlation is a notoriously difficult task, where existing research has overwhelmingly shown that the true correlation, be it the correlation of equities, bonds, and exchange rates, may fluctuate significantly over the business cycle and during extreme events.[1] Over the past two decades, the literature has offered new ways of estimating asset correlation that depart from the premise that the correlation is constant. To complement these existing methodologies, this chapter offers a new perspective based on the concept of dependence between conditional quantiles to motivate a new approach of modeling correlation structure that takes into account that the correlation may be sensitive to the performance of financial markets.

Since the early findings of Erb et al. (1994) and Longin and Solnik (1995), among others, it is well accepted among academics and practitioners that the correlation may respond to changing economic circumstances.[2] A prime example can be found in the study of how international equity markets are dependent, where the literature has provided ample evidence that the level of dependence tends to be

[1]For instance, equity returns are more highly correlated during business cycle downturns and bear markets (Erb et al. 1994; Longin and Solnik 1995, 2001; Ang and Chen 2002); the same is true for exchange rates returns (Patton 2006; Bouyè and Salmon 2009). Likewise, the deviation in the stock-bond correlation during bear markets is documented by Guidolin and Timmermann (2005).

[2]For example, Erb et al. (1994) examine the dependence of G-7 equity stock markets by computing semicorrelations and find that the correlation is generally larger when the equity returns and output growth of these countries are below than above their respective means, thus when the economies and financial markets in these countries are bearish. Longin and Solnik (1995) examine the dependence of the market returns of Switzerland plus G-7, less Italy, by using a version of the multivariate GARCH model to show that the correlation is larger in times of greater market uncertainties.

stronger when markets become bearish (e.g., Erb et al. 1994). In order to capture such salient features about correlation, it is important to address the possibility that the actual level of correlation is contingent on current market conditions. Recent techniques of modeling asset dependence are developed with this objective in mind. They include regimes-witching frameworks to model jumps in correlation between normal and bearish states (e.g., Ang and Bekaert 2002; Guidolin and Timmermann 2005); extreme value theory that facilitates estimating the level of dependence during extreme events (e.g., Longin and Solnik 2001; Ang and Chen 2002; Poon et al. 2004; Heffernan and Tawn 2004); the copula approach, where the copula is a function that expresses the dependence structure of assets that also delivers an explicit measure of the dependence between the tail distributions of these assets (e.g., Patton 2006; Bouyè and Salmon 2009; Chollete et al. 2011); and the mixed copula approach that combines several copula functions where a different copula may be specified for a different state where financial markets are found (e.g., Hu 2006; Okimoto 2008). These techniques are widely popular in financial applications and are especially powerful for eliciting the properties of correlation across various market conditions. Nevertheless, there are also some limitations in the scope of how they may be used in financial applications.

Take the extreme value theory approach of Longin and Solnik (2001), for example. Extreme value theory is relevant for the study of the conditional correlation between the extreme tail distributions of asset returns, hence is particularly useful for providing results on asymptotic tail dependence. However, as the asymptotic tail dependence is only suitable for describing the level of dependence between markets that are significantly bearish or bullish, the extreme value theory approach may not be amenable for examining the level of dependence when the extent of such market conditions is, loosely speaking, "less severe" or "mild." Similarly, the concept of tail dependence in the copula approach is an asymptotic concept, and like extreme value theory, it does not convey the level of dependence for the different degrees of how bearish or bullish markets are. Regime-switching models can paint a broader picture on the characteristics of correlation, but might also require specifying a large number of regimes that makes them computationally burdensome to estimate.

By focusing on modeling the correlation of international stock markets as the application, this chapter contributes to the existing literature by offering a simple approach to estimate the level of dependence that pertains not only to extreme market conditions but also to varying degrees of market bearishness or bullishness. It does so by simply relating the severity (or a lesser extent) of these market conditions to the distributions of the market returns under consideration. For example, we may associate the lower tail of the return distribution with a bearish market and the upper tail of this distribution with a bullish one. And between, say, the 10th and 30th percentiles of a stock return, the 10th return percentile may be perceived as associated with a market more bearish than the market associated with the 30th return percentile. By taking advantage of the fact that the quantile information of a market return reflects the market performance, concepts such as a market being "mildly" or "severely" bearish or bullish can be expressed more concretely by relating them to certain quantiles on the distribution of that market return.

To estimate the level of dependence for specific market conditions, this chapter proposes a quantile dependence approach that looks at how the conditional *quantiles* of the market returns are correlated. From this perspective, to study how the equity markets of, say, the United States and Japan are dependent when they are bearish (bullish), one suggestion is to construct the correlation between the 10th (90th) conditional percentiles of the United States and Japan market returns. When studying their dependence during less bearish (bullish) times, we may construct the correlation of their return quantiles that are further away from the left (right) tails and closer towards the center of distributions. In other words, from the perspective that the distributional information (in particular, the quantile information) of a market return is indicative of the concurrent market condition, we may model the dependence structure of equity markets in a holistic and flexible way by estimating the correlation between the conditional quantiles of their returns, so that the level of dependence pertaining to a wide range of market conditions may be uncovered.

To model the dependence of the return quantiles, a new framework combining the techniques of quantile regression and copula modeling is proposed. Quantile regression is a statistical tool for examining how the quantile of a variable is dependent on some other conditioning variables and thus is useful in this study as its main objective is to investigate the link between the distributional (or quantile) information on asset returns on the one hand and the correlation between these assets on the other. Together with the quantile regression technique, I use the copula model, which is popular among practitioners for its flexibility in modeling dependence between variables that may have complex, nonstandard joint distributions.[3] When the study of correlation is considered, the copula approach is extremely useful as the correlation structure is summarized by the parameter in the copula function. In the case of the Gaussian or Student-t copula, the copula parameter is itself the correlation coefficient.

Using these techniques to estimate how the conditional *quantile* of an asset return is correlated with the conditional *quantile* of another asset return, this chapter thereby proposes a methodology dubbed as the *copula quantile-on-quantile regression* (C-QQR) approach. The C-QQR approach is computationally convenient to implement as it is based on a two-step quantile regression procedure. Specifically, when computing the correlation between the market returns of, say, the United States and Japan, the C-QQR approach proceeds by first estimating the conditional quantile of the US return by way of quantile regression on an *auxiliary equation*, which is not an equation of main interest. Then, using information from this regression, it proceeds to estimate the correlation between the conditional quantiles of the market returns of Japan and the United States. This is achieved by implementing quantile regression on a *quantile dependence equation*, which contains a parameter that expresses the dependence between the return quantiles of these markets. As the quantile dependence equation articulates how the *quantile* of a return is related to the *quantile* of another return, it can be used to examine the entire dependence structure between these assets where their relationship is assumed to be contingent on their quantile information.

[3]Introduction to copula models can be found in Nelsen (2006) and Trivedi and Zimmer (2007).

The application in this chapter focuses on modeling how the US market return is correlated with the market returns of Australia, Hong Kong, Japan, and Singapore. I employ the C-QQR approach to estimate the correlation between the 10th–90th conditional percentiles of the US return and the 10th–90th conditional return percentiles of Australia, Hong Kong, Japan, or Singapore (in decile intervals), leading to a total of 81 different correlation estimates for each return pair.[4] These estimates exhibit substantial variation, implying that the correlation varies considerably across different levels (thus quantiles) of market returns. In particular, the C-QQR approach shows that the correlation between returns at the center of distributions, such as the correlation between the median returns, is typically weaker. This implies that equity markets are less dependent when conditions are not extreme. It also shows that the correlation is stronger between returns deep in the left tails and that the correlation between the tenth return percentiles is consistently larger than the correlation between the median returns. This observation is in line with the existing evidence that stock markets are more strongly dependent when they are bearish.

After computing the C-QQR correlations, it is straightforward to construct a correlation time series by assigning a correlation estimate (from the pool of 81 C-QQR correlation estimates) to the realized returns in each period t. Calling this the dynamic C-QQR correlation, it is interesting to compare this constructed series against estimates that are obtained using conventional methods, such as the celebrated dynamic conditional correlation (DCC) framework of Engle (2002) that is designed for the study of how correlation evolves across time. Interestingly, although the C-QQR and DCC approaches are based on completely unrelated modeling principles – the C-QQR approach is based on quantile regressions and the DCC approach is based on the GARCH framework – the dynamic C-QQR correlation turns out to have similar visual characteristics as the DCC. This underscores another strength of the C-QQR approach – its ability to capture the salient features of the actual correlation dynamics.

This chapter draws heavily from Sim (2012) on modeling quantile dependence but contains two important differences. First, it considers a different specification of the auxiliary equation as the one in Sim (2012). As it turns out, where the results are suppressed for conciseness sake, the estimation outcomes of the C-QQR correlation are not sensitive to using either the current auxiliary equation or the one in Sim (2012), suggesting at first pass that the C-QQR correlation estimates are fairly robust to mis-specification of the auxiliary equation. Second, the application in Sim (2012) is limited in scope as it focuses on modeling the correlation of the US market return with the market and sectoral returns of Australia. Extending this work, I present new results on the correlation of the US market with the stock markets of Hong Kong, Japan, and Singapore in addition to Australia.

[4]Take the US-Australia return pair, for example. The 81 correlation estimates consist of estimates of the correlation between the tenth US and tenth Australian return percentiles, the tenth US and 20th Australian return percentiles, and so on, up to the 90th US and 90th Australian return percentiles.

67.2 The C-QQR Model

67.2.1 Specification

To construct the C-QQR model, an *auxiliary equation* and a *quantile dependence equation* must be specified. In our application, the auxiliary equation is utilized for modeling the conditional quantile of the US return quantile, and the quantile dependence equation is used for relating the dependence between the conditional quantile of the US market return and the conditional quantile of the market returns of Australia, Hong Kong, Japan, and Singapore. To specify the auxiliary equation, a model is postulated to link the US return (x) to a set of conditioning variables (x) as

$$x_t = \mathbf{b}^\top \mathbf{z}_t + v_t, \tag{67.1}$$

where v_t is the innovation in x_t. In (67.1), x_t is expressed as an autoregression with 12 lags to capture any mean reverting behavior of the US stock return, but it should be emphasized that the final outcome of the correlation estimate is not sensitive to using different numbers of lags. While the auxiliary equation is specified as an autoregression of x_t, another possibility is to motivate a model such that the US return is dependent on certain macroeconomic aggregates[5] or a model that is based on a general equilibrium framework with certain restrictions imposed. For example, instead of an autoregression, Sim (2012) takes into account of these considerations by specifying the auxiliary equation as a function of the US industrial production,[6] which is an important determinant of stock return from both theoretical and empirical perspectives.[7]

To specify the quantile dependence equation, a dependence function h is postulated to relate the return of Australia, Hong Kong, Japan, or Singapore (y) to the US return (x) as

$$y_t = h(x_t; \widetilde{\varphi}(u_t, v_t)), \tag{67.2}$$

[5]According to the seminal work of Chen, Roll, and Ross (1986), asset prices could also be linked to information about the macroeconomic aggregates as they could influence the discount rate or dividend stream, given that asset price is the sum of discounted future dividend stream (e.g., McQueen and Roley 1993; Flannery and Protopapadakis 2002; Shanken and Weinstein 2006).

[6]For example, Balvers et al. (1990) show that the general equilibrium framework with a logarithmic utility function and full capital depreciation can deliver a linear econometric model of stock return on the log of output.

[7]See Balvers et al. (1990) for a theoretical justification of the importance of output as a determinant of stock return. In the empirical study of Shanken and Weinstein (2006), industrial production is found to be an important determinant of stock return. In the empirical study of Shanken and Weinstein (2006), industrial production is found to be an important determinant of stock return among other factors, such as expected and unanticipated inflation, the spread in corporate bonds, and the spread in the treasury yields.

where v_t enters the quantile dependence equation; u_t, the innovation in y_t, is independent of v_t; and y_t is monotonic in u_t for each x_t and v_t.[8] The main parameter of interest is $\widetilde{\varphi}$ which captures the dependence between y and x. In order to allow this dependence parameter to be contingent on information about the US and the other financial market, $\widetilde{\varphi}$ in (67.2) is modeled as a function of u_t and v_t so that shocks to either the United States or the other financial market may influence the extent to which these markets are dependent. To obtain a parameter that provides information about the correlation between y and x, a copula model will be used to generate the function h. In the bivariate case, the copula is a function that combines the marginal distributions of x and y to yield their joint distribution function. Specifically, for variables x and y with marginal distributions F_x and F_y and joint distribution F, there exists a unique copula function C with copula parameter (with an abuse of notation) that satisfies

$$F(x,y) = C\left(F_x(x), F_y(y); \widetilde{\varphi}\right).$$

The copula that is employed in this study is the Gaussian copula, although other copulae such as the Student-t copula may be explored. Letting $\Phi_2(\cdot)$ denote the bivariate Gaussian distribution and $\Phi(\cdot)$ denote the standard normal distribution, the Gaussian copula expresses the joint distribution of x and y as

$$F(y_t, x_t) = \Phi_2\left(\Phi^{-1}\left(F_y(y_t)\right), \Phi^{-1}(F_x(x_t)); \widetilde{\varphi}\right), \tag{67.3}$$

and is especially useful as its parameter $\widetilde{\varphi}$ is the correlation coefficient. One important advantage in adopting the Gaussian copula is the feasibility of transforming it into a regression model that is amenable to the quantile regression technique. For instance, Bouyè and Salmon (2009) show that the τ_y *copula quantile curve* based on the Gaussian copula can be derived from (67.3) as

$$\begin{aligned} \tau_y &\equiv \frac{\partial \Phi_2\left(\Phi^{-1}\left(F_y(y_t)\right), \Phi^{-1}(F_x(x_t)); \widetilde{\varphi}\right)}{\partial F_x(x_t)} \\ &= \Phi\left(\frac{\Phi^{-1}\left(F_y(y_t)\right) - \widetilde{\varphi}\,\Phi^{-1}(F_x(x_t))}{\sqrt{1-\widetilde{\varphi}^2}}\right), \end{aligned} \tag{67.4}$$

where the first line follows by definition.[9] Rewriting (67.4), a regression model can be written as[10]

[8]Monotonicity ensures that conditional on x_t and v_t, the quantile of y can be mapped from the quantile of u.

[9]See p. 726 in Bouyè and Salmon (2009).

[10]When x_t is replaced by y_{t-1}, the model becomes an autoregression in $\Phi^{-1}(F_y(y_t))$, leading to the nonlinear copula quantile autoregression model of Chen et al. (2009).

$$\Phi^{-1}\left(F_y(y_t)\right) = \widetilde{\varphi}\,\Phi^{-1}(F_x(x_t)) + \sqrt{1-\widetilde{\varphi}^2}\Phi^{-1}\left(\tau_y\right). \tag{67.5}$$

If the marginal distributions F_x and F_y are standard normal, (67.5) simplifies further into an elegant regression model of

$$y_t = \widetilde{\varphi}x_t + \sqrt{1-\widetilde{\varphi}^2}\Phi^{-1}\left(\tau_y\right). \tag{67.6}$$

Adapting from (67.6), the quantile dependence equation that is considered in the application can be expressed as

$$y_t = \widetilde{\varphi}(u_t, v_t)x_t + \sqrt{1-\widetilde{\varphi}(u_t, v_t)^2}\Phi^{-1}\left(\tau_y\right), \tag{67.7}$$

where the correlation parameter is modeled as function of u_t and v_t to be consistent with (67.2).

That $\widetilde{\varphi}$ is allowed to be influenced by u_t and v_t is a crucial feature in the quantile dependence approach. First of all, if $\widetilde{\varphi}$ were a constant, the dependence parameter will not be affected by the quantile information of the financial markets. Second, the conditional quantiles of y and x are intertwined with the quantiles of u and v, respectively. By postulating $\widetilde{\varphi}$ as a function of u_t and v_t, we are relating the dependence between y and x given by $\widetilde{\varphi}$ to the quantile information on u and v, which is in turn linked to the quantile information on y and x. Therefore, allowing $\widetilde{\varphi}$ to be dependent on the quantile information of u and v enables us to capture the level of dependence that is specific to the quantile information of y and x.

For example, let us consider how the conditional quantile of the US return can be motivated from the auxiliary equation of (67.1). Equation 67.1 shows that holding the conditioning variables fixed, any extrinsic variation in the US return (x_t) must be attributed to v_t. In other words, the conditional quantile of the US return is linked to the quantile of v_t, so that the τ_x conditional quantile of the US return is given as

$$Q_x(\tau_x|\mathbf{z}_t) = \mathbf{b}^\top\mathbf{z}_t + {F_v}^{-1}(\tau_x), \tag{67.8}$$

where the distribution function of v be $F_v(\cdot)$ and its τ_x quantile is ${F_v}^{-1}(\tau_x)$. Conditioning on $\mathbf{z}$, (67.8) therefore illustrates how the conditional quantile of x, i.e., $Q_x(\tau_x|\mathbf{z}_t)$, is intertwined with the quantile of v. By modeling $\widetilde{\varphi}$ as a function of v, $\widetilde{\varphi}$ may then vary with the quantile information of v and hence of x. Likewise, by modeling $\widetilde{\varphi}$ as a function of u, $\widetilde{\varphi}$ may vary with the quantile information of y as the quantile information of u and y are related.

To express the concept of quantile dependence using the general quantile dependence equation of (67.2) for our discussion, recall from (67.8) that $Q_x(\tau_x|\mathbf{z}_t)$ is linked with ${F_v}^{-1}(\tau_x)$. By conditioning y_t on $Q_x(\tau_x|\mathbf{z}_t)$, logical consistency requires fixing v_t at ${F_v}^{-1}(\tau_x)$ in (67.2) as well. This is because as we have seen in (67.8), v_t cannot vary freely in the construction of $Q_x(\tau_x|\mathbf{z}_t)$. With this in mind, the τ_y quantile of the Australian return conditioning on $Q_x(\tau_x|\mathbf{z}_t)$ can be expressed as

$$\begin{aligned} Q_y\left(\tau_y|Q_x(\tau_x|\mathbf{z}_t)\right) &= h\left(Q_x(\tau_x|\mathbf{z}_t);\widetilde{\varphi}\left(F_u^{-1}\left(\tau_y\right),F_v^{-1}(\tau_x)\right)\right) \\ &\equiv h\left(Q_x(\tau_x|\mathbf{z}_t);\varphi\left(\tau_y,\tau_x\right)\right), \end{aligned} \tag{67.9}$$

where I have defined $\varphi\left(\tau_y,\tau_x\right) \equiv \widetilde{\varphi}\left(F_u^{-1}\left(\tau_y\right),F_v^{-1}(\tau_x)\right)$. Through the dependence parameter $\varphi(\tau_y, \tau_x)$, (67.9) summarizes how the τ_y Australian return quantile is related to the τ_x US return quantile. If the dependence function h in (67.9) is replaced with the copula-based model in (67.7), then (67.9) expresses the *copula quantile-on-quantile regression* (C-QQR) model that is to be estimated in this chapter.

Before we proceed, it should be emphasized why the US return is chosen for the auxiliary model over the other market returns. In formulating (67.1) and (67.2), two assumptions are made. First, in relation to the auxiliary equation of (67.1), which is modeled as an autoregression of x_t, I assume that US economic fundamentals are sufficient for determining the US return beyond the economic fundamentals of the other markets. Hence, information from the other markets would not matter for driving the US return and are excluded from (67.1), which is likely too reasonable for relatively smaller economies such as Australia, Hong Kong, and Singapore and perhaps less so for Japan. By including the United States in the quantile dependence equation of (67.2), the second assumption asserts that the US return information is important for influencing the other stock markets. This would be plausible if US fundamentals contribute towards the global economic forces that drive the co-movement between the United States and the other financial markets. Therefore, if information about US fundamentals has global content, and if this is subsumed in the US return, the US return will be a powerful variable for explaining the variation in the market returns of the Australia, Hong Kong, Japan, and Singapore. This motivates placing the US return, not the other market returns, as a right-hand-side variable in the quantile dependence equation.

Furthermore, it should also be emphasized that (67.1) and (67.2) form a triangular system of simultaneous equation of the type analyzed by Ma and Koenker (2006), who study the dependence between conditional quantiles that is motivated from such a structure. However, there is a fundamental difference between this paper and Ma and Koenker (2006). While Ma and Koenker (2006) study a parametric model, this chapter does not make a parametric assumption about the function $\widetilde{\varphi}(u_t, v_t)$ in order that the data is allowed to speak with respect to the response of $\widetilde{\varphi}$ to the quantile information of u and v. Therefore, this requires an alternative method of estimation from Ma and Koenker (2006) which is discussed in the next section.

67.2.2 Estimation

This section outlines the procedure for estimating the dependence parameter in the C-QQR model, i.e., $\varphi\left(\tau_y,\tau_x\right) \equiv \widetilde{\varphi}\left(F_u^{-1}\left(\tau_y\right),F_v^{-1}(\tau_x)\right)$ in (67.9). In the context of the C-QQR model where the h function in (67.9) is replaced by the copula-based model in (67.7), $\varphi(\tau_y, \tau_x)$ expresses the correlation between τ_y conditional quantile of y and

τ_x conditional quantile of x. Even though the value of $\varphi(\tau_y, \tau_x)$ is based upon the $\widetilde{\varphi}(u_t, v_t)$ parameter in the quantile dependence equation of (67.2),[11] the estimate of $\varphi(\tau_y, \tau_x)$ cannot be obtained by a straightforward application of quantile regression on (67.2) as it is a single equation with two unobservable terms u_t and v_t. In order to estimate $\varphi(\tau_y,\tau_x)$, or equivalently $\widetilde{\varphi}\left(F_u^{-1}(\tau_y), F_v^{-1}(\tau_x)\right)$, I first anchor v_t at $F_v{}^{-1}(\tau_x)$ in the quantile dependence equation while letting u_t be "free" and then implement a τ_y -quantile regression on this resulting equation to set u_t to $F_u^{-1}(\tau_y)$ for an estimate of $\widetilde{\varphi}\left(F_u^{-1}(\tau_y), F_v^{-1}(\tau_x)\right)$.

To elaborate, let us decompose the quantile dependence equation of (67.2) into two parts:

$$y_t = h\left(x_t; \widetilde{\varphi}\left(u_t, F_v^{-1}(\tau_x)\right)\right) + \underbrace{h(x_t; \widetilde{\varphi}(u_t, v_t)) - h\left(x_t; \widetilde{\varphi}\left(u_t, F_v^{-1}(\tau_x)\right)\right)}_{\Omega}. \tag{67.10}$$

In the first part of (67.10), the v-argument in $\widetilde{\varphi}$ is anchored at $F_\upsilon{}^{-1}(\tau_x)$ so that $h\left(x_t; \widetilde{\varphi}\left(u_t, F_v^{-1}(\tau_x)\right)\right)$ is now a function of a single unobservable, u_t. If the Ω portion of (67.10) can be controlled, our target parameter $\varphi(\tau_y,\tau_x)$ can be estimated by implementing a τ_y-quantile regression on (67.10) as doing so would, in principle, deliver an estimate of the conditional function of $h\left(x_t; \widetilde{\varphi}\left(F_u^{-1}(\tau_y), F_v^{-1}(\tau_x)\right)\right)$ which contains our target. To control for Ω, I approximate it by a first-order Taylor expansion of v_t around $F_v^{-1}(\tau_x)$ as

$$\Omega \approx h_{\widetilde{\varphi}}\left(x_t; \widetilde{\varphi}\left(u_t, F_\upsilon^{-1}(\tau_x)\right)\right)\widetilde{\varphi}_v\left(u_t, F_\upsilon^{-1}(\tau_x)\right)\upsilon_t(\tau_x), \tag{67.11}$$

where (67.11) uses the definition $v_t(\tau_x) = v_t - F_v^{-1}(\tau_x)$. The function $h_{\widetilde{\varphi}}$ is the partial derivative of h with respect to $\widetilde{\varphi}$, which is a known expression given that h is specified in (67.7). The parameter $\widetilde{\varphi}_v\left(u_t, F_v^{-1}(\tau_x)\right)$ is the partial derivative of $\widetilde{\varphi}$ with respect to v_t, where its v-argument is evaluated at $F_v{}^{-1}(\tau_x)$. The functional form of this partial derivative is unknown as the functional form of $\widetilde{\varphi}\left(u_t, F_v^{-1}(\tau_x)\right)$ is not specified. With the first-order Taylor approximation leading to (67.11), the initial problem of controlling for v_t in the quantile dependence equation now becomes an issue of controlling for $v_t(\tau_x)$ in (67.11). The new variable $v_t(\tau_x)$ can be estimated as the residual following a τ_x-quantile regression on the auxiliary model of (67.1).[12] Letting $\hat{v}_t(\tau_x)$ denote the estimate of $v_t(\tau_x)$, we can control for Ω using its feasible counterpart:

$$\hat{\Omega} = h_{\widetilde{\varphi}}\left(x_t; \widetilde{\varphi}\left(u_t, F_v^{-1}(\tau_x)\right)\right)\widetilde{\varphi}_v\left(u_t, F_v^{-1}(\tau_x)\right)\hat{v}_t(\tau_x),$$

so that the quantile dependence equation to be taken to the data is

[11] $\varphi(\tau_y, \tau_x)$ can be motivated from $\widetilde{\varphi}(u_t, v_t)$ by anchoring the u , v-arguments in $\widetilde{\varphi}(u_t, v_t)$ at $F_u{}^{-1}(\tau_y)$ and $F_v{}^{-1}(\tau_x)$.

[12] Since $Q_x(\tau_x|\mathbf{z}_t) = \mathbf{b}^\top\mathbf{z}_t + F_v{}^{-1}(\tau_x)$, we may express the auxiliary regression of (67.1), i.e., $x_t = \mathbf{b}^\top\mathbf{z}_t + v_t$, as $x_t = \mathbf{b}^\top\mathbf{z}_t + F_v{}^{-1}(\tau_x) + v_t - F_v{}^{-1}(\tau_x) = Q_x(\tau_x|\mathbf{z}_t) + v_t - F_v{}^{-1}(\tau_x) = Q_x(\tau_x|\mathbf{z}_t) + v_t(\tau_x)$. Therefore, we may estimate $\upsilon_t(\tau_x)$ as the residual from a τ_x -quantile regression on (67.1).

$$
\begin{aligned}
y_t = {} & h\left(x_t; \widetilde{\varphi}\left(u_t, F_v^{-1}(\tau_x)\right)\right) \\
& + h_{\widetilde{\varphi}}\left(x_t; \widetilde{\varphi}\left(u_t, F_v^{-1}(\tau_x)\right)\right) \widetilde{\varphi}_v\left(u_t, F_v^{-1}(\tau_x)\right) \hat{v}_t(\tau_x).
\end{aligned} \tag{67.12}
$$

To a first-order approximation, u_t is the only unobservable term in (67.12). Therefore, $\varphi(\tau_y, \tau_x)$ can now be estimated by implementing a τ_y-quantile regression on (67.12). As compared to the original quantile dependence equation of (67.10), the "revised" quantile dependence equation of (67.12) now contains an additional parameter, $\widetilde{\varphi}_v$, which is to be estimated together with the main parameter of interest, $\widetilde{\varphi}$.

Let $\rho_\tau(\cdot)$ be the "check" function in Koenker and Bassett (1978), defined as $\rho_\tau(u) = u(\tau - \mathrm{I}(u < 0))$, where $\mathrm{I}(\cdot)$ is an indicator function. Using (67.12), the dependence between the quantiles of x and y can be estimated by implementing a two-step quantile regression procedure:

1. *Obtain residuals $\hat{v}_t(\tau_x)$ from a τ_x -quantile regression on* (67.1), *the auxiliary equation*, i.e.,

$$
\min_{\mathbf{b}} \sum_{t=1}^{T} \rho_\tau\left(x_t - \mathbf{b}^\top \mathbf{z}_t\right)
$$

2. *Using $\hat{v}_t(\tau_x)$, estimate $\widetilde{\varphi}$ from a τ_y -quantile regression on* (67.12), *the quantile dependence equation*, i.e.,

$$
\min_{(\widetilde{\varphi}, \widetilde{\varphi}_v)} \sum_{t=1}^{T} \rho_\tau\left(y_t - h(x_t; \widetilde{\varphi}) - h_{\widetilde{\varphi}}(x_t; \widetilde{\varphi}) \widetilde{\varphi}_v \hat{v}_t(\tau_x)\right).
$$

Step 1 is a standard linear quantile regression and Step 2 is a standard nonlinear quantile regression. The second-step estimate of $\widetilde{\varphi}$ will yield the desired estimate of $\varphi(\tau_y, \tau_x)$. Because the C-QQR approach involves a two-step quantile regression procedure, it can be implemented using statistical packages for quantile regression such as the quantreg package of Koenker (2009) within the R software.[13]

67.3 Application

Monthly returns of Australia, Hong Kong, Japan, Singapore, and the United States are constructed from the Datastream-MSCI indices.[14] Table 67.1 provides the summary statistics for the period between March 1974 and February 2010. Among Australia, Hong Kong, Japan, and Singapore, the US market is most strongly correlated with the Singapore market (at 0.60) and most weakly correlated with the Japan market (at 0.41). But these are sample correlation coefficients that could differ significantly from actual levels of correlation especially when equity markets are bearish. By considering how

[13]For instance, the Steps 1 and 2 regressions can be implemented using the rq and nlrq commands of the quantreg package in R.

[14]The monthly returns are constructed as 100 multiplied by the change in the log of the index.

Table 67.1 Summary statistics on monthly stock returns (March 1974 to February 2010)

	Mean	Min	Max	Standard deviation	Correlation with United States
Australia	0.67	−44.79	20.11	5.80	0.58
Hong Kong	1.22	−62.50	36.36	9.19	0.49
Japan	0.39	−24.38	17.51	5.32	0.41
Singapore	0.71	−44.55	44.92	7.70	0.60

the quantiles of these market returns are correlated, the C-QQR approach offers a way of examining their level of dependence that is contingent on specific market conditions.

In this application, the correlation between the τ_x conditional quantile of the US return and the τ_y conditional quantile of the return of Australia, Japan, Hong Kong, or Singapore is computed. The values of τ_y and τ_x are defined on the grid [0.1, 0.2, ..., 0.9], where the total number of τ_y and τ_x combinations on this grid is 81. The C-QQR approach is implemented to estimate the correlation between the 10th–90th percentiles of the US return and the 10th–90th percentiles of the other market return (in decile intervals), for a total of 81 different correlation estimates for each return pair.

To visualize the C-QQR correlation, Fig. 67.1 plots the C-QQR correlation surfaces by relating the level of correlation in the *z*-axis to the return quantiles of Australia, Hong Kong, Japan, or Singapore on the *x*-axis and the return quantiles of the United States on the *y*-axis. The correlation surfaces offer some important insights. First, they show that the level of dependence varies substantially across the distributions of returns. Second, across all the four return pairs, the C-QQR correlations share many common features that echo our existing understanding about how correlations behave under various circumstances. For instance, the C-QQR correlation has the tendency to be weak at the center of the return distributions, implying that markets are less dependent when they are neither bearish nor bullish. As a side remark, while the correlation between centrally located quantiles may be interpreted as the level of dependence when markets are neither bearish or bullish, measures such as the sample correlation coefficient may not deliver the same interpretation. In fact, Table 67.2 shows that the sample correlation coefficient appears to be different from the correlation between centrally located return quantiles such as the median returns, where the correlation between the median returns is *always* smaller than the sample correlation coefficient. This perhaps is not surprising as the sample correlation coefficient is computed without cleaving out the consequence of extreme events that lead to inflating the actual level of dependence when they happen.

Of particular relevance to the practice of risk management is the fact that stock markets are more strongly dependent when they are bearish. Through the C-QQR approach, a similar point is made in terms of the stronger dependence between return quantiles in the left tail distributions. This is demonstrated in Fig. 67.2, which plots the correlation along the main diagonals in Fig. 67.1, i.e., the correlation along $\tau_y = \tau_x$. Figure 67.2 reveals that the correlation usually peaks at around $\tau_y = \tau_x = 0.1$ (the 10th percentiles of returns). Because these lower return quantiles are associated with markets that are bearish, Fig. 67.2 reiterates a familiar result in the existing literature that the correlation between bear markets would be stronger than usual (e.g., Longin and Solnik 2001; Hu 2006; Chollete et al. 2011).

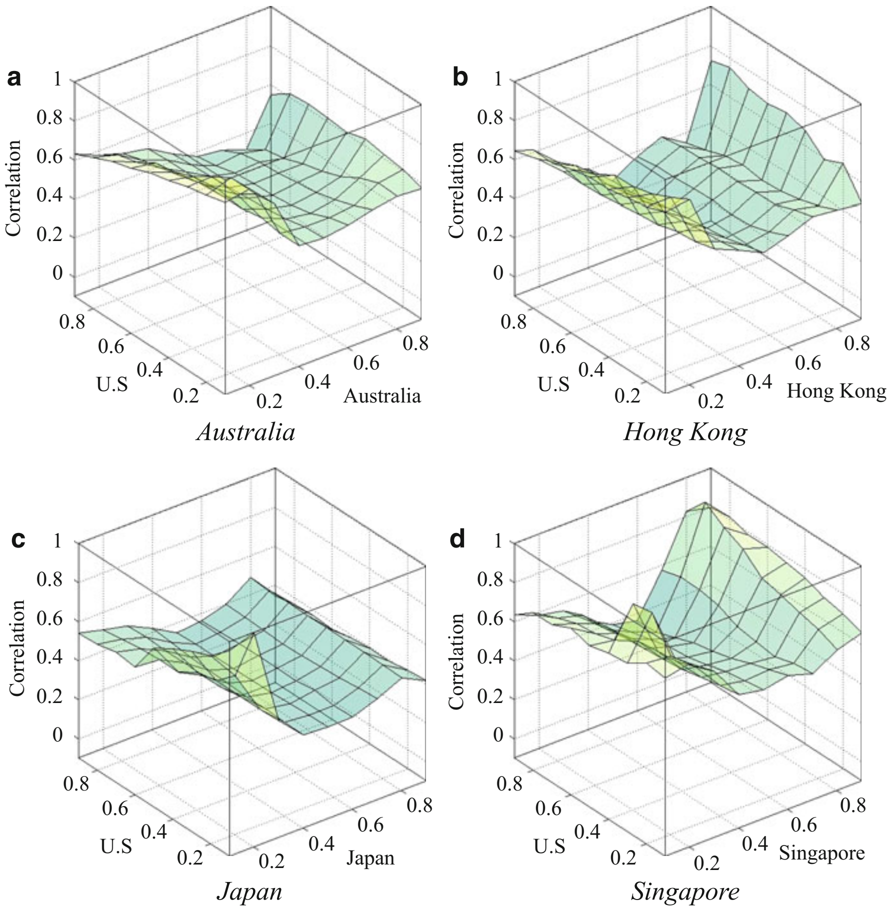


Fig. 67.1 C-QQR correlation
This figure plots the correlation between the quantile of the US market return and the quantile of the market return of Australia, Hong Kong, Japan, or Singapore using the C-QQR approach. The x-axis marks the quantiles of the US return and the y-axis marks the quantile of the other return

Note that the C-QQR approach is useful for showing the key locations in the return distributions where most of the movements in correlation take place. Specifically, it offers a new insight that the correlation may deviate sharply even at less extreme return quantiles. In the case of the US-Australia and US-Singapore return pairs, Fig. 67.2 shows that the correlation rises substantially starting from the 30th return percentiles. If 30th return percentiles are associated with markets that are mildly bearish, this implies that markets do not have to be severely bearish in order to trigger a nontrivial increase in correlation.

To evaluate the "performance" of the C-QQR approach informally, it is useful to compare the average correlation based on the C-QQR approach with the sample

Table 67.2 Sample and C-QQR correlations

	Australia	Hong Kong	Japan	Singapore
Panel A				
Sample	0.5795	0.4867	0.4123	0.5987
Average C-QQR	0.5697	0.5194	0.4657	0.6318
Panel B				
$\tau_y = \tau_x = 0.1$	0.8856	0.8358	0.9541	0.8132
$\tau_y = \tau_x = 0.5$	0.4625	0.3583	0.3552	0.4757
$\tau_y = \tau_x = 0.9$	0.5480	0.7212	0.3439	0.7976

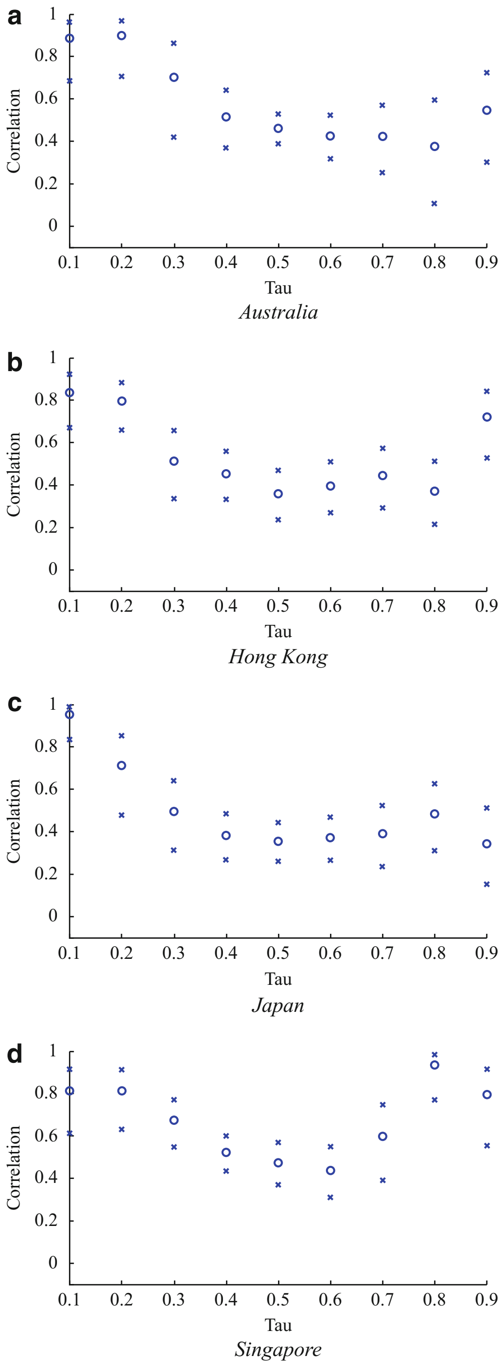
Sample reports the sample correlation coefficient. *C-QQR* reports the average of the 81 correlation estimates, where each estimate $\hat{\varphi}(\tau_y, \tau_x)$ is specific for $\boldsymbol{\tau_y}$ and $\boldsymbol{\tau_x}$ defined on the grid [0.1, . . ., 0.9]. $\tau_y = \tau_x = k$ reports the correlation between the k quantile of the US market return and k quantile of the market return of Australia, Hong Kong, Japan, or Singapore, i.e., $\hat{\varphi}(k, k)$

correlation coefficient, which itself reflects the average level of dependence. This "average C-QQR correlation" can be obtained by averaging up the 81 C-QQR correlation estimates for each return pair. Panel A of Table 67.2 compares the average C-QQR correlation with the sample correlation coefficient and shows that the two measures of average dependence are quite similar. For example, looking at the dependence between the US-Australia return pair, the average C-QQR correlation is 0.57 (rounded to the nearest two decimal places), which is very close the sample correlation of 0.58. However, the closeness between the average C-QQR correlation and the sample correlation coefficient is not specific to the US-Australia correlation. For instance, in the case of the US-Hong Kong and US-Singapore return pairs, their average C-QQR correlations are 0.52 and 0.63, respectively, which are very close to their sample correlation coefficients of 0.49 and 0.60. Even in the largest case of disparity between the two correlation measures, which is found for US-Japan return pair, the difference between the average C-QQR correlation and the sample correlation is only about 0.05. Given that the average C-QQR correlation delivers a reasonable measure of the average level of dependence as benchmarked by the sample correlation coefficient, one may interpret the C-QQR approach as a technique for decomposing the level of average dependence into levels that are specific to different points in the distribution of returns and thus to a wide spectrum of market conditions.

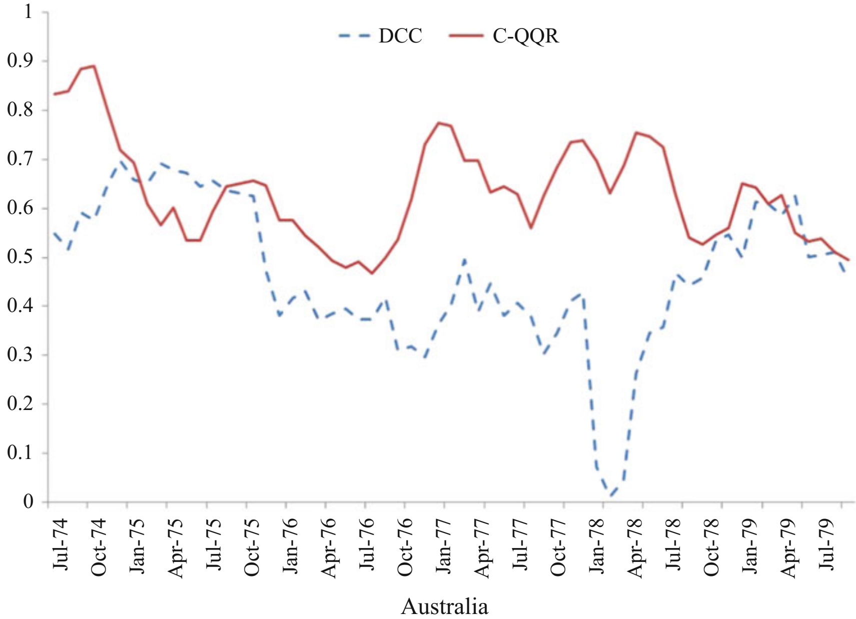
67.3.1 Dynamic C-QQR Correlation

Having obtained the C-QQR correlation estimates, it is straightforward to construct a historical series of correlation. This construction is especially useful for shedding light on the behavior of correlation across time, and in this regard, the C-QQR approach is related to the celebrated Dynamic Conditional Correlation (DCC) framework of Engle (2002) that is designed for the study of the time series behavior of correlation. Therefore, another informal evaluation of the C-QQR approach is to compare its estimates directly with the DCC. This comparison is interesting as the two approaches are completely unrelated – the C-QQR approach is based

Fig. 67.2 C-QQR correlation, $\tau_y = \tau_x$
For $\tau_y = \tau_x$, this figure plots the correlation between the τ_y US return quantile and the τ_x quantile of the market return of Australia, Hong Kong, Japan, or Singapore using the C-QQR approach. The 95 % bootstrap confidence band is provided



a



b

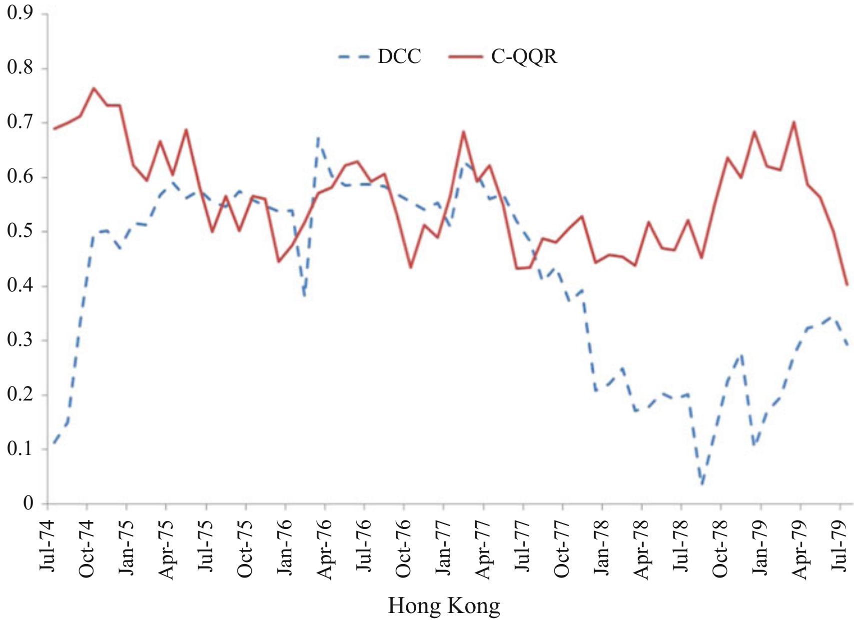


Fig. 67.3 (continued)

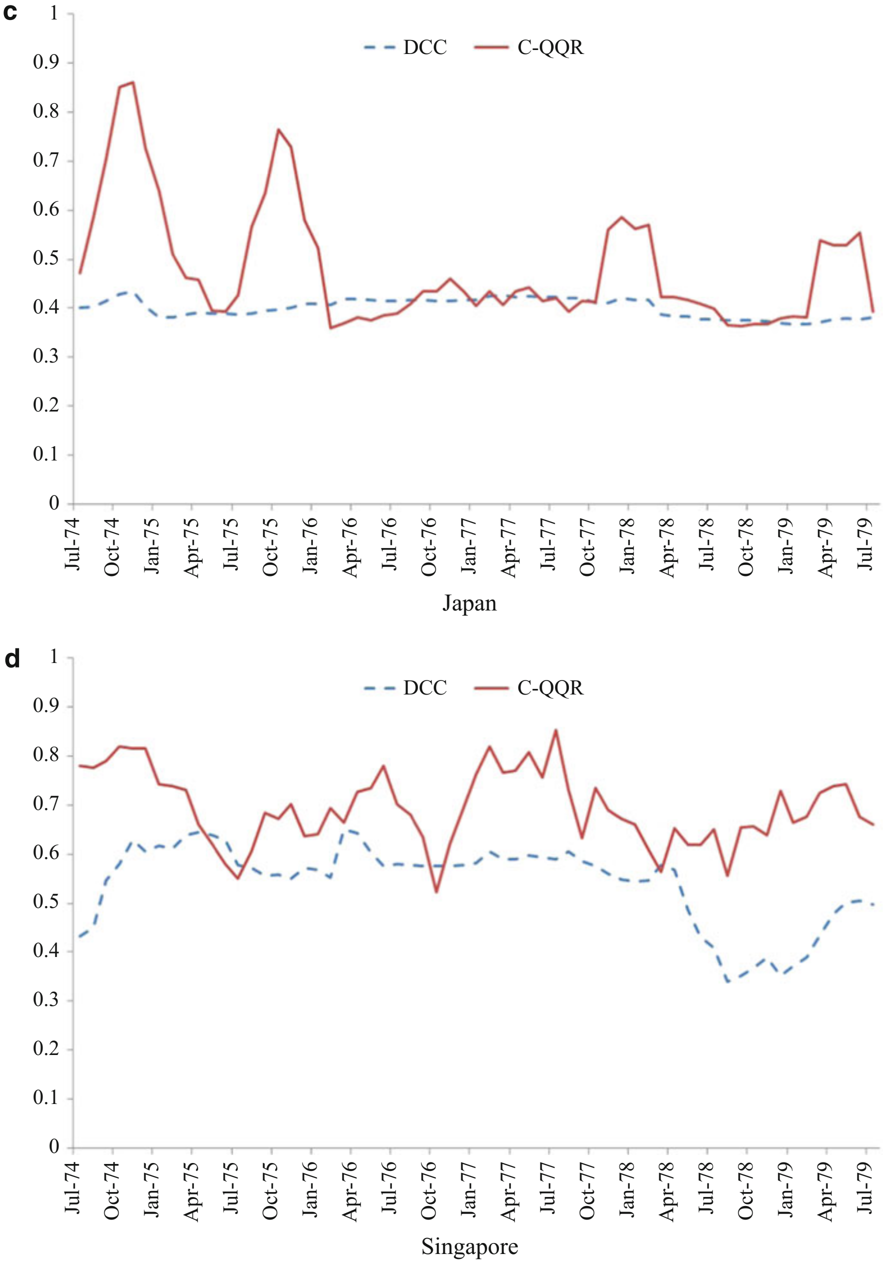


Fig. 67.3 Dynamic C-QQR correlation and DCC in levels, 1974–1979
The C-QQR correlation and the DCC are plotted for the correlation of the US market return with the market return of Australia, Hong Kong, Japan, or Singapore. The *solid line* corresponds to the 4-month moving average C-QQR correlation and the *dotted line* corresponds to the DCC that is estimated using the DCC-GARCH(1,1) specification

on quantile regressions and the DCC approach is based on the GARCH framework. As it turns out, despite the clear difference in the two modeling approaches, the C-QQR approach produces estimates that in some ways are visually comparable to the DCC.

The time series of correlation can be constructed by first matching each realized stock return at period t to the nearest return quantile that is used when estimating the C-QQR correlation. Once that is done for the US return and the return of the other market, we may then match a C-QQR correlation estimate (from the pool of 81 estimates) to these quantiles to obtain an approximate realized correlation at period t. By employing this matching approach for each period, we may construct an approximation of the historical correlation. Calling this the "dynamic C-QQR correlation," the detailed procedure of computing it is outlined as follows:

1. Compute the empirical distributions for the returns to the US market (x) and to the market returns of Australia, Hong Kong, Japan, or Singapore (y) for each period t. Denote them by $\widetilde{F}_x(x_t)$ and $\widetilde{F}_y(y_t)$.
2. If $\widetilde{F}_x(x_t)$ or $\widetilde{F}_y(y_t)$ is less than 0.05, add 0.05. If $\widetilde{F}_x(x_t)$ or $\widetilde{F}_y(y_t)$ is greater than 0.95, subtract 0.05. Call the new series $\widetilde{F}_x(1)(x_t)$ and $\widetilde{F}_y(1)(y_t)$.
3. Round $\widetilde{F}_x(1)(x_t)$ and $\widetilde{F}_y(1)(y_t)$ to the nearest first decimal place. The new series $\widetilde{F}_x(2)(x_t)$ and $\widetilde{F}_y(2)(y_t)$ will be on the grid [0.1, . . ., 0.9].
4. The C-QQR correlation estimates is a 9×9 matrix, where each point on the matrix corresponds to a combination of points on two [0.1, . . ., 0.9] grids, with each grid representing the return percentiles of the US market and the other market, respectively. The correlation at time t is obtained by matching $\widetilde{F}_x(2)(x_t)$ and $\widetilde{F}_y(2)(y_t)$ to the C-QQR correlation matrix of estimates.

For a close-up comparison, Fig. 67.4 plots the dynamic C-QQR correlation (4-month moving average) and the DCC for the first 5-year period in the sample from 1973 to 1979, and Fig. 67.5 plots these correlations for the last 5-year period in the sample from 2005 to 2010. Choosing the first and last 5-year periods allows us to observe how the dynamic C-QQR correlation and DCC compare across the two most distant 5-year periods in the sample. Besides comparing their levels, it is also useful to compare their first difference as doing so would help us to gain further insights on how the dynamic C-QQR correlation behaves relative to the DCC.

Focusing on the 1973–1979 period, Fig. 67.4 shows that the dynamic C-QQR correlation is similar to the DCC in terms of movements, although not necessarily in terms of magnitude. For instance, comparing the C-QQR correlation and the DCC for the US-Australia return pairs, Fig. 67.1 shows a decline in the C-QQR correlation from around July 1974 to April 1975, while the DCC manifests a similar downward motion from January to October 1975. Likewise, the C-QQR correlation trends upwards from April 1976 to January 1977 with the DCC following suit from October 1976 to around the same time. Focusing on their first difference, Fig. 67.5 shows that the peaks in the first difference of the dynamic C-QQR correlation are

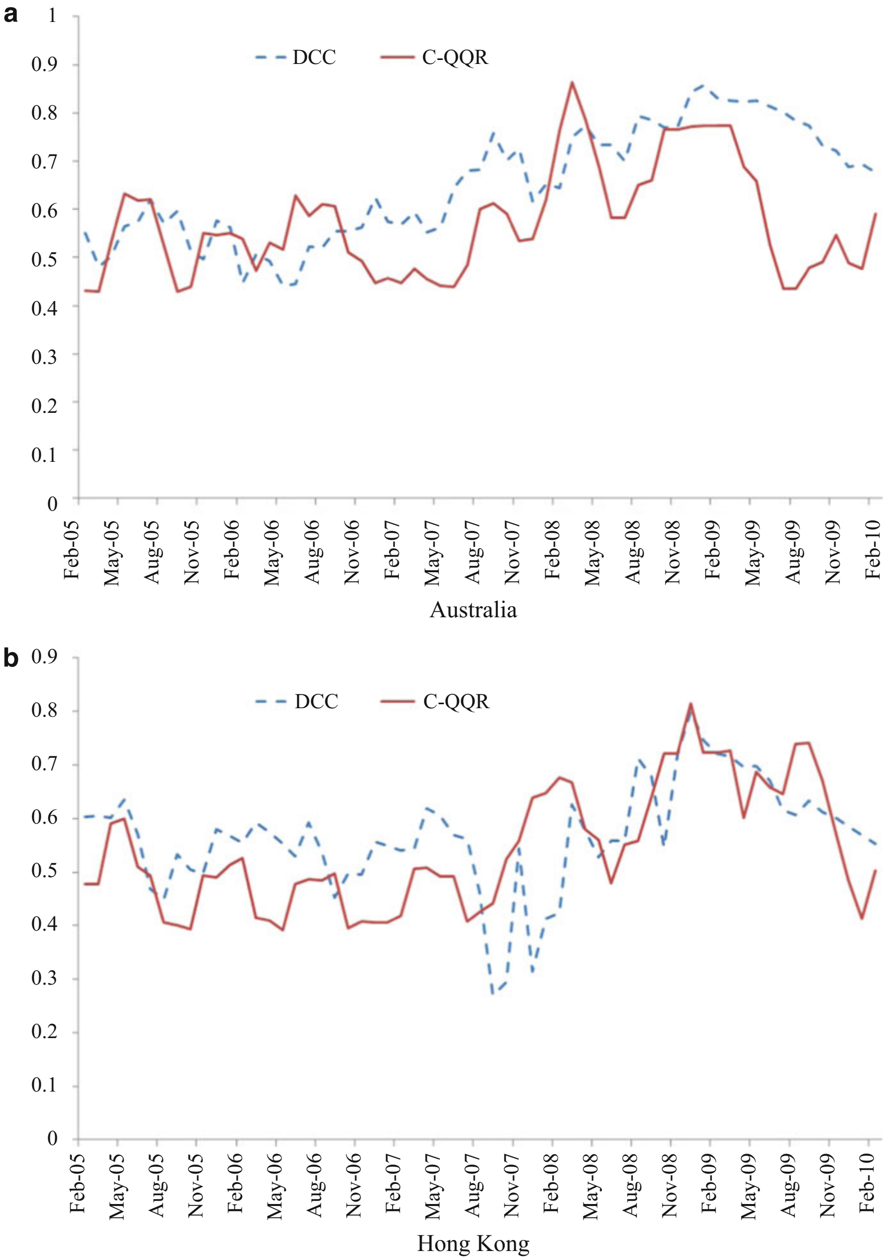
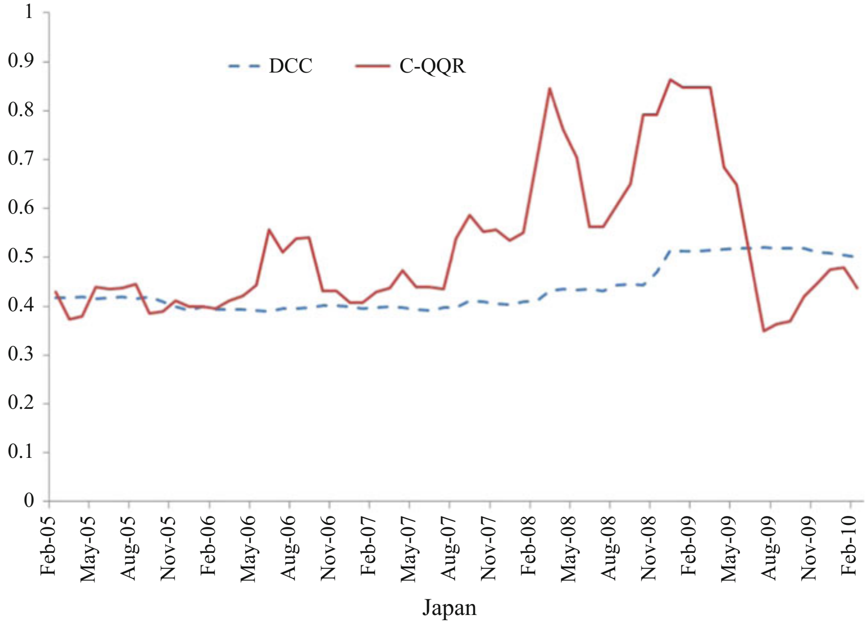


Fig. 67.4 (continued)

c



d

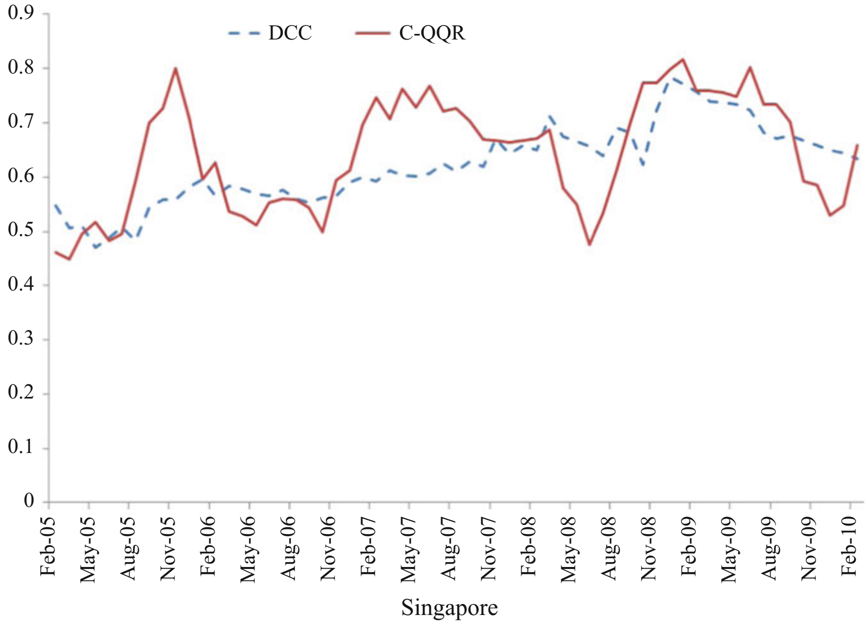


Fig. 67.4 Dynamic C-QQR correlation and DCC in levels, 2005–2010
The C-QQR correlation and the DCC are plotted for the correlation of the US market return with the market return of Australia, Hong Kong, Japan, or Singapore. The *solid line* corresponds to the 4-month moving average C-QQR correlation and the *dotted line* corresponds to the DCC that is estimated using the DCC-GARCH(1,1) specification

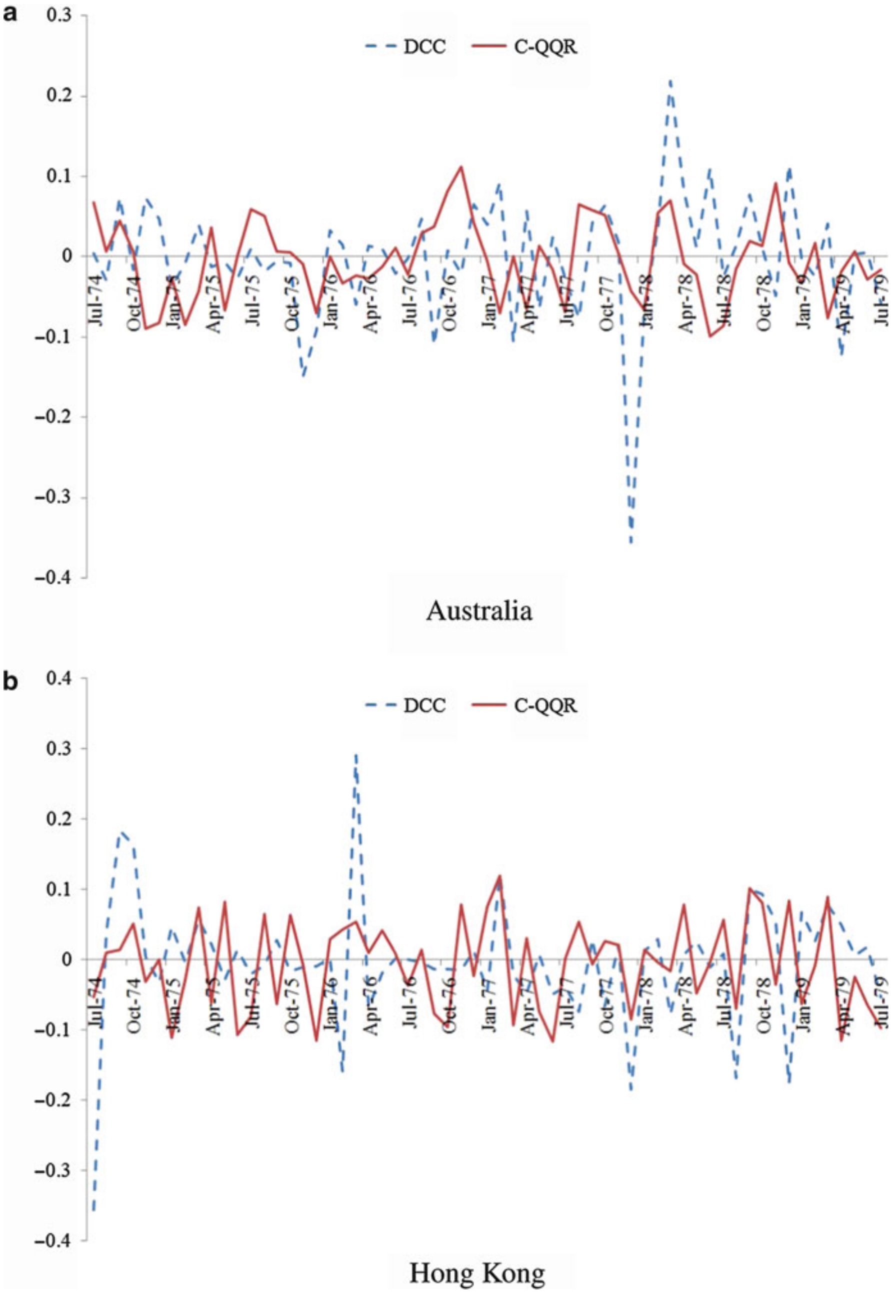
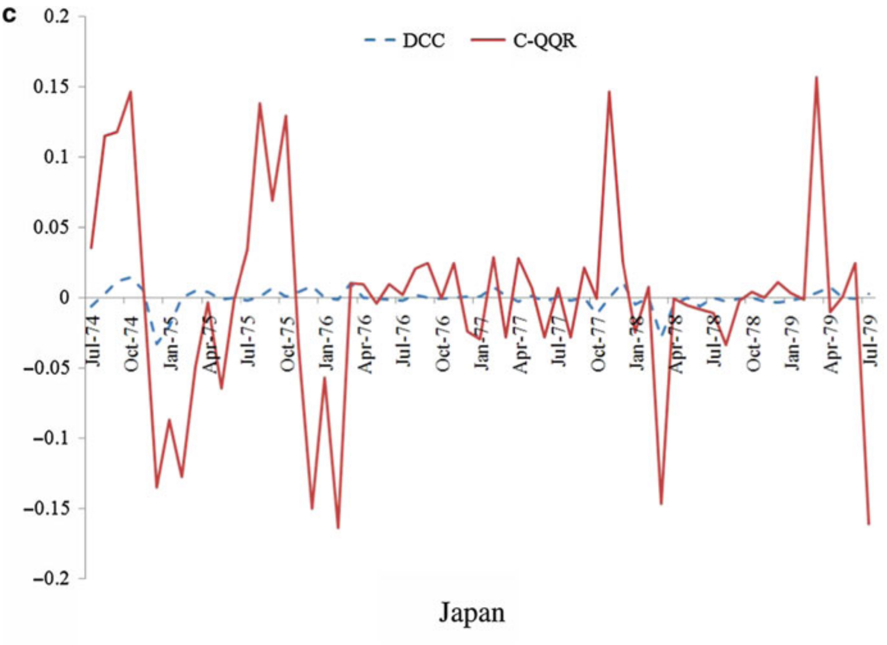


Fig. 67.5 (continued)



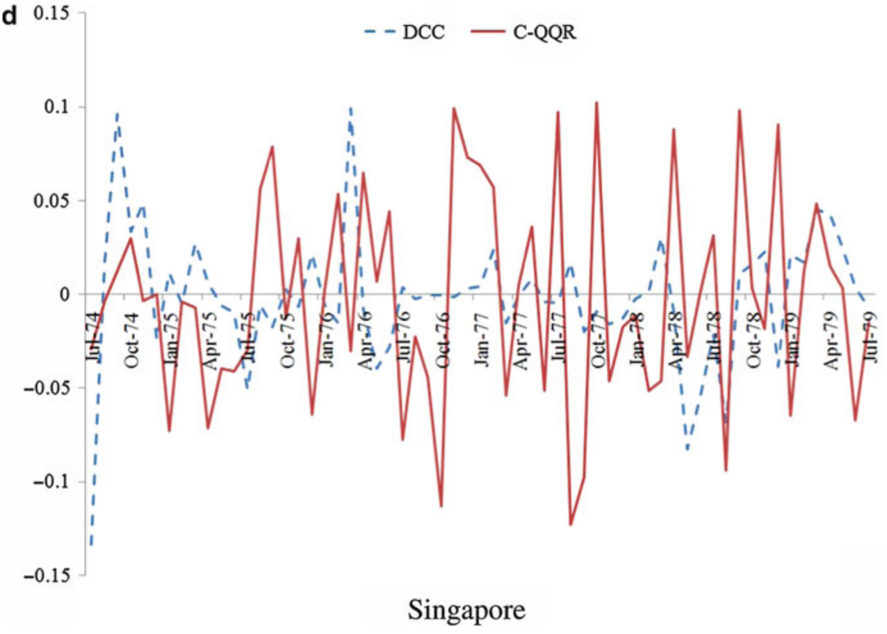


Fig. 67.5 Dynamic C-QQR correlation and DCC in first difference, 1974–1979
The C-QQR correlation and the DCC, in first difference, are plotted for the correlation of the US market return with the market return of Australia, Hong Kong, Japan, or Singapore. The *solid line* corresponds to the 4-month moving average C-QQR correlation and the *dotted line* corresponds to the DCC that is estimated using the DCC-GARCH(1,1) specification

followed by similar peaks in the first difference of the DCC in many occasions, although the size of these changes across the C-QQR correlation and the DCC are somewhat different. Besides the US-Australia return pairs, the C-QQR correlation and the DCC display noticeable similarities with respect to the US-Hong Kong and US-Singapore return pairs, but least resemble each other in the case of US-Japan.

During the 2005–2010 period, not only do the dynamic C-QQR correlation and the DCC display similar trends, they also appear to be moving along the same path. For instance, when focusing on the US-Australia and US-Hong Kong return pairs, Fig. 67.4 shows that the C-QQR correlation and the DCC track each other closely with similar levels. In terms of their first difference, Fig. 67.6 also shows that while changes in the C-QQR correlation and the DCC occur nearly in tandem, the changes in the DCC are usually preceded by changes in the C-QQR correlation in the same direction. For example, in the case of the US-Australia return pairs, Fig. 67.6 shows that the peaks in the first difference of the C-QQR correlation around April 2005, November 2005, and June 2006 are followed by peaks in the first difference of the DCC about a month or two later.

The US-Japan return pair presents an interesting case in the comparison between the dynamic C-QQR correlation and the DCC. For example, I find that the DCC in this context to be more or less steady throughout the sample period, including the 1973–1979 and 2005–2010 periods that saw the 1974–1975 US and global recession triggered by the tripling of the price of oil and the current global financial crisis that started in 2007. Interestingly, while the C-QQR correlation typically meanders around a steady level throughout the sample period, it is also characterized by sharp upward movements during 1974–1975 and starting from the end of November 2007, the periods when global financial markets are bearish. And, during the times when the C-QQR correlation is free from these large deviations, it is nearly identical to the DCC. This can be seen by comparing the two correlation series during April 1976 to August 1977 and April 1978 to March 1979 in Fig. 67.3, and prior to February 2006 in Fig. 67.4, where the two correlation estimates nearly coincide.

While this exercise does not offer a statistical evaluation of the closeness between the dynamic C-QQR correlation and the DCC, visual inspection of the two correlation series reveals some common features between them especially during 2005–2010. That there are some similarities between the dynamic C-QQR correlation and the DCC is somewhat surprising since from the outset, the C-QQR and the DCC approaches based on completely different modeling paradigms.

67.4 Conclusion

This chapter discusses a new perspective of modeling correlation, based on the C-QQR approach, which focuses on the correlation between the conditional quantiles of asset returns as a way of uncovering the level of dependence for specific market conditions. The C-QQR approach has the ability to replicate key features about the correlation between stock returns that have been noted before. For instance, it shows that the correlation between lower return quantiles is stronger than that between

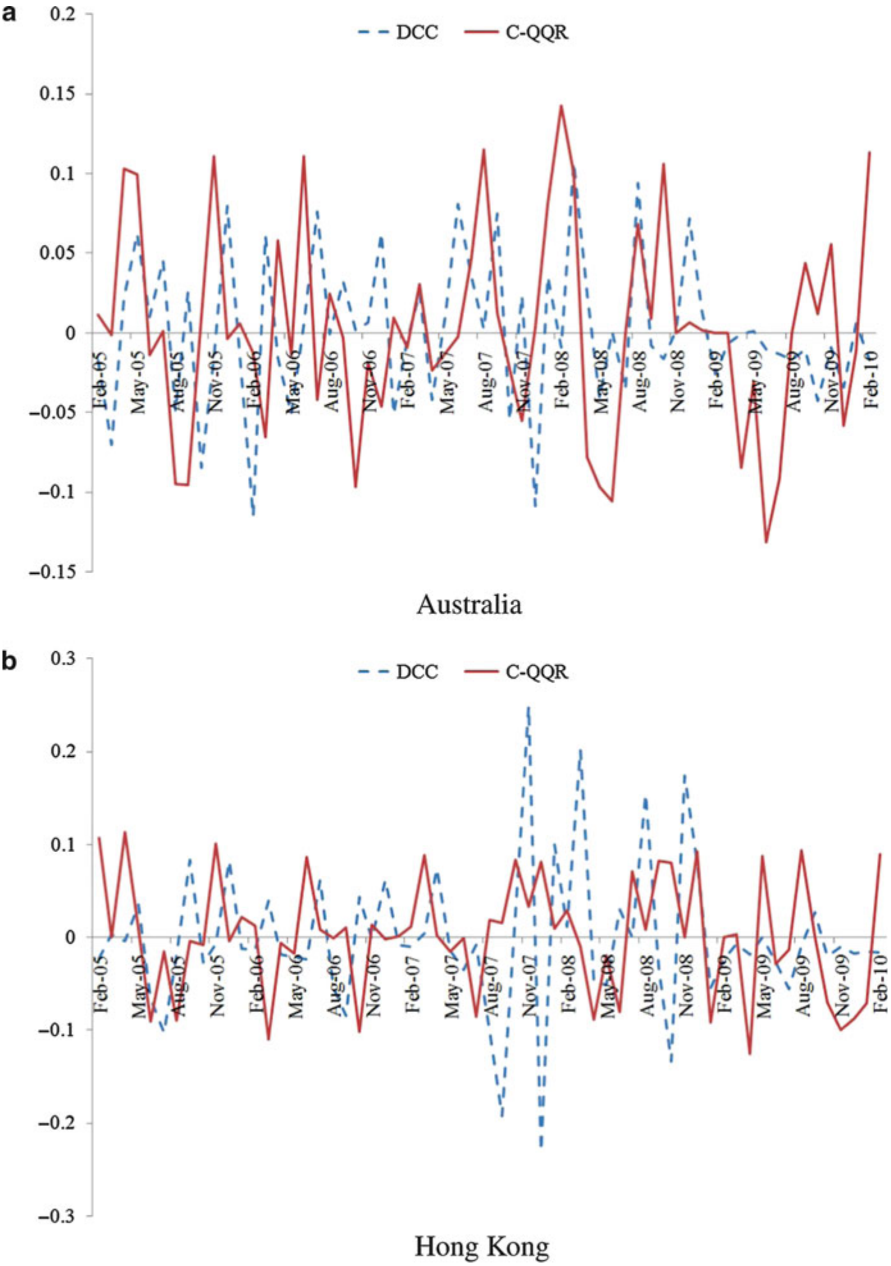


Fig. 67.6 (continued)

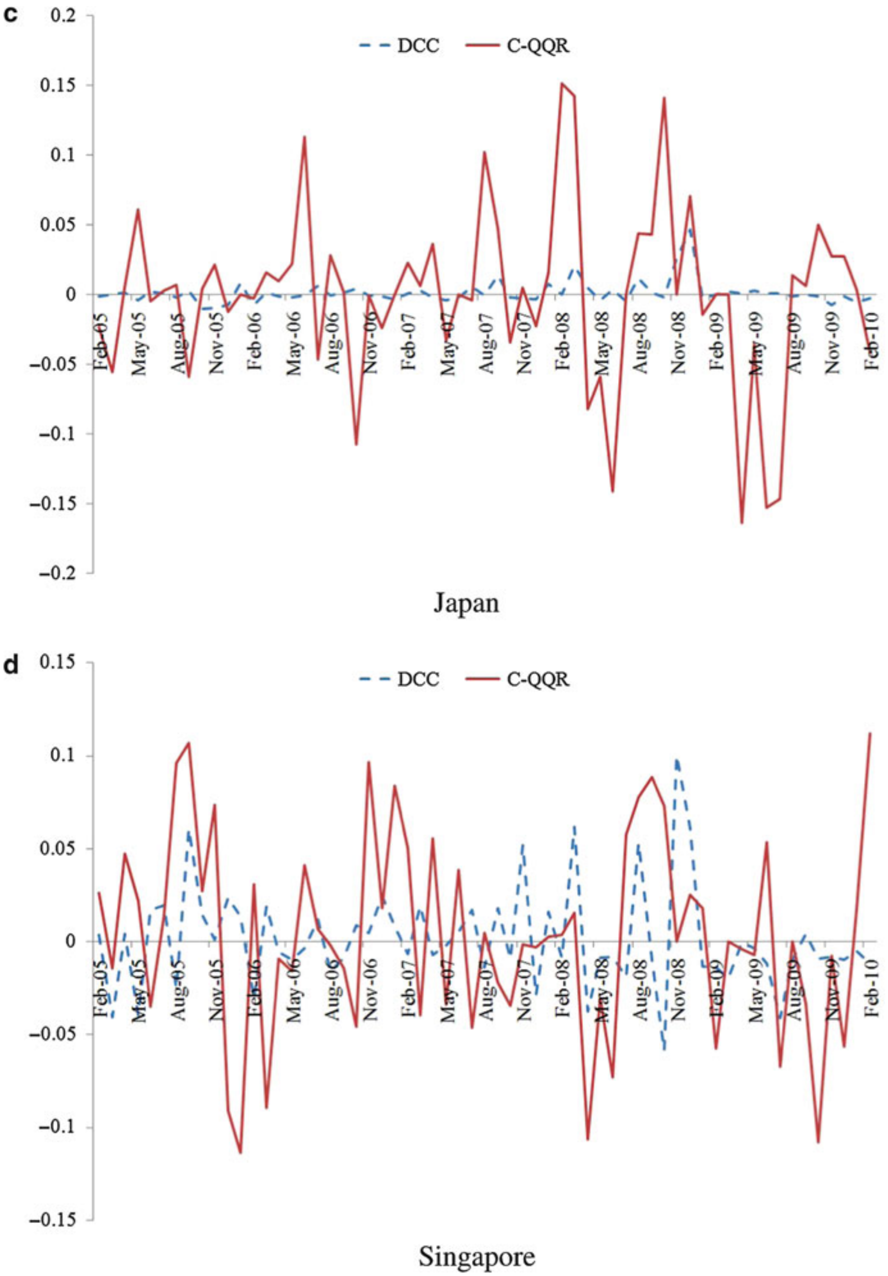


Fig. 67.6 Dynamic C-QQR correlation and DCC in first difference, 2005–2010
The C-QQR correlation and the DCC, in first difference, are plotted for the correlation of the US market return with the market return of Australia, Hong Kong, Japan, or Singapore. The *solid line* corresponds to the 4-month moving average C-QQR correlation and the *dotted line* corresponds to the DCC that is estimated using the DCC-GARCH(1,1) specification

centrally located return quantiles, which corroborates the familiar observation that bear markets are more strongly correlated. The C-QQR approach also has the ability to produce a constructed time series of correlation that resembles the DCC even though the C-QQR and DCC approaches are completely unrelated. Given that the C-QQR framework produces correlation estimates that match the findings of existing non-dynamic-based approaches of modeling correlation and the estimates of the dynamic based DCC approach, it therefore empirically bridges the gap between the dynamic and non-dynamic-based paradigms of modeling correlation.

Nevertheless, the application of the concept of quantile dependence in financial economics is still in the early stage where some issues could be addressed going forward. Firstly, I presented a bivariate version of the C-QQR model for the analysis of pairwise correlation. As financial markets are interrelated, an extension to the multivariate case would be an important direction. Secondly, I use the auxiliary equation to model the US return, and it would be interesting to investigate the implications on the final correlation estimate of doing the opposite, that is, use the auxiliary equation to model the market returns of Australia, Hong Kong, Japan, and Singapore. Finally, in terms of applications, it should be emphasized that the relevance of the C-QQR approach is not confined to the study of equities alone. For instance, one could also look at the correlation between stocks and bonds through the lens of the C-QQR approach and examine "flight to quality" hypothesis, which emphasizes the tendency of investors to underweight equities in favor of bonds in the face of market uncertainties (e.g., Connolly et al. 2005). It would also be interesting to apply the C-QQR approach to examine issues in macroeconomics such as studying the nonlinearity in the relationships between macroeconomic aggregates, which has been a topic of considerable interest in recent research.

References

Ang, A., & Bekaert, G. (2002). International asset allocation with regime shifts. *Review of Financial Studies, 15*, 1137–1187.

Ang, A., & Chen, J. (2002). Asymmetric correlations of equity portfolios. *Journal of Financial Economics, 63*, 443–494.

Balvers, R. J., Cosimano, T. F., & McDonald, B. (1990). Predicting stock returns in an efficient market. *Journal of Finance, 45*, 1109–1128.

Bouyè, E., & Salmon, M. (2009). Dynamic copula quantile regressions and tail area dynamic dependence in Forex markets. *European Journal of Finance, 15*, 721–750.

Chen, N. F., Roll, R. R., & Ross, S. A. (1986). Economic forces and the stock market. *Journal of Business, 59*, 383–403.

Chen, X., Koenker, R., & Xiao, Z. (2009). Copula-based nonlinear quantile regression autoregression. *Econometrics Journal, 12*, S50–S67.

Chollete, L., de la Penã, V., & Lu, C. (2011). International diversification: A copula approach. *Journal of Banking and Finance, 35*, 403–417.

Connolly, R. A., Stivers, C. T., & Sun, L. (2005). Stock market uncertainty and the stock-bond return relation. *Journal of Financial and Quantitative Analysis, 40*, 161–194.

Engle, R. F. (2002). Dynamic conditional correlation: A simple class of multivariate generalized autoregressive conditional heteroskedasticity models. *Journal of Business and Economic Statistics, 20*, 339–350.

Erb, C. B., Harvey, C. R., & Viskanta, T. E. (1994). Forecasting international equity correlations. *Financial Analysts Journal, 50*, 32–45.

Flannery, M. J., & Protopapadakis, A. A. (2002). Macroeconomic factors do influence aggregate stock returns. *Review of Financial Studies, 15*, 751–782.

Guidolin, M., & Timmermann, A. (2005). Economic implications of bull and bear regimes in UK stock and bond returns. *Economic Journal, 115*, 111–143.

Heffernan, J. E., & Tawn, J. A. (2004). A conditional approach for multivariate extreme values. *Journal of the Royal Statistical Society B, 66*, 497–546.

Hu, L. (2006). Dependence patterns across financial markets: A mixed copula approach. *Applied Financial Economics, 16*, 717–729.

Koenker, R. (2009). Quantile regression in R: A vignette. Mimeo.

Koenker, R., & Bassett, G. (1978). Regression quantiles. *Econometrica, 46*, 33–50.

Longin, F., & Solnik, B. (1995). Is the correlation in international equity returns constant: 1960–1990? *Journal of International Money and Finance, 14*, 3–26.

Longin, F., & Solnik, B. (2001). Extreme correlations in international equity markets. *Journal of Finance, 56*, 649–676.

Ma, L., & Koenker, R. (2006). Quantile regression methods for recursive structural equation models. *Journal of Econometrics, 134*, 471–506.

McQueen, G., & Roley, V. V. (1993). Stock prices, news, and business conditions. *Review of Financial Studies, 6*, 683–707.

Nelsen, R. B. (2006). *An introduction to copulas* (2nd ed.). New York: Springer.

Okimoto, T. (2008). New evidence of asymmetric dependence structures in international equity markets. *Journal of Financial and Quantitative Analysis, 43*, 787–815.

Patton, A. J. (2006). Modelling asymmetric exchange rate dependence. *International Economic Review, 47*, 527–555.

Poon, S. H., Rockinger, M., & Tawn, J. (2004). Extreme value dependence in financial markets: Diagnostics, models and financial implications. *Review of Financial Studies, 17*, 581–610.

Shanken, J., & Weinstein, M. I. (2006). Economic forces and the stock market revisited. *Journal of Empirical Finance, 13*, 129–144.

Sim, N. (2012). Modeling the correlations of asset returns by the dependence between their quantiles. Working paper.

Trivedi, P. K., & Zimmer, D. A. (2007). *Copula modeling: An introduction for practitioners*. Hanover: Now Publishers.

Multi-criteria Decision Making for Evaluating Mutual Funds Investment Strategies

68

Shin Yun Wang and Cheng-Few Lee

Contents

S.Y. Wang (✉)
National Dong Hwa University, Shou-Feng, Hualien, Taiwan
e-mail: gracew@mail.ndhu.edu.tw

C.-F. Lee
Department of Finance and Economics, Rutgers Business School, Rutgers, The State University of New Jersey, Piscataway, NJ, USA

Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
e-mail: cflee@business.rutgers.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_68,

Abstract

Investors often need to evaluate the investment strategies in terms of numerical values based upon various criteria when making investment. This situation can be regarded as a multiple criteria decision-making (MCDM) problem. This approach is oftentimes the basic assumption in applying hierarchical system for evaluating the strategies of selecting the investment style. We employ the criteria measurements to evaluate investment style. To achieve this objective, first, we employ factor analysis to extract independent common factors from those criteria. Second, we construct the evaluation frame using hierarchical system composed of the above common factors with evaluation criteria and then derive the relative weights with respect to the considered criteria. Third, the synthetic utility value corresponding to each investment style is aggregated by the weights with performance values. Finally, we compare with empirical data and find that the model of MCDM predicts the rate of return.

Keywords

Investment strategies • Multiple criteria decision making (MCDM) • Hierarchical system • Investment style • Factor analysis • Synthetic utility value • Performance values

68.1 Introduction

The number of mutual funds has increased exceeding the number of stocks listed on the organized exchange, hence making the selection of mutual funds an onerous task for the investor. In addition, the mutual funds are moving rapidly towards financial market development in response to increasing market demand and the mutual fund industry. Therefore the mutual funds have huge market potential and have been gaining momentum in the financial market. The complexities are numerous, and overcoming these complexities to offer successful selections is a mutual fund manager's challenge.

The mutual fund managers need to evaluate aquatic return so as to reduce its risk to find the optimal combination of invested stocks out of many feasible stocks and distribute the amount of investing funds to many stocks. Because of the limited amount of funds invested into mutual funds, the solution of the portfolio selection problem proposed by Markowitz (1952) has a tendency to increase the number of stocks selected for mutual funds. In a real investment, a fund manager first makes a decision on how much proportion of the investment should go to the market, and then he invests the fund to which stocks which is the stock selection ability. After that, many researchers explained in the presence of market-timing ability that actions will affect the performance of mutual funds. When investing mutual funds, some reports also point out that there are 90 % of investors who will consider the rate of return firstly and then the reputation of mutual fund corporation and investment risk. Maximizing the mutual fund performance is the primary goal of mutual fund managers in a corporation. Usually, the mutual fund return reflects the financial performance of a fund corporation for operating and development. This study explores which criteria can lead to high mutual fund performance.

To achieve this purpose, we use the method of multi-criteria decision making (MCDM). MCDM is one of the most widely used decision methodologies in engineering, medicine, economics, law, the environment, and public policy and business. The theory, methodology, and practice of MCDM have experienced a revolutionary process during the last five decades. MCDM methods aim at improving the quality of decisions by making the process more explicit, rational, and efficient. One intriguing problem is that oftentimes, different methods may yield different answers to the same decision problem. Thus, the issue of evaluating the relative performance of different MCDM methods is raised. One evaluating procedure is to examine the stability of an MCDM method's mathematical process by checking the validity of its proposed rankings. However, within a dynamic and diversified decision-making environment, the traditional quantitative method does not solve the non-quantity problems of investment selection. Therefore, what is needed is a useful and applicable strategy that addresses the issues of investment selection. We thus propose a MCDM method to evaluate the hierarchy system for selecting investment strategies.

In this study the hierarchical analytic approach is used to determine the weights of criteria from subjective judgment, and a nonadditive integral technique is utilized to evaluate the performance of investment style. Traditionally, researchers have used additive techniques to evaluate the synthetic performance of each criterion. The rest of the chapter is organized as follows: Mutual fund literature is discussed in the next section. The method of MCDM including the hierarchical analytic approach and non- additive integral evaluation process for MCDM problems is derived in Sect. 68.3. Then an illustrative example is presented in Sect. 68.4, which applies the MCDM method of investment. After which we discuss and show how the MCDM methods in this chapter are effective in Sect. 68.5. Finally, the conclusions are presented in Sect. 68.6.

68.2 Review of Mutual Fund Investment

Mutual fund research abounds in finance literature, and the investment performance of mutual fund managers has been extensively examined. Most of these studies employ a method developed by Jensen (1968, 1969) and later refined by Black et al. (1972) and Blume and Friend (1973). Such a method compares a particular manager's performance with that of a benchmark index fund. Connor and Korajczyk (1986) develop a method of portfolio performance measurement using a competitive version of the arbitrage pricing theory (APT). However, they ignore any potential market timing by managers. One weakness of the above approach is that it fails to separate the aggressiveness of a fund manager from the quality of the information he/she possesses. It is apparent that superior performance of a mutual fund manager occurs because of his/her ability to “time” the market and the ability to forecast the returns on individual assets.

Jensen (1968) demonstrates that the presence of market-timing ability is an important factor in mutual fund selections. Grant (1977) explains how market-timing actions will affect the results of empirical tests that focus only on microforecasting skills. Fama (1972) indicates that there are two ways for fund managers to obtain abnormal returns. The first one is security analysis, which is the ability of fund managers to identify the potential winning securities. The second one is market timing, which is the ability of portfolio managers to time market cycles and takes advantage of this ability in trading securities. Treynor and Mazuy (1966) add a quadratic term to the Jensen function to test for market-timing ability. Chen and Stockum (1986) employ a generalized varying parameter model, which treats Treynor and Mazuy (1966) as a special case, to study the mutual fund’s stock selectivity and market-timing ability. They find mutual funds as a group exhibits some evidence of stock selection ability yet no market-timing ability. Jensen (1972) develops theoretical structures for the evaluation of micro- and macroforecasting performance of fund managers where the basis for evaluation is a comparison of the ex post performance of the fund manager with the returns on the market. Merton (1981) and Henriksson’s (1984) model differs from Jensen’s formulation in that their forecasters follow a more qualitative approach to market timing. Chang and Lewellen (1984) and Henriksson (1984) employ the Merton-Henriksson model in evaluating mutual fund performance and find no evidence of market timing by fund managers. Bhattacharya and Pfleiderer (1983) extend the work of Jensen (1972). By correcting an error made in Jensen, they show that one can use a simple regression technique to obtain accurate measures of timing and selection ability.

Lehmann and Modest (1987) combine the APT performance evaluation method with the Treynor and Mazuy (1966) quadratic regression technique. They found statically significant abnormal timing and selectivity performance by mutual funds. They also examine the impact of alternative benchmarks on the performance of mutual funds and find that performance measures are quite sensitive to the benchmark chosen. Also, Henriksson (1984) finds a negative correlation between the measures of stock selection and market-timing ability. Finally, Lee and Rahman (1990) also empirically examine market timing and selectivity performance of mutual funds. Furthermore, Jorge et al. (2006) deal with the relevance of benchmark choice for mutual fund performance behavior, and Spitz (1970) researches the relationship between mutual fund performance and cash inflows. Blake and Morey (2000) verify the mutual fund performance of Morningstar ratings.

The above mentioned studies concentrate on a fund manager’s security selection and market-timing skills. However, external evaluation, human judgment, and subjective perception also affect the performance of mutual funds. In a real-world setting, the performance of mutual funds involves many criteria. In this article we will discuss these criteria and performance at the same time.

68.3 The Method of Multi-criteria Decision Making

Kuosmanen (2004) and Kopa and Post (2009) use the stochastic dominance criterion test on a portfolio optimality and efficient diversification. In this section we employ factor analysis to extract four independent common factors from those criteria. At the same time we construct the evaluation frame using AHP (analytic hierarchy process), which is composed of the above four common factors with sixteen evaluated criteria, and derive the relative weights with respect to the considered criteria and the synthetic utility value corresponding to each mutual fund investment style.

According to the literature review and our questionnaire survey, we employ factor analysis to extract independent common factors from criteria. At the same time we construct the evaluation framework using a hierarchical system composed of the above common factors with evaluation criteria and derive the relative weights pertinent to the considered criteria. Then the synthetic utility value corresponding to each investment style is aggregated by the weights with performance values. Traditional analytic hierarchy process (AHP) assumes that there is no interaction between any two criteria within the same hierarchy. However, in reality, a criterion is inevitably correlated with another one. In this section, we give a brief to some notions from the theory of measure and integral. We describe a hierarchical analytic approach to determine the weighting of subjective judgments.

68.3.1 λ-Measure

The specification for general measures requires the values $\cap$ of a measure for all subsets in X. Let (X, β, g) be a measure space: $\lambda \in (-1, \infty)$. If A $\in \beta$, B $\in \beta$; and A $\cap$ B$=\phi$, and

$$g(A \cup B) = g(A) + g(B) + \lambda g(A)g(B) \tag{68.1}$$

If this holds, then measure g is λ-additive. This kind of measure is named λ-measure, or the Sugeno measure. In this chapter we denote this λ-measure by g_i to differentiate from other measures. Based on the axioms above, the λ-measure of the finite set can be derived from densities, as indicated in the following equation:

$$g_\lambda(\{x_1, x_2\}) = g_1 + g_2 + \lambda g_1 g_2 \tag{68.2}$$

Where $g_{1,}$ g_2 represents the density.

Let set $X = \{x_1, x_2, .., x_n\}$ and the density of measure $g_i = g_\lambda(\{x_i\})$, which can be formulated as follows:

$$g_\lambda(\{x_1, x_2, \ldots, x_n\}) = \sum_{i=1}^{n} g_i + \lambda \sum_{i_1=1}^{n-1} \sum_{i_2=i_1+1}^{n} g_{i_1} g_{i_2} + \cdots + \lambda^{n-1} g_1 g_2 \cdots g_n \tag{68.3}$$

For an evaluation case with two criteria, A and B, there are three cases based on the above properties:

Case 1: If $\lambda > 0$, i.e., $g_\lambda(A \cup B) > g_\lambda(A) + g_\lambda(B)$, implying that A and B have a multiplicative effect.

Case 2: If $\lambda = 0$, i.e., $g_\lambda(A \cup B) = g_\lambda(A) + g_\lambda(B)$, implying that A and B have an additive effect.

Case 3: If $\lambda < 0$, i.e., $g_\lambda(A \cup B) < g_\lambda(A) + g_\lambda(B)$, implying that A and B have a substitutive effect.

The measure is often used with the integral for aggregating information evaluation by considering the influence of the substitutive and multiplication effect among all criteria.

68.3.2 The Integral

In a measure space (X, β, g), let h be a measurable set function defined in the measurable space. Then the definition of the integral of h over A with respect to g is

$$\int_A h(x)dg = \sup_{\alpha \in [0,1]} [\alpha \wedge g(A \cap H_\alpha)] \tag{68.4}$$

where $H_a = \{$x belonging to Xlh(x) $\geq \alpha\}$. A is the domain of the integral. When A = X, then A can be taken out.

Next, the integral calculation is described in the following. For the sake of simplification, consider a measure g of X, $\aleph$ where X is a finite set. Let $h : X \to [0,1]$ and assume without loss of generality that the function $h\ (x_j)$ is monotonically decreasing with respect to j, i.e., $h(x_1) \geq h(x_2) \geq \cdots \geq h(x_n)$. To achieve this, the elements in X can be renumbered. With this, we then have

$$\int h(x)dg = \mathop{\vee}_{i=1}^{n} [f(x_i) \wedge g(X_i)] \tag{68.5}$$

where $X_i = \{x_1, x_2, \cdots, x_i\}$, i = 1, 2, $\cdots$, n.

In practice, h is the evaluated performance on a particular criterion for the alternatives, and g represents the weight of each criterion. The integral of h with respect to g gives the overall evaluation of the alternative. In addition, we can use the same measure using Choquet's integral, defined as follows:

$$\int hdg = h(x_n)g(X_n) + [h(x_{n-1}) - h(x_n)]g(X_{n-1}) + \cdots + [h(x_1) - h(x_2)]g(X_1) \tag{68.6}$$

The integral model can be used in a nonlinear situation since it does not need to assume the independence of each criterion.

68.3.3 The Integral Multi-criteria Assessment Methodology

The integral is used in this study to combine assessments primarily because this model does not need to assume independence among the criteria. A brief overview of the integral is presented here.

Assume under general conditions, $h(x_1^k) \geq \cdots \geq h(x_i^k) \geq \cdots \geq h(x_n^k)$ where $h(x_i^k)$ is the performance value of the kth alternative for the ith criterion, the integral of the measure $g_\lambda(X_n^k)$ with respect to $h(x_n^k)$ on $\aleph$ (g: $\aleph \longrightarrow [0, 1]$) can be defined as follows:

$$\int^k hdg = h\left(x_n^k\right) g_\lambda\left(X_n^k\right) + \left[h\left(x_{n-1}^k\right) - h\left(x_n^k\right)\right] g_\lambda\left(X_{n-1}^k\right) + \cdots + \left[h\left(x_1^k\right) - h\left(x_2^k\right)\right] g_\lambda\left(X_1^k\right) \tag{68.7}$$

where $g_\lambda(X_1^k) = g_\lambda(\{x_1^k\})$, $g_\lambda(X_2^k) = g_\lambda(\{x_1^k, x_2^k\})$, …, $g_\lambda(X_n^k) = g_\lambda(\{x_1^k, x_2^k, \cdots, x_n^k\})$

The measure of each individual criterion group $g_\lambda(X_n^k)$ can be expressed $\sum_{i=1}^{n} g_\lambda\left(x_i^k\right) + \lambda \sum \sum g_\lambda(\{x_i\}) g_\lambda\left(\{x_j\}\right) + \cdots \lambda^{n-1} g_\lambda(\{x_1\}) \cdots g_\lambda(\{x_n\})$ as follows:

$$\begin{aligned} g_\lambda\left(X_n^k\right) &= g_\lambda\left(\left\{x_1^k, x_2^k \cdots x_n^k\right\}\right) = \sum_{i=1}^{n} g_\lambda\left(x_i^k\right) + \\ &\lambda \sum \sum g_\lambda(\{x_\mathrm{i}\}) g_\lambda\left(\{x_j\}\right) + \cdots \\ &\lambda^{n-1} g_\lambda(\{x_1\}) \cdots g_\lambda(\{x_n\}) = \frac{1}{\lambda}\left[\prod_{i=1}^{n}\left(1 + \lambda g_\lambda\left(x_i^k\right)\right) - 1\right] \end{aligned} \tag{68.8}$$

for $-1 < \lambda < +\infty$.

λ is the parameter that indicates the relationship among related criteria (if λ=0, Eq. 68.7 is an additive form; if $\lambda \neq 0$, Eq. 68.7 is a nonadditive form).

68.4 Evaluation Model for Prioritizing the Investment Strategy

We build up a hierarchical system for evaluating investment strategies of Wang and Lee (2011). Its analytical procedures stem from three steps: (i) factor, (ii) criteria, and (iii) investment style. We employ factor analysis to extract four independent common factors from various criteria, and these factors are (1) market timing, (2) stock selection ability, (3) fund size, and (4) teamwork. We construct the evaluation frame using hierarchical system composed of the above four common factors with sixteen evaluated criteria. We then derive the relative weights pertinent to the considered criteria. According to the risk of investment, mutual funds with different investment styles are classified as S1, asset allocation style; S2, aggressive growth style; S3, equity income style; S4, growth style; and S5, growth income style. Based on the review of literature, personal experience, and interviews with senior mutual fund managers, relevance trees are used to create hierarchical strategies for developing the optimal selection strategy of mutual funds.

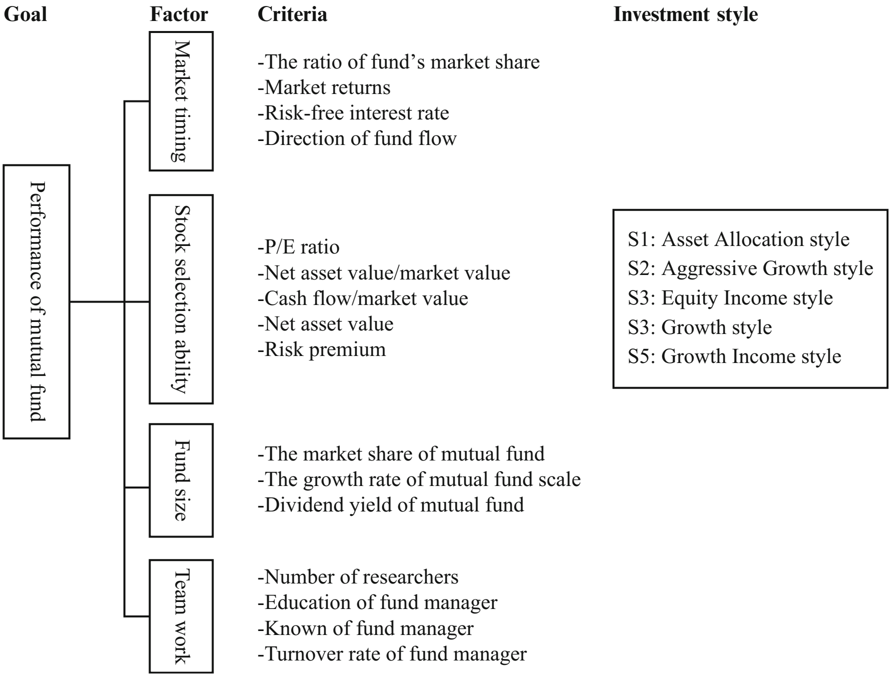


Fig. 68.1 Relevance system of hierarchy tree for evaluating mutual fund strategy

The elements (nodes) of relevance trees are defined and identified in hierarchical strategies, the combination of which consists of an evaluating mechanism for selecting a mutual fund strategy, as shown in Fig. 68.1.

68.4.1 Evaluating the Mutual Fund Strategy Hierarchy System

Minimum risk or maximum return is usually used as the measurement index in traditional financial evaluation methods. Based on the risk of investment, mutual funds are classified into five investment styles and we evaluate the funds' performance by the rate of return. Within a dynamic and diversified decision-making environment, the traditional quantitative method does not solve the non-quantity problems of mutual fund selection. Therefore, what is needed is a useful and applicable strategy that addresses the issues of selecting mutual funds. We propose an MCDM method to evaluate the hierarchy system for selecting mutual fund strategies.

The performance of mutual fund architecture includes four components: market timing, stock selection ability, fund size, and teamwork. We first discuss conceptual and econometric issues associated with identifying four components of mutual fund performance. We have chosen multiple criteria evaluation method for selecting and prioritizing the mutual fund strategies to optimize the real scenarios faced by managers or investors.

68.4.2 The Process for Evaluating and Prioritizing Mutual Fund Strategies

In this study, we use this MCDM method to evaluate various mutual fund strategies and rank them by performance. The following subsection describes the method of MCDM.

68.4.2.1 The Weights for the Hierarchy Process

An evaluator always perceives the weight of a hierarchy subjectively. Therefore, consider the uncertain, interactive effects coming from other criteria when calculating the weight of a specified criterion.

The weights w_j corresponding to each criterion is as follows:

$$w_j = r_j \otimes (r_1 \oplus \cdots \oplus r_m)^{-1} \tag{68.9}$$

where r_j is the geometric mean of each row of AHP reciprocal matrix

$$r_j = \left(a_{j1} \otimes \cdots \otimes a_{jm}\right)^{1/m} \tag{68.10}$$

68.4.2.2 The Synthetic Decision

The weight of the different criteria and the performance value needs to be operated using integral techniques to generate the synthetic performance of each strategy within the same dimension.

Furthermore, we have calculated the synthetic performance of each alternative strategy using different λ values. Additionally, the synthetic performance is conducted by a simple additive weight method assuming the criteria are independent in an environment. Since each individual criterion is not completely independent from the others, we use the nonadditive integral technique to find the synthetic performance of each alternative and to investigate the order of the synthetic performance of different λ values.

68.5 Empirical Examinations and Discussions

To demonstrate the practicality of our proposed method of evaluating mutual fund strategies, we conducted an empirical study based on survey of a total of 30 valid samples from managers of 12 Taiwanese mutual fund companies and researchers of eight research institutions and universities. The majority of the respondents are fund managers responsible for financial or general management. The mutual fund strategy selection process is examined below.

68.5.1 Evaluating the Weights of Criteria

By using the MCDM method, the weights of the factors and criteria are found and are shown in Table 68.1. The empirical evidence shown in Table 68.1 indicates that the weights of each criterion are market timing, 0.524; stock selection ability, 0.318; fund size, 0.141; and teamwork, 0.017, respectively. Therefore, the market timing is the most important factor influencing the performance of mutual funds, followed by the stock selection ability. Prior studies have simultaneously estimated the magnitudes of these portfolio performance evaluation measures. For example, some results show that, on average, mutual fund managers have positive security selection ability but negative market-timing ability (e.g., Chen and Stockum 1986). Since our results suggest that market timing has heavier weight than the stock selection ability, to enhance their performance, mutual fund managers should improve their ability of market timing.

68.5.2 Evaluation and Prioritization of the Mutual Fund Strategy

In this study, the surveyors define their individual range (from 0 to100) for the linguistic variables based on their judgments. By ranking weights and synthetic performance values, we can determine the relative importance of criteria and decide on the best strategies. We apply a λ-measure and nonadditive integral technique to evaluate investment strategies. The synthetic performance of each alternative using different λs is shown in Table 68.2. By ranking the synthetic performance in different λs in Table 68.2, we obtain mutual fund strategy ranking in Table 68.3. In Table 68.3, our empirical results show that when $\lambda < 0$, the aggressive growth style is the most important strategy and growth style is the second most important strategy. When $\lambda \geq 0 – 5$, the results show that growth style is the most important strategy, and equity income style is the second most important strategy. When $\lambda = 10 – 30$, growth style is the most important strategy, followed by the growth income style strategy. When $\lambda \geq 40$, the results show that growth income style replaces growth style becoming the second ranked. On the other hand, when $\lambda \geq 0$, asset allocation style is the worst strategy with the smallest synthetic performance. We can thus infer that the less risky the funds are, the less performance of the funds will be.

68.5.3 Comparing with the Empirical Data

Monthly returns from January 1980 to September 1996 (201 months) for a sample of 65 US mutual funds are used in this study to generate mutual fund performance. The random sample of mutual funds is provided by the MorningStar. The MorningStar segregates mutual funds into four basic investment styles on the basis of manager's portfolio characteristics. Our sample consists of 8 asset allocation (S1), 14 aggressive growth (S2), 10 equity income (S3), 16 growth (S4),

Table 68.1 The weights of criteria for evaluating mutual funds

Criteria	Weight
Market timing	**0.524**
The ratio of fund's market share	0.252
Market returns	0.495
Risk-free interest rate	0.212
Direction of fund flow	0.041
Stock selection ability	**0.318**
P/E ratio	0.252
Net asset value/market value	0.150
Cash flow/market value	0.080
Net asset value	0.187
Risk premium	0.331
Fund size	**0.141**
The market share of mutual fund	0.341
The growth rate of mutual fund scale	0.155
Dividend yield of mutual fund	0.504
Teamwork	**0.017**
Number of researchers	0.293
Education of fund manager	0.182
Known of fund manager	0.429
Turnover rate of fund manager	0.096

Table 68.2 The synthetic performance of mutual fund style

λ	−1.0	−0.5	0.0	1.0	3.0	5.0	10.0	20.0	40.0	100.0	150.0	200.0
S1	395.3	538.0	300.7	299.9	298.8	297.1	296.1	295.9	293.7	292.5	291.6	291.0
S2	616.8	901.2	311.9	310.8	307.6	306.0	303.2	300.2	297.3	294.0	292.0	290.9
S3	457.0	682.5	313.0	311.8	309.5	307.8	304.5	302.1	297.6	294.5	291.7	290.5
S4	552.2	855.7	315.7	312.8	311.7	310.2	308.6	304.4	300.3	298.9	295.8	294.6
S5	363.1	382.9	310.5	309.3	307.2	305.3	304.9	303.5	302.1	301.9	300.1	299.5

S1 is the asset allocation fund, S2 is the aggressive growth, S3 is the equity income, S4 is the growth, and S5 is the growth income fund

Table 68.3 The evaluation results of mutual fund strategy

Mutual fund strategy ranking	
$\lambda = -1, -0.5$	S2 $\succ$ S4 $\succ$ S3 $\succ$ S1 $\succ$S5
$\lambda = 0, 1, 3$	S4 $\succ$ S3 $\succ$ S2 $\succ$ S5 $\succ$ S1
$\lambda = 5$	S4 $\succ$ S3 $\succ$ S5 $\succ$ S2 $\succ$ S1
$\lambda = 10, 20$	S4 $\succ$ S5 $\succ$ S3 $\succ$ S2 $\succ$ S1
$\lambda = 40, 100$	S5 $\succ$ S4 $\succ$ S3 $\succ$ S2 $\succ$ S1
$\lambda = 150, 200$	S5$\succ$ S4 $\succ$ S2 $\succ$ S3 $\succ$ S1

Where S1, asset allocation style; S2, aggressive growth style; S3, equity income style; S4, growth style; and S5, growth income style and 17 growth income (S5) mutual funds. The monthly returns on the S&P 500 Index are used for the market returns. Monthly observations of the 30-day Treasury-bill rate are used as a proxy for the risk-free rate.

Appendix 3 contains summary statistics for the returns of mutual funds. All values are computed in excess of the returns on the US T-bills closest to 30 days to maturity. Data contains mean, standard deviation, maximum, and minimum. Averages of each investment style show that the asset allocation style has the smallest expected return and it also has the smallest standard deviation. However, the aggressive growth style has the largest maximum return but it also has the smallest minimum return and the largest standard deviation. In other words, the more aggressive the funds are, the more volatile the fund returns will be.

The primary purpose of comparing with mutual fund performance data is to find out the true value of λ. Given the true λ value, we can infer other mutual funds' performance during the same period. For example, in Appendix 3, we find the pecking order of mutual funds' performance based upon investment styles is S4 $>$ S5 $>$ S3 $>$ S2 $>$ S1, which is in the same order as shown in Table 68.3 when $\lambda = 10, 20$. Therefore, we find the λ value for certain period when comparing with a sample of mutual fund performance data. Based upon this λ value, we can easily predict the performance of other mutual funds.

68.6 Conclusions

This study focuses on providing a mutual fund strategy for the mutual fund managers so that they could be successful in their decision making. Our empirical study demonstrates the validity of this method. In this study, the mutual fund strategy stems from four aspects: market timing, stock selection ability, fund size, as well as teamwork. Picking a mutual fund from the thousands is not an easy task. Mutual fund managers have difficulty in selecting the proper strategy for reasons such as the uncertain and dynamic environment and numerous criteria that they are facing. Managers are hence overwhelmed by this vague scenario and do not make proper decisions or allocate resources efficiently. The hierarchical method guides the manager how to select the investment style of mutual funds in the uncertainty environment.

We compare our results with the empirical data and find that the model of MCDM predicts the rate of return well in certain ranges of λ. Furthermore, we can use this λ value to compute the performance of different mutual funds; thus the nonadditive integral technique is an effective method to predict the mutual fund performance. By ranking weights and synthetic performance values, we determine the relative importance of criteria, which allows us to decide on the best strategies. We apply a λ-measure and nonadditive integral technique to evaluate investment strategies. By ranking the synthetic performance in different λs, we obtain mutual fund strategy ranking. Our empirical results show that when $\lambda < 0$, the aggressive growth style is the most important strategy; when $\lambda \geq 0 - 5$, the growth style is the most important strategy; when $\lambda = 10 - 30$, growth style is the most important strategy followed by the growth income style strategy. However, when $\lambda \geq 40$, growth income style replaces growth style becoming the second ranked. On the other hand, when $\lambda \geq 0$, asset allocation style is the worst strategy with the smallest synthetic performance. We can thus infer that the less risky the funds are, the less performance of the funds will be, and the more aggressive the funds are, the higher the volatility of the fund performance will be.

Few studies have addressed mutual fund strategy planning. Proposed in this study is a first attempt to formally model the formulation process for a mutual fund strategy using MCDM. We believe that the analysis presented is a significant contribution to the literature and will help to establish groundwork for future research. Even though we are dedicated to setting up the model as completely as possible, there are additional criteria (e.g., tax, expenses, dividend) and methods that could be adopted and added in future research. The mutual fund industry is growing rapidly in the financial markets in response to increasing demand. Therefore, what is needed is a useful and applicable method that addresses the selection of mutual funds. We use a MCDM method to achieve this goal.

Appendix 1

The Description of Evaluative Criteria of Mutual Funds

Criteria	Description
Market timing	The ability of portfolio managers to time market cycles and take advantage of this ability in trading securities
The ratio of fund market share	The ratio of fund invested in securities
The return of market	The fraction of ups or downs of deep bid index in current period divided by the deep bid index in last period
Riskless interest rate	The risk-free interest rate is the interest rate that it is assumed can be obtained by investing in financial instruments with no default risk. In practice most professionals and academics use short-dated government bonds of the currency in question. For Taiwan investments, usually Taiwan bank 1-month deposit rate is used
Flowing of cash	Cash flow refers to the amounts of cash being received and spent by a business during a defined period of time, sometimes tied to a specific project. Measurement of cash flow can be used to evaluate the state or performance of a business or project
Stock selection ability	The ability of fund managers to identify the potential winning securities
P/E ratio	The P/E ratio (price per share/earnings per share) of a mutual fund is used to measure how cheap or expensive its share price is. The lower the P/E, the less you have to pay for the mutual fund, relative to what you can expect to earn from it
Net value/market value	The value of an entity's assets less the value of its liabilities divided by market value

(*continued*)

Criteria	Description
Cash flowing/market value	It equals cash receipts minus cash payments over a given period of time divided by market value or equivalently, net profit plus amounts charged off for depreciation, depletion, and amortization (business) divided by market value
Net value	Net value is a term used to describe the value of an entity's assets less the value of its liabilities. The term is commonly used in relation to collective investment schemes
Risk premium	A risk premium is the minimum difference between the expected value of an uncertain bet that a person is willing to take and the certain value that he is indifferent to
Fund size	The volume and scale of mutual funds
The market share of mutual fund	It can be expressed as a company's sales revenue (from that market) divided by the total sales revenue available in that market. It can also be expressed as a company's unit sales volume (in a market) divided by the total volume of units sold in that market
The growth rate of mutual fund scale	The fraction of the increase or decrease of the fund scale in current period divided by the fund scale in last period
Dividend yield of mutual fund	The dividend yield on a company mutual fund is the company's annual dividend payments divided by its market cap or the dividend per share divided by the price per share
Teamwork	The culture of mutual fund company
Number of researcher	The number of researcher of each fund
Education of fund manager	Fund manager's seniority, quality, and performance
Known of fund manager	Fund manager's rate of exposed in the medium and number of win a prize
Turnover rate of fund manager	Fund manager leaves his job temporarily

Appendix 2

Summary Statistics for Returns of the Mutual Funds

The notations and definition of the investment style of mutual funds are in panel 2.1.
Panel 2.1

Classifications	Investment style	Description
Aa	Asset allocation	A large part of financial planning is finding an asset allocation that is appropriate for a given person in terms of their appetite for and ability to shoulder risk. The designation of funds into various categories of assets
Ag	Aggressive growth	Regardless of the investment style or the size of the companies purchased, funds vary widely in their risk and price behavior which is likely to have a high beta and high volatility

(*continued*)

Classifications	Investment style	Description
Ei	Equity income	It will invest in common stock but will have a portfolio beta closer to 1.0 than to 2.0. It likely favors stocks with comparatively high dividend yields so as to generate the income its name implied
G	Growth	The pursuit of capital appreciation is the emphasis with growth funds. This class of funds includes those called aggressive growth funds and those concentrating on more stable and predictable growth
Gi	Growth income	It pays steady dividends, and it is still predominately an investment in stocks, although some bonds may be included to increase the income yield of the fund

Monthly mutual funds are from January 1980 to September 1996 for a sample of 65 US mutual funds. The data are from Morningstar Company.

Panel 2.2

Fund name	Investment style	Mean	Standard deviation	Maximum	Minimum
General Securities	Aa	0.477	5.084	15.389	−17.151
Franklin Asset Allocation	Aa	0.407	3.743	10.424	−19.506
Seligman Income A	Aa	0.394	2.414	8.474	−7.324
USAA Income	Aa	0.316	2.024	9.381	−5.362
Valley Forge	Aa	0.293	1.803	9.980	−5.573
Income Fund of America	Aa	0.566	2.552	9.166	−8.836
FBL Growth Common Stock	Aa	0.273	3.599	10.466	−24.088
Mathers	Aa	0.220	3.910	14.405	−14.750
Asset allocation average	Aa	0.391	2.550	8.962	−9.464
American Heritage	Ag	−0.905	6.446	28.976	−33.101
Alliance Quasar A	Ag	0.644	6.547	15.747	−39.250
Keystone Small Co Grth (S-4)	Ag	0.433	7.053	19.250	−38.516
Keystone Omega A	Ag	0.473	6.112	18.873	−33.240
Invesco Dynamics	Ag	0.510	6.009	17.378	−37.496
Security Ultra A	Ag	0.222	6.940	16.297	−43.468
Putnam Voyager A	Ag	0.808	5.781	17.179	−29.425
Stein Roe Capital Opport	Ag	0.578	6.783	17.263	−32.135
Value Line Spec Situations	Ag	0.145	6.240	13.532	−37.496
Value Line Leveraged Gr Inv	Ag	0.601	4.970	14.617	−29.025
WPG Tudor	Ag	0.726	6.010	14.749	−33.658

(*continued*)

Fund name	Investment style	Mean	Standard deviation	Maximum	Minimum
Winthrop Aggressive Growth A	Ag	0.476	5.596	17.012	−34.921
Delaware Trend A	Ag	0.787	6.536	14.571	−42.397
Founders Special	Ag	0.564	5.900	12.905	−31.861
Aggressive growth average	Ag	0.459	5.814	13.142	−35.335
Smith Barney Equity Income A	Ei	0.601	3.270	7.813	−18.782
Van Kampen Am Cap Eqty-Inc A	Ei	0.510	3.530	12.292	−22.579
Value Line Income	Ei	0.423	3.357	9.311	−18.242
United Income A	Ei	0.714	4.037	11.852	−13.743
Oppenheimer Equity Income A	Ei	0.555	3.422	10.071	−16.524
Fidelity Equity Income	Ei	0.706	3.612	10.608	−19.627
Delaware Decatur Income A	Ei	0.547	3.615	10.269	−20.235
Invesco Industrial Income	Ei	0.601	3.705	9.349	−20.235
Old Dominion Investors	Ei	0.360	3.699	11.498	−21.092
Evergreen Total Return Y	Ei	0.508	3.220	8.074	−13.857
Equity income average	Ei	0.527	3.238	9.094	−18.718
Guardian Park Avenue A	G	0.740	4.391	11.321	−27.965
Founders Growth	G	0.718	4.986	13.055	−25.108
Fortis Growth A	G	0.724	5.983	14.520	−30.771
Franklin Growth I	G	0.570	4.050	12.907	−11.706
Fortis Capital A	G	0.682	4.791	12.818	−21.585
Growth Fund of America	G	0.625	4.722	12.226	−23.962
Hancock Growth A	G	0.484	5.381	15.708	−25.236
Franklin Equity I	G	0.469	5.156	12.818	−32.135
Nationwide growth	G	0.598	4.370	11.444	−27.570
Neuberger&Berman Focus	G	0.434	4.366	12.187	−25.108
MSB	G	0.517	4.665	13.452	−31.178
Neuberger&Berman Partners	G	0.661	3.612	9.311	−19.385
Neuberger&Berman	G	0.606	5.095	11.574	−30.500
Manhattan	G	0.710	4.067	10.125	−19.385
Nicholas	G	0.225	5.234	11.321	−31.451
Oppenheimer A	G	0.727	5.802	19.120	−37.207
New England growth A	G	0.608	4.505	11.121	−26.081

(*continued*)

Fund name	Investment style	Mean	Standard deviation	Maximum	Minimum
Growth average	G	0.594	4.775	12.649	−26.255
Pioneer II A	Gi	0.517	4.386	10.912	−29.693
Pilgrim America Magna Cap A	Gi	0.611	3.949	10.843	−22.704
Pioneer	Gi	0.410	4.339	12.293	−28.361
Philadelphia	Gi	0.244	4.004	11.074	−23.457
Penn Square Mutual A	Gi	0.504	3.907	11.852	−20.724
Oppenheimer Total Return A	Gi	0.507	4.451	13.861	−22.829
Vanguard/Windsor	Gi	0.726	4.078	10.746	−18.542
Van Kampen Am Cap Gr & Inc A	Gi	0.570	4.781	15.349	−32.135
Van Kampen Am Cap Comstock A	Gi	0.599	4.539	13.167	−34.921
Winthrop Growth & Income A	Gi	0.430	3.987	10.717	−24.088
Washington Mutual Investors	Gi	0.723	3.882	11.409	−20.113
Safeco Equity	Gi	0.587	4.797	14.263	−31.042
Seligman Common Stock A	Gi	0.553	4.224	11.785	−23.331
Salomon Bros Investors O	Gi	0.583	4.194	11.785	−24.980
Security Growth & Income A	Gi	0.233	3.825	10.161	−19.674
Selected American	Gi	0.650	3.969	13.142	−19.385
Putnam Fund for Grth & Inc A	Gi	0.637	3.540	8.456	−22.081
Growth income average	Gi	0.544	3.940	10.380	−24.469

Appendix 3

Summary Statistics for Returns of the Mutual Funds

Fund name	Investment style	Mean	Standard deviation	Maximum	Minimum
Asset allocation average	S1	0.391	2.550	8.962	−9.464
Aggressive growth average	S2	0.459	5.814	13.142	−35.335
Equity income average	S3	0.527	3.238	9.094	−18.718
Growth average	S4	0.594	4.775	12.649	−26.255
Growth income average	S5	0.544	3.940	10.380	−24.469

Appendix 4

The MCDM proposed approach consists of eight steps: define the problem, define the evaluation criteria, initial screen, define the preferences on evaluation criteria, define the MCDM method for selection, evaluate the MCDM methods, choose the most suitable method, and conduct sensitivity analysis.

Step 1: Define the problem. The characteristics of the decision-making problem under consideration are addressed in the problem definition step, such as identifying the number of alternatives, attributes, and constraints. The available information about the decision-making problem is the basis on which the most appropriate MCDM techniques will be evaluated and utilized to solve the problem.

Step 2: Define the evaluation criteria. The proper determination of the applicable evaluation criteria is important because they have great influence on the outcome of the MCDM method selection process. However, simply using every criterion in the selection process is not the best approach because the more criteria used, the more information is required, which will result in higher computational cost. In this study, the characteristics of the MCDM methods will be identified by the relevant evaluation criteria in the form of a questionnaire. Ten questions are defined to capture the advantages, disadvantages, applicability, computational complexity, etc. of each MCDM method, as shown in the following. The defined evaluation criteria will be used as the attributes of an MCDM formulation and as the input data of decision matrix for method selection:

1. Is the method able to handle MADM, MODM, or MCDM problem?
2. Does the method evaluate the feasibility of the alternatives?
3. Is the method able to capture uncertainties existing in the problem?
4. What input data are required by the method?
5. What preference information does the method use?
6. What metric does the method use to rank the alternatives?
7. Can the method deal changing alternatives or requirements?
8. Does the method handle qualitative or quantitative data?
9. Does the method deal with discrete or continuous data?
10. Can the method handle the problem with hierarchy structure of attributes?

Step 3: Initial screen in the initial screen step. The dominated and infeasible MCDM methods are eliminated by dominance and conjunctive. An alternative is dominated if there is another alternative which excels it in one or more attributes and equals it in the remainder. The dominated MCDM methods are eliminated by the dominance method, which does not require any assumption or any transformation of attributes. The sieve of dominance takes the following procedures. Compare the first two alternatives, and if one is dominated by the other, discard the dominated one; then compare the un-discarded alternative with the third alternative and discard any dominated alternative; and then introduce the fourth alternative and repeat this process until the last alternative has been compared. A set of non-dominated alternatives may possess unacceptable or infeasible attribute values. The conjunctive method is employed to remove the unacceptable alternatives, in which the decision maker sets up the cutoff value he/she will accept for each of the attributes. Any alternative which has an attribute value worse than the cutoff values will be eliminated.

Step 4: Define the preferences on evaluation criteria. Usually, after the initial screen step is completed, multiple MCDM methods are expected to remain; otherwise we can directly choose the only one left to solve the decision-making problem. With the ten evaluation criteria defined in step 2, the decision maker's preference information on the evaluation criteria is defined. This will reflect which criterion is more important to the decision maker when he/she makes decisions on method selection.

Step 5: Define the MCDM method for selection. Existing commonly used MCDM methods are identified and stored in the method base as candidate methods for selection. The simple additive weighting (SAW) method is chosen to select the most suitable MCDM methods considering its simplicity and general acceptability. Basically, the SAW method provides a weighted summation of the attributes of each method, and the one with the highest score is considered as the most appropriate method. Though SAW is used in this study, it is worth noting that other MCDM methods can be employed to handle the same MCDM methods selection problem.

Step 6: Evaluate the MCDM methods. Mathematical formulation of appropriateness index (AI) is used to rank the MCDM methods. The method with the highest AI will be recommended as the most appropriate method to solve the problem under consideration.

Step 7: Choose the most suitable method. For optimization of specification of grinding wheel, the MCDM method which has the highest AI will be selected as the most appropriate method to solve the given decision-making problem. If the DM is satisfied with the final results, he/she can implement the solution and move forward. Otherwise, he/she can go back to step 2 and modify the input data or preference information and repeat the selection process until a satisfied outcome is obtained. Be displayed to provide guidance to DM is provided guidance about how to get the final solution by using the selected method. In addition, the detailed mathematical calculation steps are also built in the MATLAB-based DSS, which highly facilitates the decision-making process. Thus, the DM can input their data according to the instruction and get the final results by clicking one corresponding button.

Step 8: Conduct analysis. In this section, selection of an optimized specification of grinding wheel problem is conducted to improve the capabilities of the grinding operation products by proposed MCDM decision support system. It is observed that different decision makers often have different preference information on the evaluation criteria and different answers to the ten questions; thus, analysis should be performed on the MCDM method selection algorithm in order to analyze its robustness with respect to parameter variations, such as the variation of decision maker's preference information and the input data. If the decision maker is satisfied with the final results, he/she can implement the solution and move forward. Otherwise, he/she can go back to step 2 and modify the input data or preference information and repeat the selection process until a satisfied outcome is obtained. In this implementation, emphasis is put on explaining the holistic process of the intelligent MCDM decision support system. Thus, the step-by-step problem-solving process is explained and discussed for this decision-making problem.

References

Bhattacharya, S., & Pfleiderer, P. (1983). A note on performance evaluation. Technical report 714, Stanford University, Graduate School of Business.

Black, F., Jensen, M. C., & Scholes, M. (1972). The capital asset pricing model: Some empirical tests. In M. C. Jensen (Ed.), *Studies in the theory of capital markets*. New York: Praeger.

Blake, C. R., & Morey, M. R. (2000). Morningstar ratings and mutual fund performance. *Journal of Financial & Quantitative Analysis, 35*(3), 451–483.

Blume, M. E., & Friend, I. (1973). A new look at the capital asset pricing model. *Journal of Finance, 28*, 19–33.

Chang, E. C., & Lewellen, W. G. (1984). Market timing and mutual fund investment performance. *Journal of Business, 57*, 57–72.

Chen, C. R., & Stockum, S. (1986). Selectivity, market timing, and random beta behavior of mutual funds: A generalized model. *Journal of Financial Research, 9*(1), 87–96.

Connor, G., & Korajczyk, R. (1986). Performance measure with the arbitrage pricing theory: A new framework for analysis. *Journal of Financial Economics, 15*(3), 373–394.

Fama, E. F. (1972). Components of investment performance. *Journal of Finance, 27*(3), 551–567.

Grant, D. (1977). Portfolio performance and the cost of timing decisions. *Journal of Finance, 32*, 837–846.

Henriksson, R. D. (1984). Market timing and mutual fund performance: An empirical investigation. *Journal of Business, 57*(1), 73–96.

Jensen, M. C. (1968). The performance of mutual funds in the period 1945−1964. *Journal of Finance, 23*(2), 389–416.

Jensen, M. C. (1969). Risk, the pricing of capital assets, and the evaluation of investment portfolios. *Journal of Business, 42*(2), 167–247.

Jensen, M. C. (1972). Optimal utilization of market forecasts and the evaluation of investment performance. In G. P. Szego & K. Shell (Eds.), *Mathematical methods in investment and finance*. Amsterdam: Elsevier.

Jorge, S., Pilar, G., & Luis, M. D. (2006). Mutual fund performance and benchmark choice: The Spanish case. *Applied Financial Economics Letters, 2*(5), 317–321.

Kopa, M., & Post, T. (2009). A portfolio optimality test based on the first-order stochastic dominance criterion. *Journal of Financial & Quantitative Analysis, 44*(5), 1103–1124.

Kuosmanen, T. (2004). Efficient diversification according to stochastic dominance criteria. *Management Science, 50*(10), 1390–1406.

Lee, C. F., & Rahman, S. (1990). Market timing, selectivity, and mutual fund performance: An empirical investigation. *Journal of Business, 63*(2), 261–278.

Lehmann, B. N., & Modest, D. M. (1987). Mutual fund performance evaluation: A comparison of benchmarks and benchmark comparisons. *Journal of Finance, 42*(2), 233–265.

Markowitz, H. (1952). Portfolio selection. *Journal of Finance, 7*(1), 77–91.

Merton, R. C. (1981). On market timing and investment performance I: An equilibrium theory of value for market forecasts. *Journal of Business, 54*, 363–406.

Spitz, A. E. (1970). Mutual fund performance and cash inflows. *Applied Economics, 2*(2), 141–145.

Treynor, J. L., & Mazuy, K. K. (1966). Can mutual funds outguess the market? *Harvard Business Review, 44*, 131–136.

Wang, S. Y., & Lee, C. F. (2011). Fuzzy multi-criteria decision making for evaluating mutual fund strategies. *Applied Economics, 43*(24), 3405–3414.

Econometric Analysis of Currency Carry Trade

69

Yu-Jen Wang, Huimin Chung, and Bruce Mizrach

Contents

Abstract

The carry trade is a popular strategy in the currency markets whereby investors fund positions in high interest rate currencies by selling low interest rate currencies to earn the interest rate differential. In this article, we first provide an overview of the risk and return profile of currency carry trade; second, we introduce two popular models, the regime-switch model and the logistic smooth transition regression model, to analyze carry trade returns because the carry trade returns are highly regime dependent. Finally, an empirical example is illustrated.

Keywords

Carry trade • Uncovered interest parity • Markov chain Monte Carlo • Regime-switch model • Logistic smooth transition regression model

Y.-J. Wang (✉) • H. Chung
Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
e-mail: yjwang1.iof96g@g2.nctu.edu.tw; chunghui@mail.nctu.edu.tw

B. Mizrach
Department of Economics, Rutgers, The State University of New Jersey, New Brunswick, NJ, USA
e-mail: mizrach@econ.rutgers.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_69,

69.1 Introduction

According to the international financial theory such as uncovered interest parity (UIP), the exchange rate will depreciated in the future if the country has a high interest rate. Although investors can potentially profit from the interest rate differential between the two countries, the exchange rate differential may offset the interest rate differential. However, in the past two decades, an enormous amount of empirical research has refuted the UIP theory. These evidences suggest that when exchange rate returns combine with time-varying premium, UIP is not usually present in the past.[1]

Carry trade, a common international investment strategy, is a good example that runs contradictory to UIP theory. Carry trade is built by borrowing currency from a lower interest rate country and then investing in a higher interest rate country. A large body of research shows that carry trade is profitable over the long horizon. How the carry trade strategy earns a persistent excess return is an open question. The literature suggests several explanations for the forward premium puzzle. Engel (1984) and Fama (1984) provide the straightforward and theoretically convincing explanation based on the existence of time-varying risk premia for this puzzle. Burnside et al. (2011) refer to the peso problem as an explanation for the high average payoff to the carry trade.[2] Baillie and Chang (2011) provide another explanation for this puzzle that focuses on the trading behavior.

However, because UIP is not always held in the short term, previous empirical studies have adopted several models that allow for temporary deviations from UIP and have discussed regime dependence among other factors to fit the carry trade return. For example, Ichiue and Koyama (2008) provide the regime-switch model to detect how the exchange rate volatility influences UIP. The failures of UIP usually happen at relatively low volatility environment. In particular, they argue that the rapidly unwinding carry trade affects the exchange rate volatility. Recently, using daily data from 1985 to 2008, Christiansen and Ranaldo (2011) analyze carry trade returns with the multifactor model. Their main findings suggest high regime dependence of the carry trade return. Clarida et al. (2009) also find significant volatility regime sensitivity for Fama regressions estimated over low and high volatility periods. Sarno et al. (2006)

---

[1]Meese and Rogoff (1983) assume that exchange rates follow the "near-random walk" model and provide the evidence to reject UIP. Fama (1984) applies the concept of forward rate contained in the time-varying premium to analyze the relation between the forward exchange rate and spot exchange rate and points out that high interest rate currencies tend to appreciate rather than depreciate. Froot and Thaler (1990) replace time-varying premium with the mean return theory to explain foreign exchange anomalies. Burnside et al. (2009) emphasize that the forward premium puzzle can be construed as the adverse selection problems between participants in foreign exchange markets. Brunnermerier et al. (2008) use the liquidity risk factor to explain the excess return of the carry trade. They add the change of VIX index or the TED spread variable to be the liquidity risk factors in the regression and suggest that the market liquidity factor may explain the carry trade's risk premium.

[2]The peso problem is as a generic term for the effects of small probabilities of large events in empirical work. Burnside et al. (2011) approach relies on analyzing the payoffs to a version of the carry trade strategy that does not yield high negative payoffs in the peso state.

provide empirical evidence the deviations from UIP display significant nonlinearities that consistent with theories based on transactions costs or limits to speculation. Menkhoff et al. (2012) explain that low returns occur in times of unexpected high volatility, when low interest rate currencies provide a hedge by yielding positive returns. Empirically, however, the literature has serious problems convincingly claiming that carry trade returns as regime dependence or that the *higher profit opportunities of carry trade usually* occurred in lower volatility regime.

In sum, massive losses associated with the carry trade usually occurred in the higher volatility regime, the positive return usually exists in carry trade at the low volatility regime, and the carry trade return process does not follow the traditional linear models. Baillie and Chang (2011) further find that momentum trading increases carry trade volatility. Brunnermerier et al. (2008) also find the levered market participants gradually build up positions in high-yielding currencies, causing high-yielding currencies to appreciate over time along with speculators' larger positions. In addition, Brunnermerier et al. (2008) find that higher market volatility is associated with carry trade losses. Clarida et al. (2009) use the Fama regression, which produces a positive coefficient that is greater than unity when volatility is in the top quartile.[3] Baillie and Chang (2011) find that UIP is more likely to hold in a regime when volatility is unusually high. These results suggest that the momentum effect in the carry trade perhaps exists in the low volatility regime, but the empirical results are mixed in the high volatility regime. However, the carry trade perhaps confronts crash risk in the high volatility regime whether UIP holds.

Because the carry trade return is highly regime dependent, regime conditions must be considered in the return process model. Ichiue and Koyama (2008) use the regime-switch model to investigate the relation among exchange rate returns, volatilities, and interest rate differentials. Baillie and Chang (2011) use the logistic smooth transition regression (LSTR) model to identify whether the forward FX market is in a regime where the anomaly is present or whether it is in a regime where UIP tends to hold. They find that UIP is more likely to hold in a regime where volatility is unusually high, which may be explained by previous theoretical work that links momentum trading to increased volatility and more pronounced reversion to fundamentals.

In estimating the model, we suggest the Markov chain Monte Carlo (MCMC) methods. The major advantage of the Bayesian MCMC approach is its extreme flexibility. Because the parameters are generated by the random variable from posterior distributions, the MCMC method can avoid the thorny problem of maximum such as the maximum likelihood estimation. This method is well suited to fit realistic models to complex data sets with threshold value, measurement error, censored or missing observations, multilevel or serial correlation structures, and multiple endpoints. This method is suitable for the regime-dependent models because these models usually have a threshold point. For instance, Ichiue and Koyama (2008) use the Bayesian Gibbs sampling method to estimate the parameters

[3]The empirical result of Clarida et al. (2009) indicates that UIP is violated in the high volatility regime.

of the regime-switch model, and early investigations also used MCMC method, such as Albert and Chib (1993), Kim et al. (1998), and Kim and Nelson (1999).

The study provides a more clear analysis of carry trade return behavior in the different volatility regimes. Our main purpose is to construct an investment strategy that can create lower volatility and higher return of the carry trade, where the carry trade returns are produced by exchange rate and 3-month interbank rate data and the volatility variables are determined by VIX, which is calculated from the S&P500 equity-options market, generalized autoregressive conditional heteroskedasticity (GARCH), and exponential GARCH (EGARCH).[4] According to our empirical outputs, we find several results. First, we confirm that carry trade returns display a significant momentum effect at the low volatility regime. Second, the VIX performance of carry trade is better than other volatility measures during the subprime period (2007–2008). Third, we can create lower downside risk strategy than the buy-and-hold strategy in the long run when we used the GARCH volatility measure. Finally, we still have a higher Sortino ratio by using the GARCH volatility measure.

The remainder of this study is organized as follows. We introduce two financial econometrics models, the regime-switch model and LSTR model, in Sect. 69.2.1. Section 69.2.2 discusses our method, based on the MCMC method. We discuss the carry trade return behavior and provide a new carry trade trading strategy of carry trade in Sect. 69.3. Section 69.4 concludes.

69.2 Overview of Model and Methodology for Carry Trade

Based on past study, the regime-dependent model is most popular to analyze carry trade returns because the behaviors of carry trade returns are diversely in the different volatility regimes. We will introduce two popular regime-dependent models, the regime-switching and logistic smooth transition regression models, in this section. The regime-dependent models usually have threshold value such as regime-switching model. We suggest using the MCMC methods to avoid this thorny problem of parameter estimation.

69.2.1 The Regime-Switching and Logistic Smooth Transition Regression Models

The covered interest parity (CIP) is a non-arbitrage condition. It postulates that the nominal interest differential between two countries ($i_t^* - i_t$) should equal the forward premium ($f_t - s_t$). It is expressed as

$$E_t[\Delta s_{t+1}] = i_t^* - i_t = f_t - s_t, \tag{69.1}$$

[4]Wang et al. (2012) suggest that GARCH models with skew density innovations may be another suitable volatility measure for carry trade return.

where s_t is the logarithm of the spot exchange rate quoted as the foreign price of domestic currency, f_t is the logarithm of the forward rate for a one-period ahead transaction, and i_t and i_t^* are the one-period risk-free domestic and foreign interest rates, respectively. The standard test of UIP to estimate the regression is

$$\Delta s_{t+1} = \alpha + \beta(f_t - s_t) + u_{t+1}. \tag{69.2}$$

Under UIP, the null hypothesis is that $\alpha = 0$ and $\beta = 1$ that the error term, u_{t+1}, is serially uncorrelated. The forward premium anomaly generally refers to the widespread phenomenon of a negative slope coefficient being obtained by the ordinary least square estimation of Eq. 69.2. Baillie and Chang (2011) further use the LSTR model, which postulates that the slop coefficient is related nonlinear to the degree of carry and momentum trading over time; a natural approach is to specify the UIP relation in terms of the LSTR model:

$$\begin{aligned}\Delta s_{t+1} = {} & [\alpha_1 + \beta_1(f_t - s_t)](1 - G(z_t;\gamma,c)) \\ & + [\alpha_2 + \beta_2(f_t - s_t)]G(z_t;\gamma,c) + u_{t+1},\end{aligned} \tag{69.3}$$

where $G(\cdot)$ is a logistic transition function as follows:

$$G(z_t;\gamma,c) = \left(1 + \exp\left(\frac{-\gamma(z_t - c)}{\sigma_{z_t}}\right)\right)^{-1}, \gamma > 0, \tag{69.4}$$

where z_t is the transition variable, σ_{zt} is the standard deviation of z_t, γ is a slope parameter, and c is a location parameter. Baillie and Chang choose various transition variables, z_t, related to carry and consider momentum trading. Specifically, they use the interest differentials and the conditional volatility of exchange rates as measured by GARCH(1,1) models of spot exchange rate returns. This model approach works well for carry trade analysis because it allows for smooth and continuous adjustment between regimes.

Another popular regime-dependent model is the regime-switch model. After Hamilton (1989) proposed the regime-switching model to examine the persistency of recessions and booms, many studies applied this model to exchange rate data. Engel and Hamilton's (1990) two-regime model specifies currency returns as

$$s_{t+1} - s_t = \alpha_i + \sigma_i \eta_{t+1}, \tag{69.5}$$

where $i \in \{1, 2\}$ denotes the regime; α_i and σ_i denote the trend of exchange rate and the volatility of exchange rate return under regime i, respectively; and $\eta_{t+1} \sim^{i.i.d.} N(0, 1)$. Ichiue and Koyama (2008) employ the four-regime model to discuss carry trade returns. First, they use the following nesting model:

$$s_{t+1} - s_t = \alpha_i + \beta_i\left(i_t - i_t^*\right) + \sigma_i \eta_{t+1}. \tag{69.6}$$

According to the views of market participants, regime switches in exchange rate returns should be interpreted as switches in the relation between the returns and interest rate differentials or switches in market participants' activities between the carry trade and its unwinding, rather than just switches in trends. Ichiue and Koyama assume that the intercept α_i does not switch, that is, $\alpha_i = \alpha$ for all $i = 1, 2$. Second, they define a regime indicator variable that spans the regime space for both the slope and volatility regimes as

$$S_t = \begin{cases} 1 \;\; if \;\; S_{\beta_t} = 1 \; and \; S_{\sigma_t} = 1 \\ 2 \;\; if \;\; S_{\beta_t} = 2 \; and \; S_{\sigma_t} = 1 \\ 3 \;\; if \;\; S_{\beta_t} = 1 \; and \; S_{\sigma_t} = 2 \\ 4 \;\; if \;\; S_{\beta_t} = 2 \; and \; S_{\sigma_t} = 2 \end{cases}, \tag{69.7}$$

where S_{β_t} is the slope regime, which indicates the relation between exchange rate returns and interest rate differentials, and at time t is β_i when $S_{\beta_t} = i, i = 1, 2$. S_{σ_t} is defined as the volatility regime, with the volatility at time t being σ_j when $S_{\sigma_t} = j$, $j = 1, 2$. Finally, Ichiue and Koyama's model can be described as

$$s_{t+n} - s_t = \alpha + \beta_t \left(i_{t,n} - i^*_{t,n} \right) + \sigma_t \eta_{t+n}. \tag{69.8}$$

$$\beta_t = \beta_1 (I_{1t} + I_{3t}) + \beta_2 (I_{2t} + I_{4t}). \tag{69.9}$$

$$\sigma_t = \sigma_1 (I_{1t} + I_{2t}) + \sigma_2 (I_{3t} + I_{4t}), \tag{69.10}$$

where $I_{kt} = 1$ if $S_t = k$ and $I_{kt} = 0$ if $S_t \neq k$, $k = 1, 2, 3, 4$, where $\beta_1 < \beta_2$ and $0 < \sigma_1 < \sigma_2$. And $\Pr[S_{t+1} = k | S_t = l] = p_{kl}$ are the transition probabilities in the transition matrix P^n, for $k, l = 1, 2, 3, 4$, where $\sum_{k=1}^{4} p_{kl} = 1$. This model provides four kinds of regimes for carry trade returns. For instance, when $S_t = 1$, the carry trade return will be with lower slope coefficient β_1 and low volatility. Ichiue and Koyama refer to this regime the negative/low regime. The remainder of the regimes are positive/low, when $S_t = 2$; negative/high, when $S_t = 3$; and positive/high, when $S_t = 4$.[5]

The parameters of these regime-dependent models can be estimated by the MCMC method, especially regime-switch models, because regime-switch models have threshold points to separate the regimes. We introduce this method in the next subsection.

69.2.2 The MCMC Method

Bayesian inference using the MCMC method is a popular technique for parameter estimation. We can estimate parameters easily via this technique. MCMC includes

[5]The empirical results shown β_1 and β_2 are negative and positive, respectively.

two major algorithms, Gibbs sampling and Metropolis-Hustings (M-H) algorithm. The algorithms are sampled from probability distributions based on constructing a Markov chain that has the desired distribution as its equilibrium distribution.

Metropolis et al. (1953) introduced the M-H algorithm. Setting the interest distribution $\pi = \{\pi_j, j \in S\}$ with $\sum_j \pi_j = 1$ and an arbitrary irreducible transition matrix $Q = [q_{ij}]$ on the state space S with the elements satisfying $q_{ij} = q_{ji}$, the MCMC $\{x_t, t \geq 0\}$ is constructed by the following steps. First, set $x_k =$ i where i is any realization from π. Second, generate j from the $\{q_{ij}; j = 1, 2, \ldots\}$ where q_{ij} is often called the proposal or candidate generating function. Third, set $\alpha = \pi_j/\pi_i$; if $\alpha \geq 1$, set $x_{k+1} = q_{ij}$. Otherwise generate $u \sim U(0, 1)$ if $u < \alpha$; set $x_{k+1} = q_{ij}$, else $x_{k+1} = x_k$. Finally, set $k = k + 1$ and go to the first step.

The Gibbs sampler is a generally applicable method, which is an algorithm for generation values from the full conditional distributions. The algorithm for Gibbs sampling is as follows: First, give an initial value $\boldsymbol{\Theta}^{(0)} = (\theta_1^{(0)}, \theta_2^{(0)}, \ldots, \theta_p^{(0)})$ for parameter $\boldsymbol{\Theta} = (\theta_1, \theta_2, \ldots, \theta_p)$. Second, draw θ_i from the full conditional $\pi(\theta_i | \boldsymbol{\Theta}^{(k)}_{-i})$ for $i = 1, \ldots, p$ at k th iteration, where $\boldsymbol{\Theta}^{(k)}_{-i}$ is defined by $\boldsymbol{\Theta}^{(k)}_{-i} = \left(\theta_1^{(k)}, \theta_2^{(k)}, \ldots, \theta_{i-1}^{(k)}, \theta_{i+1}^{(k-1)}, \ldots, \theta_p^{(k-1)}\right)$. Third, set $k = k + 1$; return to second step until convergence is achieved. We can use the M-H algorithm to generate the random variables for parameters if they are difficult to generate by full conditional distributions.

For example, we review the regime-switch model. We can find Eq. 69.8 following the normal distribution as

$$(s_{t+n} - s_t) \sim N\left(\alpha + \beta_t\left(i_{t,n} - i^*_{t,n}\right), \sigma_t^2\right). \tag{69.11}$$

Ichiue and Koyama (2008) employ the noninformation priors for all the parameters of the model as $\alpha \sim N(0, 10)$, $\beta_1 \sim N(-1, 10)$, $\beta_2 \sim N(1, 10)$, $\sigma_1^2 \sim IG(4{,}300)$, $\sigma_2^2/\sigma_1^2 \sim IG(4{,}8)$, and $(p_{k1}, \ldots, p_{k4}) \sim Dirichlet(p_{0,k1}, \ldots, p_{0,k4})$ where $p_{0,\,kk} = 4$ and $p_{0,\,kl} = 1$ if $k \neq l$. The selection of priors can be based on prior knowledge or user experience. The noninformation (diffuse) priors are useful to be able to conveniently calculate full conditional distributions. For instance, the conditional distribution (posterior distribution) of α can be found as

$$\pi\left(\alpha \middle| \boldsymbol{\Theta}^{(k)}_{-i}\right) \sim N\left(s_{t+n} - s_t, \sigma_t^2\right), \tag{69.12}$$

where $\boldsymbol{\Theta}^{(k)}_{-i} = \left(s_{t+n} - s_t, \beta_1^{(k)}, \beta_2^{(k)}, \sigma_1^{(k)}, \sigma_2^{(k)}, P^{n(k)}\right)$. All parameters can be draw from full conditional distributions via the MCMC method.

69.3 Empirical Result

In this section, we discuss the relation between carry trade return and volatility measures. We calculate a series of long Australian dollars (AUD) and short

Japanese Yen (JPY) weekly carry trade return and three kinds of volatility measure (VIX, GARCH, and EGARCH) from 4 January 2001 to 18 August 2010. Our sample consists of 503 observations.

We calculate weekly return of short JPY and long AUD carry trades using the following formula:

$$y_{t+1} = \log\left(1 + (i_t^* - i_t)\right) - \Delta s_{t+1}, \tag{69.13}$$

where i_t^* is Australia interest rate at time t and i_t is Japan interest rate. Therefore, log $(1 + (i_t^* - i_t))$ is log return of interest rate differential and $\Delta s_{t+1} = s_{t+1} - s_t$ with S_j denotes logarithm of the nominal exchange rate at time j (unit of AUD per JPY).

Table 69.1 presents the summary statistics for the weekly return from investing in a long position in the Australian dollar (AUD) financed by borrowing in the Japanese Yen (JPY) and three volatility measures, VIX, GARCH, and EGARCH, from 4 January 2001 to 18 August 2010. Panels B and C show the summary statistics for the weekly return in the low and high volatility regimes, respectively. We can observe that, relative to the normal distribution, the density of the weekly returns exhibits skewness and excess kurtosis, especially the returns in the high volatility regime. Brunnermerier et al. (2008) explained that carry trade returns have crash risk that caused the negative skewness and excess kurtosis. They pointed out that "exchange rates go up by the stairs and down by the elevator." Based on the VaR results, we find the losses at high volatility regime are larger than low volatility regime. However, Table 69.1 clearly shows that the behaviors of carry trade returns are diversely in the different volatility regimes.

First, we use the AR(1) model to discuss the relation between momentum effect and volatility regimes.[6] We consider following the AR(1) model:

$$y_{t+1} = \phi_0 + \phi_1 y_t + \varepsilon_{t+1}, \tag{69.14}$$

where ε_{t+1} is serially uncorrelated. If parameter ϕ_1 is positive, we assume that the momentum effect exists because the directions of returns are the same between the now and future. We show the results in the four kinds of volatility regimes in Table 69.2, in which we use three volatility measures, VIX, GARCH, and EGARCH, to capture the volatility of carry trade.

However, based on results of Table 69.2, we create a new carry trade trading strategy that results in higher return and downside risk. At the low volatility, we have a momentum trading strategy of carry trade. On the other hand, we do nothing during the high volatility regime. For robustness, we separate our sample into 2, 5, and 10 years. We use buy-and-hold strategy as our benchmark to

[6]Based on Eq. 69.6, we divide the long AUD and short JPY carry trade from 4 January 2001 to 18 August 2010 into four kinds of volatility regimes and try to use the MCMC method to test whether UIP is existing in these intervals, but we cannot find any evidences to support it.

Table 69.1 Descriptive statistics of weekly returns for JPY carry trades and three volatility measures from 4 January 2001 to 18 August 2010

	Mean	S.D.	Skew.	Kur.	Size	Q1	Q2	Q3[a]	Min.	Max.	VaR[b]
Panel A: Weekly returns and volatility measures											
Returns	0.0043	0.0231	−1.5872	10.4449	503	−0.0054	0.0075	0.0172	−0.1487	0.0828	0.0338
VIX	0.2195	0.1020	1.9379	8.7591	503	0.1442	0.1969	0.2594	0.1002	0.7913	–
GARCH	0.0219	0.0111	2.6467	12.3584	503	0.0149	0.0185	0.0248	0.0113	0.0896	–
EGARCH	0.0205	0.0076	1.3201	4.7824	503	0.0152	0.0186	0.0238	0.0091	0.0514	–
Panel B: Weekly returns in the low volatility regime[c]											
	0.0068	0.0146	−0.1561	4.0953	251	−0.0017	0.0080	0.0154	−0.0413	0.0562	0.0172
Panel C: Weekly returns in the high volatility regime											
	0.0018	0.0291	−1.3876	7.6085	252	−0.0091	0.0064	0.0192	−0.1487	0.0828	0.0460

[a]Denote the first, second, and third quartiles as Q1, Q2, and Q3, respectively
[b]The 95 % VaRs are calculated by the normal distribution
[c]If the return in the low volatility regime where the EGARCH volatility measure is smaller than Q2

Table 69.2 The parameters are calculated by AR(1) model in the four kinds of volatility regimes that are distinguished by the first, second, and third quartiles as Q1, Q2, and Q3, respectively. The volatility measures, V, are found by VIX, GARCH, and EGARCH. The weekly carry trade returns are calculated by Eq. 69.13 from 4 January 2001 to 18 August 2010

	V<Q1	Q1<V<Q2	Q2<V<Q3	V>Q3
Panel A V: EGARCH				
ϕ_0	0.0035***	0.0055***	0.0062***	0.0008
p-value	0.0002	0.0000	0.0009	0.8130
ϕ_1	0.1090*	0.0379	−0.0679	−0.0233
p-value	0.0631	0.6040	0.4530	0.8110
Panel B V: GARCH				
ϕ_0	0.0065***	0.0069***	0.0066***	−0.0037
p-value	0.0000	0.0000	0.0001	0.2870
ϕ_1	0.0210	0.0232	0.0441	−0.0665
p-value	0.6970	0.7080	0.5689	0.5110
Panel C V: VIX				
ϕ_0	0.0058***	0.0060***	0.0049**	−0.0002
p-value	0.0000	0.0001	0.0206	0.9470
ϕ_1	0.1149	0.0627	−0.1244	0.0142
p-value	0.1820	0.4841	0.1444	0.8780

*, **, and *** denote significance at the 10 %, 5 %, and 1 % levels, respectively
Based on these data, we find results consistent with previous investigations: the carry trade displays the momentum effect only in the low volatility regime. In addition, the EGARCH volatility measure has a significant result. The negative relation between now and future returns are usually found in the high volatility regime. This phenomenon may imply the carry trade's exposure to crash risk but the parameter ϕ_1 is always insignificant

Table 69.3 We compare the total return of four types of strategy. First, we calculate the third quartiles (Q3) with three different volatility measures. Second, if the return in the low volatility regime where the volatility measure is smaller than Q3, we do the momentum trading strategy of carry trade. Otherwise, we do nothing

Period	VIX	GARCH	EGARCH	Buy and hold
10 years (2001–2010)	1.0277	1.0413	0.8416	2.1070[a]
5 years (2006–2010)	0.8313[a]	0.7935	0.7963	0.8207
5 years (2001–2005)	0.1773	0.4198	0.5439	1.2768[a]
2 years (2009–2010)	0.0634	0.1193	−0.1907	0.4260[a]
2 years (2007–2008)	0.5285[a]	0.4814	0.4550	0.0669
2 years (2005–2006)	0.2665	0.2646	0.3658	0.6015[a]
2 years (2003–2004)	0.1906	0.3474	0.2758	0.5779[a]
2 years (2001–2002)	−0.1088	−0.0112	0.0868	0.4009[a]

[a]Denotes the maximum cumulative return that an investment has gained or lost over sample period

compare with our strategy performance (cumulative return), which is created by three volatility measures.

Table 69.3 shows the results. The buy-and-hold strategy usually has the best performance. During the financial crisis (2007–2008), the VIX threshold produces

Table 69.4 We compare the Sortino ratio of four types of strategy, where the semi-standard deviations (downside risks) are calculated by negative deviations for each return interval

Period	VIX	GARCH	EGARCH	Buy and hold
10 years (2001–2010)	66.9450	94.3632[a]	67.5042	86.4720
5 years (2006–2010)	55.2632	82.3461[a]	66.5947	27.4893
5 years (2001–2005)	12.6324	39.8917	59.5527	110.3050[a]
2 years (2009–2010)	2.2658	10.9572	−7.2599	15.6953[a]
2 years (2007–2008)	34.0658	41.1898[a]	37.8497	1.8875
2 years (2005–2006)	30.7356	36.8381	42.5293	88.2519[a]
2 years (2003–2004)	16.5059	61.9929[a]	26.1353	60.9510
2 years (2001–2002)	−10.7429	−1.5258	8.5799	30.3857[a]

[a]Denotes the maximum Sortino ratio among the four strategies

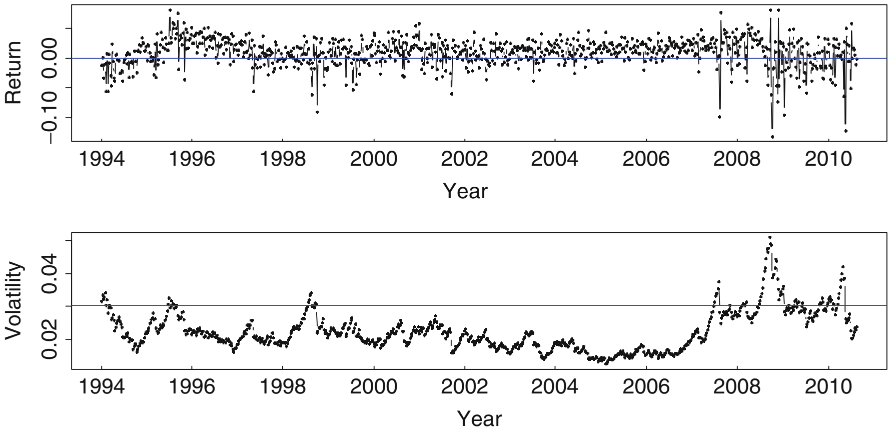


Fig. 69.1 Weekly return of long AUD and short JPY carry trade from 13 January 1994 to 18 August 2010 and weekly volatilities, calculated by AR(1)-EGARCH(1,1) model

the highest performance (0.5285), which explains why prior studies commonly employ VIX as a proxy variable for global risk. During the same tumultuous period, the buy-and-hold strategy shows lower returns than other periods. The reason may be that traders suffered the risk from the low interest rate country's currency appreciated, and thus UIP tends to hold. Although the buy-and-hold strategy often has the highest returns, it also may suffer from the highest crash risk. To consider the downside risk of the carry trade strategy, we use the Sortino ratio criteria to compare the performance of strategies. Sortino ratio measures how many units of return are received per unit of downside risk experienced (Riddles 2001). Table 69.4 provides the results of the comparison of the Sortino ratio of four types of strategy.

As Table 69.4 shows, using the GARCH volatility measure, our trading strategy produces the highest return per unit of the downside risk in the long run.

It also shows that the GARCH is the best variable to explain our trading strategy, which usually has a higher Sortino ratio. Thus, although the buy-and-hold strategy often has the highest returns, our strategy effectively avoids loss at high volatility regime. Figure 69.1 explicitly expresses the benefits of our strategy.

Figure 69.1 shows that carry trade returns are negatively related to the EGARCH volatility measure. Base on the threshold value, we can precisely capture the 1997–1998 Asian financial crisis and the 2007–2008 American subprime risk. The volatility level is far above our threshold during these tumultuous episodes. In addition, we find that the carry trade usually has a positive return during the low volatility regime.

69.4 Conclusion

The literature provides several explanations for why the carry trade has persistent excess returns including time-varying risk premia, illiquidity spirals, the peso problem, and trading behavior. Regime-dependent models, such as the regime-switch model and the LSTR model, are commonly used for currency markets. This study employs characteristics of the carry trade to build a new trading strategy that can earn higher returns with lower volatility. We use three measures of volatility, VIX, GARCH, and EAGRCH, to capture the volatility of carry trade returns.

Our main results are threefold. First, we find that carry trade returns have a significant momentum effect at the low volatility regime. Second, the results show that although the buy-and-hold strategy often has higher returns, our strategy effectively avoids losses at the high volatility regime. Third, carry trade returns often suffer higher losses at the high volatility regime. Compared with the buy-and-hold strategy, our method thus bears less downside risk. We use the Sortino ratio criteria to provide this evidence.

Questions remain to be addressed. The United States is maintaining a low interest rate policy right now. Is the US dollar the new carry trade currency? This and other questions are worth discussing. For example, future research may be usefully directed at incorporating such trading strategies with more conventional models of regime dependence. The emerging and developing currency market has dissimilar volatility structures, which may be another issue discussed in the future.

Appendix: Empirical Study Process

1. We find the first, second, and third quartiles from each volatility measures to determine the volatility regimes.
2. We sort the carry trade returns by the volatility regimes.
3. We test the momentum effects for AR(1) model.

4. We use the third quartile of the volatility measures to discriminate returns between high and low volatility regimes.
5. Based on our trading strategy, we calculate the cumulative return and compare the performance with the buy and hold strategy.
6. We use the cumulative returns to calculate the Sortino ratio for each strategy.

These processes can run in the R programming. If you need programming assistance, please contact authors.

References

Albert, J. H., & Chib, S. (1993). Bayes inference via Gibbs sampling of autoregressive time series subject to Markov mean and variance shifts. *Journal of Business and Economic Statistics, 11*(1), 1–15.

Baillie, R. T., & Chang, S. S. (2011). Carry trades, momentum trading and the forward premium anomaly. *Journal of Financial Markets, 14*, 441–464.

Brunnermerier, M. K., Nage, S., & Pedersen, L. H. (2008). Carry trades and currency crashes. Working paper in NBER macroeconomics annual.

Burnside, C., Eichenbaum, M., & Rebelo, S. (2009). Understanding the forward premium puzzle: A microstructure approach. *American Economic Journal: Macroeconomics, American Economic Association, 1*(2), 127–54.

Burnside, C., Eichenbaum, M., Kleshchelski, I., & Rebelo, S. (2011). Do peso problems explain the returns to the carry trade? *Review of Financial Studies*, Oxford University Press for Society for Financial Studies *24*(3), 853–891.

Christiansen, C., & Ranaldo, A. (2011). The time-varying systematic risk of carry trade strategies. *Journal of Financial and Quantitative Analysis, 46*(4), 1107–1125.

Clarida, R., Davis, J., & Pedersen, N. (2009). Currency carry trade regimes: Beyond the Fama regression. *Journal of International Money and Finance, 28*, 1375–1389.

Engel, & Charles. (1984). Testing for the absence of expected real profits from forward market speculation. *Journal of International Economics, 17*, 299–308.

Engel, C., & Hamilton, J. D. (1990). Long swings in the dollar: Are they in the data and do markets know it? *American Economic Review, 80*, 689–713.

Fama, E. F. (1984). Forward and spot exchange rates. *Journal of Monetary Economics, 14*, 319–338.

Froot, K. A., & Thaler, R. H. (1990). Anomalies foreign exchange. *Journal of Economic Perspectives, 4*(3), 179–192.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica, 57*, 357–384.

Ichiue, H., & Koyama, K. (2008). Regime switches exchange rate volatility and uncovered interest parity. Working paper.

Kim, C., & Nelson, C. (1999). Has the U.S. economy become more stable? A Bayesian approach based on a Markov-switching model of the business cycle. *Review of Economics and Statistics, 81*(4), 608–616.

Kim, C., Nelson, C., & Startz, R. (1998). Testing for mean reversion in Heteroskedastic data based on Gibbs-sampling-augmented randomization. *Journal of Empirical Finance, 5*, 131–154.

Meese, R., & Rogoff, K. (1983). Empirical exchange rate models of the seventies. Do they fit out of sample? *Journal of International Economics, 14*, 3–24.

Menkhoff, L., Sarno, L., Schmeling, M., & Schrimpf, A. (2012). Carry trades and global foreign exchange volatility. *Journal of Finance, 67*(2), 681–718.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equations of state calculations by fast computing machines. *Journal of Chemical Physics, 21*, 1087–1092.

Neil, R. (2001). A portfolio manager's view of downside risk. In F. A. Sortino & S. E. Satchell (Eds.), *Managing downside risk in financial markets*. Oxford: Butterworth-Heinemann.

Rogers, D. S., & Van Dyke, C. J. V. (2006). Measuring the volatility of hedge fund returns. *Journal of Wealth Management, 9*(1), 45–53.

Sarno, L., Valente, G., & Leon, H. (2006). Nonlinearity in deviations from uncovered interest parity: An explanation of the forward bias puzzle. *Review of Finance, 10*, 443–482.

Wang, Y. J., Chung, H., & Guo, J. H. (2012). A value-at-risk analysis of carry trades using skew-GARCH models. *Studies in Nonlinear Dynamics & Econometrics*, (in press).

Evaluating the Effectiveness of Futures Hedging

70

Donald Lien, Geul Lee, Li Yang, and Chunyang Zhou

Contents

D. Lien (✉)
University of Texas at San Antonio, San Antonio, TX, USA
e-mail: don.lien@utsa.edu

G. Lee • L. Yang
University of New South Wales, Sydney, Australia
e-mail: geul.lee@student.unsw.edu.au; l.yang@unsw.edu.au

C. Zhou
Shanghai Jiaotong University, Shanghai, China
e-mail: cyzhou@sjtu.edu.cn

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_70,

Abstract

This chapter examines the Ederington hedging effectiveness (EHE) comparisons between unconditional OLS hedge strategy and other conditional hedge strategies. It is shown that OLS hedge strategy outperforms most of the optimal conditional hedge strategies when EHE is used as the hedging effectiveness criteria. Before concluding that OLS hedge is better than the others, however, we need to understand under what circumstances the result is derived. We explain why OLS is the best hedge strategy under EHE criteria in most cases and how most conditional hedge strategies are judged as inferior to OLS hedge strategy by an EHE comparison.

Keywords

Futures hedging • Portfolio management • Ederington hedging effectiveness • Variance estimation • Unconditional variance • Conditional variance • OLS hedging strategy • GARCH hedging strategy • Regime switching hedging strategy • Utility-based hedging strategy

70.1 Introduction

Futures market provides a useful tool for hedgers to reduce the overall risk. The extent of the usefulness is, however, determined by the hedging strategy adopted by the hedger. In this regard, the hedging effectiveness measure proposed by Ederington (1979) has been the most popular criterion to evaluate the usefulness. Different hedging strategies are compared in terms of Ederington hedging effectiveness (EHE). The strategy possessed with the greatest EHE is deemed the best strategy.

Specifically, EHE is the percentage reduction in the return variance of the hedged portfolio relative to the return variance of the unhedged portfolio. While the variance could be conditional or unconditional, in empirical studies EHE is always calculated on the basis of unconditional variance. This is natural as Ederington (1979) considers only unconditional constant hedge strategies. Further development in futures hedging literature focuses on conditional dynamic hedge strategies. However, EHE remains the major criterion to evaluate the usefulness of these strategies. This approach is inappropriate since the conditional hedge strategy is constructed to minimize conditional variance, but its usefulness is measured by unconditional variance. As long as there is not a linear relationship between conditional and unconditional variances, the EHE should not serve as a benchmark to evaluate the conditional hedge strategy.

This paper examines the EHE comparisons between the OLS hedge strategy (i.e., the unconditional strategy) with various conditional hedge strategies, assuming spot and futures returns are described by different statistical framework. It is shown that, for most statistical models, the OLS hedge strategy is most likely to outperform the optimal conditional hedge strategy. For example, in a vector error correction model (VECM), the optimal conditional hedge ratio should take into account the cointegration relationship. The resulting EHE from this optimal ECM hedge ratio, however, underperforms the OLS hedge ratio (where the cointegration relationship is ignored). Similarly, when spot and futures returns follow a multivariate generalized autoregressive conditional heteroskedasticity (MGARCH) model, the GARCH hedge ratio is likely to be inferior to the OLS hedge ratio in terms of EHE.

The above results are not surprising as OLS hedge ratio is chosen to minimize the unconditional variance, whereas ECM and GARCH hedge ratios minimize their corresponding conditional variances. By definition, EHE is biased in favor of OLS hedge ratio over other hedge ratios. The only possible exception is the regime switching (RS) hedge ratio. It is analytically shown that RS-OLS hedge ratio would outperform the conventional OLS hedge ratio under certain assumptions.

Besides EHE, another popular hedging effectiveness measure is certainty equivalent derived from expected utility comparisons from hedged and unhedged portfolios. It is shown that the sample certainty equivalent estimator, similar to the sample EHE estimator, is biased. On the other hand, this utility-based effectiveness measure does not necessarily favor the OLS hedge ratio except when the futures price is a martingale or when the hedger is extremely risk averse.

The remaining of the paper is organized as follows. In Sect. 70.2, we discuss the Ederington hedging effectiveness measure and demonstrate the superiority of the OLS hedge ratio. The next sections consider two specific dynamic hedging strategies; Sect. 70.3 examines the GARCH specifications and Sect. 70.4 the regime switching models. The two models provide contradicting conclusions regarding the relative performance to the OLS hedge ratio. In Sect. 70.5, we analyze the utility-based hedging effectiveness. Finally, conclusions are provided in Sect. 70.6.

70.2 Ederington Hedging Effectiveness

The fundamental idea of Ederington (1979) originates from Johnson (1960) and Stein (1961) who introduce portfolio theory into the area of hedging. Most of the previous hedging theories consider only the "naive" hedging, which is done by trading the hedging instrument in the same amount as the asset being hedged. Ederington shows that the hedge ratio, which is the ratio of the amount of the hedging instrument being used relative to the amount of the asset being hedged, must be adjusted to obtain the maximum hedging effectiveness. To derive this result, Ederington proves that there exists an optimal hedge ratio which minimizes the variance of the portfolio value.

70.2.1 Definition

Ederington shows that if we construct a hedged portfolio P which consists of the asset being hedged, S, and a hedging instrument, F, the optimal hedge ratio is the value where the partial derivative of the portfolio return variance with respect to the hedge ratio becomes zero. This partial derivative is given by

$$\frac{\partial Var(p)}{\partial h} = X_s^2[2hVar(f) - 2Cov(s,f)], \tag{70.1}$$

where X_s and X_f are the positions of the asset and the hedging instrument, respectively; $h = -X_f/X_s$ is the hedge ratio; s and f are the returns of S and F, respectively; p is the return of the portfolio: $p = X_s s + X_f f = X_s(s - hf)$. Var(.) and Cov(.,.) are the variance and covariance operators, respectively. Given this formula, the optimal hedge ratio h^* can be easily derived by setting Eq. 70.1 equal to zero, i.e.,

$$h^* = Cov(s,f)/Var(f). \tag{70.2}$$

As shown in Eqs. 70.1 and 70.2, Ederington regards hedging as an act of “minimizing variance.” When devising his measure of hedging effectiveness, he also takes this property as the main criteria. Specifically, the Ederington hedging effectiveness (EHE hereafter) is defined as

$$H = 1 - \frac{Var(p)}{Var(s)}. \tag{70.3}$$

Equation 70.3 shows that EHE is directly related to the percentage reduction of the variance in the asset return after hedging.

70.2.2 Some Properties

The most evident characteristic of EHE is its simplicity. The variance and covariance in Eq. 70.3 are both unconditional and are assumed to be constant over time. While this aspect of EHE is one of the reasons why it is being widely used, it also has caused some controversies about its appropriateness. It was argued that, given the information set, the hedger is concerned with the conditional variance of the portfolio return. Accordingly, unconditional variance and covariance in Eq. 70.3 should be replaced by their conditional counterparts. Various variables were taken into account to derive conditional variance and covariance, including past prices and inventories. In addition, recent research emphasizes the nonconstancy nature of conditional variance and covariance and recommends time-varying hedge ratios.

In empirical implementation, the complete sample is divided into two subsamples. The first subsample is applied to construct the most appropriate (within-sample) statistical models for conditional variance and covariance. Based upon the estimated model, optimal hedge ratios are obtained for the second subsample. Returns for the hedged portfolio are calculated for this subsample. The unconditional variance of the return series is adopted to calculate the so-called post-sample

EHE which serves as a benchmark to compare various hedge strategies. Thus, the within-sample model is chosen to minimize conditional variance, whereas the post-sample hedging effectiveness is evaluated at unconditional variance. The inconsistency in criteria raises a concern for the appropriateness of EHE to be used as the criteria to compare conditional hedge strategies. Moreover, one would suspect the hedge strategy constructed by minimizing the within-sample unconditional variance may have the best out-of-sample EHE.

To address this question, Lien (2005a) introduces several assumptions:

1. The size of the estimation sample is sufficiently large.
2. The size of the evaluation sample is sufficiently large.
3. There is no structural change between the estimation sample and the evaluation sample.

Under these conditions, it is shown that the ratio of unconditional covariance to the conditional variance provides the best EHE. This ratio can be obtained by the ordinary least squares (OLS) method when regressing the spot price change on the futures price change.

The first two assumptions ensure sample unconditional variance and sample unconditional covariance both to be close enough to their population counterparts. The third assumption requires the estimation and evaluation samples to be drawn from the same population such that the hedge ratio derived from the former is applicable to the latter. When there is a structural change across the two samples, nothing can be guaranteed. However, Lien (2005a, b) warns that it is a tautology to prove the superiority of OLS hedge ratio to other hedge ratios with EHE. Since the OLS hedge ratio is the hedge ratio which produces the minimal unconditional variance, it cannot be inferior to any other hedge ratios when compared in terms of EHE, which measures the unconditional variance reduction. Alternatively, one can argue that it is not appropriate to compare conditional hedge strategies on the basis of EHE.

Lien (2005b) illustrates the superiority of the OLS hedge ratio with a simple example:

$$f_t = \alpha_0 + \alpha_1 f_{t-1} + u_t, \tag{70.4}$$

$$s_t = \beta_0 + \beta_1 s_{t-1} + v_t, \tag{70.5}$$

where both $\{u_t\}$ and $\{v_t\}$ are white noises. Let σ_u^2 and σ_v^2 denote the variances of u_t and v_t, respectively, and let σ_{uv} denote the covariance between u_t and v_t. In addition, to ensure stationarity, we require $|\alpha_1| < 1$ and $|\beta_1| < 1$. The unconditional hedge ratio (i.e., OLS hedge ratio) is

$$h_u = \left(\frac{1-\alpha_1^2}{1-\alpha_1\beta_1}\right)\left(\frac{\sigma_{uv}}{\sigma_u^2}\right), \tag{70.6}$$

whereas the conditional hedge ratio is simply $h_c = \sigma_{uv}/\sigma_u^2$. By construction, h_u performs better than h_c in terms of EHE.

Upon incorporating the cointegration relationship between spot and futures prices into Eqs. 70.4 and 70.5, Lien (2005a) demonstrates the superiority of the OLS hedge ratio over the error correction hedge ratio. For the importance of the error correction term for futures hedging and further comparisons between the two hedge ratios, see Lien (1996, 2004).

70.2.3 Estimation Bias

In Lien (2006), it is shown that the usual EHE estimator is downward biased and, therefore, tends to underestimate the true hedging performance, even when the estimator for optimal hedge ratio is unbiased. This is because the estimated optimal hedge ratio itself is a random variable so that its variance affects the expected EHE. Lien explains this by decomposing the EHE formula (70.3) as follows.

Let $M = I_k - (e_k e_k^{'}/k)$ where I_k is a $(k \times k)$-dimensional identity matrix and e_k is a k-dimensional vector such that all elements are equal to 1. Then Eq. 70.3 can be decomposed to

$$H = 1 - \frac{w^{'}Mw}{p^{'}Mp}, \tag{70.7}$$

where p and w are k-dimensional vectors consisting of k unhedged asset returns and hedged portfolio returns, respectively. Because the estimated hedge ratio $\hat{h}$ substitutes the optimal hedge ratio h^*, the EHE one calculates based on $\hat{h}$ is also in fact an estimated EHE, $\hat{H}$. That is,

$$\hat{H} = 1 - \frac{\hat{w}^{'}M\hat{w}}{p^{'}Mp}, \tag{70.8}$$

where $\hat{w}$ is a k-dimensional vector consisting of the portfolio returns which are hedged with $\hat{h}$. Since $\hat{w} = w + \left(h^* - \hat{h}\right)f$, Eq. 70.7 can be rewritten as

$$1 - \hat{H} = \frac{w^{'}Mw}{p^{'}Mp} + 2\left(h^* - \hat{h}\right)\left[\frac{f^{'}Mw}{p^{'}Mp}\right] + \left(h^* - \hat{h}\right)^2\left[\frac{f^{'}Mf}{p^{'}Mp}\right], \tag{70.9}$$

and therefore,

$$\hat{H} = H - 2\left(h^* - \hat{h}\right)\left[\frac{f^{'}Mw}{p^{'}Mp}\right] - \left(h^* - \hat{h}\right)^2\left[\frac{f^{'}Mf}{p^{'}Mp}\right]. \tag{70.10}$$

As a consequence,

$$E(\hat{H}) = H + 2b\left[\frac{f'Mw}{p'Mp}\right] - \left[b^2 + Var(\hat{h})\right]\left[\frac{f'Mf}{p'Mp}\right], \tag{70.11}$$

where $b = E(\hat{h} - h^*)$ and $b^2 + Var(\hat{h})$ are the estimation bias and mean squared error of $\hat{h}$, respectively. If $\hat{h}$ is an unbiased estimator of h^*, b becomes zero and Eq. 70.11 is reduced to

$$E(\hat{H}) = H - Var(\hat{h})\left[\frac{f'Mf}{p'Mp}\right]. \tag{70.12}$$

This shows that $\hat{H}$ is a downward biased estimator of H, even when $\hat{h}$ is an unbiased estimator of h^*.

We provide two further remarks. First, Chen and Sutcliffe (2007) examine the benefits of a composite hedge where multiple hedging instruments are adopted over a simple hedge where only one hedging instrument is adopted. The benefit is measured by the improvement in EHE. Lien (2008) demonstrates the empirical estimator is biased. Secondly, through empirical studies, Lien and Shreshta (2008) concludes that the downward bias of the EHE estimator is negligible and therefore bias correction seems to be redundant.

70.3 GARCH Hedging Strategy

The previous analysis assumes the conditional second moments of spot and futures returns are constant over time. This assumption is frequently rejected through empirical data analysis. To describe time-varying second moments, researchers rely upon different versions of multivariate GARCH (generalized autoregressive conditional heteroskedasticity) models. The standard univariate GARCH model is an extension of the ARCH (autoregressive conditional heteroskedasticity) model proposed by Engle (1982).

70.3.1 GARCH Specification

Specifically, consider a time series $\{y_t\}$ such that

$$y_t = b'x_t + \varepsilon_t, \tag{70.13}$$

where x_t is the vector of exogenous variables contained in the information set previous to time $t-1$, ϑ_{t-1}. The error term is normally distributed conditional on the information set, i.e., $\varepsilon_t|\vartheta_{t-1} \sim N(0, \sigma_t^2)$. Bollerslev (1986) proposed the following process for the conditional variance:

$$\sigma_t^2 = \omega + \sum_{i=1}^{q} \alpha_i \varepsilon_{t-i}^2 + \sum_{j=1}^{p} \beta_j \sigma_{t-j}^2. \tag{70.14}$$

The model is termed as GARCH (p, q) process. If we set $p = 0$ and $q \neq 0$, the process is reduced to an ARCH(q) process. When $p = q = 0$, it is further reduced to a process with a constant variance.

For futures hedging purpose, we need to consider multivariate GARCH models which specify the process for the conditional covariance as well. Various specifications are available in the literature, e.g., constant correlation, BEKK, and DCC (dynamic conditional correlation) models. Different dynamic hedge ratios are then generated and compared on the basis of EHE; for example, see Baillie and Myers (1991), Myers (1991), Kroner and Sultan (1993), Dawson et al. (2000), and Kavussanos and Visvikis (2008).

70.3.2 GARCH Hedging Strategy

Under GARCH models, the optimal hedge ratio is determined by the ratio of the conditional covariance to the conditional variance,

$$h_{t-1}^c = \frac{Cov_{t-1}(s_t, f_t)}{Var_{t-1}(f_t)}, \tag{70.15}$$

where $Cov_{t-1}(s_t, f_t)$ is the conditional covariance between spot and futures returns at time t based upon information available at time $t-1$ and $Var_{t-1}(f_t)$ is the conditional variance of the futures return at time t based upon information available at time $t-1$. As both conditional moments are time varying, the conditional hedge ratio is expected to change over time as well.

That is, although OLS and GARCH hedge ratios have the same object of variance minimization, they differ in terms of the target variance. While OLS hedge ratio considers the unconditional variance, GARCH hedge ratio focuses on conditional variance under the GARCH assumptions. This difference suggests a concern about the appropriate procedure of assessing and comparing their effectiveness. Since their objectives are different, the relative superiority of one hedge ratio over the other can vary when one applies a different effectiveness measure. In particular, since EHE depends upon the reduction in the unconditional variance, OLS hedge ratio is naturally favored. Lien (2009) explains this result as follows.

70.3.3 EHE and GARCH Hedging Strategy

Let us assume that there are two portfolios P0 and P1, both consisting of an asset S and a hedging instrument F. The first portfolio is constructed from the OLS hedge ratio, $h_0 = Cov(s_t,f_t)/Var(f_t)$, whereas the second portfolio is constructed from the GARCH hedge ratio, $h^c_{t-1} = Cov_{t-1}(s_t,f_t)/Var_{t-1}(f_t)$. We can decompose the unconditional variance of the return from portfolio P1 as

$$Var(P1) = Var\left(s_t - h^c_{t-1}f_t\right) = E\left[Var_{t-1}\left(s_t - h^c_{t-1}f_t\right)\right] + Var\left[E_{t-1}\left(s_t - h^c_{t-1}f_t\right)\right], \tag{70.16}$$

where $E(.)$ is the unconditional expectation operator and $E_{t-1}(.)$ is the conditional expectation operator based upon information available at time $t-1$. We can rewrite the first term of Eq. 70.16 as follows:

$$E\left[Var_{t-1}(s_t) - \frac{Cov^2_{t-1}(s_t,f_t)}{Var_{t-1}(f_t)}\right] = Var(s_t) - E\left[\frac{Cov^2_{t-1}(s_t,f_t)}{Var_{t-1}(f_t)}\right], \tag{70.17}$$

using the definition of h^c_{t-1}. Suppose that the sample size is sufficiently large, we can approximate the second term of Eq. 70.17 by

$$E\left[\frac{Cov^2_{t-1}(s_t,f_t)}{Var_{t-1}(f_t)}\right] \approx \frac{E\left[Cov^2_{t-1}(s_t,f_t)\right]}{E[Var_{t-1}(f_t)]} = \frac{Cov^s(s_t,f_t)}{Var(f_t)}. \tag{70.18}$$

Consequently,

$$Var(P1) \approx Var(s_t) - \frac{Cov^2(s_t,f_t)}{Var(f_t)} + Var[E_{t-1}(s_t - h_{t-1}f_t)]. \tag{70.19}$$

On the other hand, by the definition of h_0, the unconditional variance of the return from portfolio P0 is

$$Var(P0) = Var(s_t - h_0 f_t) = Var(s_t) - \frac{Cov^2(s_t,f_t)}{Var(f_t)}. \tag{70.20}$$

Therefore,

$$Var(P1) \approx Var(P0) + Var[E_{t-1}(s_t - h_{t-1}f_t)], \tag{70.21}$$

implying $Var(P1)$ tends to be larger than $Var(P0)$. That is, the OLS hedge ratio is likely to have a greater hedging effectiveness than the GARCH hedge ratio, in terms of EHE.

Note that the derivation of Eqs. 70.16, 70.17, 70.18, 70.19, 70.20, and 70.21 does not rely on any specific properties of the GARCH model. The above conclusion, therefore, applies to any general dynamic hedge strategy that aims at minimizing the conditional variance; see also Lien (2010). In other words, when adopting EHE as the effectiveness measure, the OLS hedge ratio is likely to outperform any dynamic hedge ratio. However, we should be careful when interpreting this result. As Lien (2005a) points out, EHE is focused on the unconditional variance and it would be an abuse to use EHE to assess a conditional variance minimization strategy.

Kavussanos and Nomikos (2000) suggest that for the GARCH hedge strategy to outperform the OLS hedge strategy, the variability of the resulting GARCH ratio must be sufficiently large. On the other hand, Park and Jei (2010) find an inverse relationship between the variability of the GARCH hedge ratio and corresponding hedging effectiveness (i.e., EHE).

70.4 Regime Switching Hedging Strategy

Lien (2010) provides a theoretical analysis on the relationship between the variability of the hedge ratio and hedging performance in support of the finding from Park and Jei (2010). Extending the result to general dynamic hedge strategy, there is a small window for the strategy to outperform the OLS strategy, that is, when the variability of the hedge ratio cannot be too small or too large. We therefore turn to regime switching hedge strategies.

70.4.1 Definition of Regime Switching

Both GARCH and regime switching models belong to the family of nonlinear time series. Hamilton (1988, 1989) characterizes the concept of "regime switching" (RS hereafter) and proposes an approach to model the RS process. The simplest RS model specification of RS is the first-order Markov process with two states. If $S_t \in \{0, 1\}$ denotes the (not directly observable) state of the system in which the source of the time-series data exists, the transition between two states is driven by the following first-order Markov process:

$$\begin{aligned}
&\Pr\left(S_t = 1 | S_{t-1} = 1\right) = p,\\
&\Pr\left(S_t = 0, | S_{t-1} = 1\right) = 1 - p,\\
&\Pr\left(S_t = 0 | S_{t-1} = 0\right) = q,\\
&\Pr\left(S_t = 1, | S_{t-1} = 0\right) = 1 - q.
\end{aligned} \tag{70.22}$$

Thus, the probability of state transition depends only upon the state of the previous period.

In each state, the spot and futures returns can be described by linear models such as ECM or nonlinear models such as GARCH processes. For the former case, there will be two constant hedge ratios each pertaining to one state; for the latter case, there will be two dynamic hedge ratios instead. The literature on RS hedge strategies began with the former case and recently extended to the latter case. For example, Sarno and Valente (2000) and Alizadeh et al. (2008) combine RS with ECM; Alizadeh and Nomikos (2004) and Lee and Yoder (2007a, b) add RS into the GARCH models; Lee (2010) combines RS with dynamic conditional correlation (DCC) models. In most cases, it is shown that the hedging performance is improved when regime switching is incorporated into the econometric framework.

70.4.2 RS Hedging Strategy

Although RS can be introduced in various ways, it is most understandable when we combine RS with the OLS hedging strategy. Lien (2012b) explains the basic framework of the RS-OLS strategy as follows. Suppose that RS process is given as Eq. 70.22. Assume futures returns in the two states equal to each other. When $S_{t-1} = 1$, the OLS hedge ratio in Eq. 70.2 is modified to

$$h_1^* = \frac{pCov_1(s_t, f_t) + (1-p)Cov_0(s_t, f_t)}{pVar_1(f_t) + (1-p)Var_0(f_t)}, \tag{70.23}$$

where $Var_n(.)$ and $Cov_n(.,.)$ denote the variance and covariance operators in state n, respectively, $n = 0, 1$. Similarly, when $S_{t-1} = 0$, the corresponding OLS hedge ratio is

$$h_0^* = \frac{qCov_0(s_t, f_t) + (1-q)Cov_1(s_t, f_t)}{qVar_0(f_t) + (1-q)Var_1(f_t)}. \tag{70.24}$$

The pair of hedge ratios (h_0^*, h_1^*) constitutes the optimal RS-OLS hedge ratio. To apply this hedge strategy, it requires the hedger to be able to identify the state at the moment of making the hedging decision.

Lien (2012b) compares the RS-OLS hedge strategy to the conventional OLS hedge strategy. To calculate the conventional OLS hedge ratio under the RS framework, we first derive the steady-state probability for each state. Let α and $1 - \alpha$ denote the steady-state probability of state 1 and 0, respectively. Thus,

$$(1-\alpha)p + \alpha(1-q) = \alpha, \tag{70.25}$$

or equivalently

$$\alpha = \frac{1-q}{2-p-q}. \tag{70.26}$$

Given the steady-state probability of each state, we can obtain the conventional OLS hedge ratio as follows:

$$h^* = \frac{\alpha Cov_1(s_t, f_t) + (1-\alpha)Cov_0(s_t, f_t)}{\alpha Var_1(f_t) + (1-\alpha)Var_0(f_t)}. \tag{70.27}$$

Let $h^*_{RS} = \alpha h^*_1 + (1-\alpha)h^*_0$, the expected RS-OLS hedge ratio. Lien (2012b) shows the expected RS-OLS hedge ratio exceeds the conventional OLS hedge ratio:

$$h^*_{RS} \geq h^*. \tag{70.28}$$

Thus, more transaction cost is incurred when implementing the RS-OLS hedge strategy.

70.4.3 Hedging Effectiveness

To compare the hedging effectiveness, let $V(h)$ denote the variance of the return from the hedged portfolio, where $h = h^*_0$, h^*_1, or h^*. The expected variance of the RS-OLS hedged portfolio is then $V_{RS} = \alpha V(h^*_1) + (1-\alpha)V(h^*_0)$. Lien (2012b) demonstrates that

$$V_{RS} \leq V(h^*), \tag{70.29}$$

that is, the RS-OLS hedged portfolio has a smaller variance than the conventional OLS hedged portfolio. Consequently, the RS-OLS strategy outperforms the OLS strategy in terms of EHE.

While the RS-OLS seems to be very promising, a serious problem with this result is that, as Lien (2012a) points out, the superiority of the RS-OLS strategy is based on the assumption that a hedger can always correctly identify the prevailing state at the decision time correctly. To successfully conduct the above hedging strategy, we must succeed in at least three tasks to complete the correct identification:

1. We must identify the entire set of possible states.
2. We must identify the prevailing state.
3. We must identify the relationship between spot and futures returns in each state.

In reality, it is unlikely to complete any of these tasks without errors. Hamilton (1989) is well aware of these issues and emphasizes the importance of "optimal probabilistic inference" to find the turning points. One may try to go around this problem by a weighted average strategy such that the optimal hedge ratio is chosen to be

$$\hat{h}^* = \beta h^*_1 + (1-\beta)h^*_0, \tag{70.30}$$

where β is the estimated probability that the prevailing state is state 1. However, as Lien (2012b) points out, this again dilutes the relative superiority of RS-OLS strategy, since at least one of the states is false at any time t. We can conclude that, therefore, one must succeed in the structural definition of possible states and correct identification of the current state to fully take advantage of the RS framework.

70.5 Utility-Based Hedging Effectiveness

Up to now, we assume that the sole objective of hedging is variance reduction, and correspondingly the optimal hedge ratio is the one that minimizes variance. This is quite intuitive because risk minimization is the most important reason that hedging is actually being done. In the real world, however, the variance-minimizing hedge ratio is not always the optimal one. To understand why this is true, we must know that there are some other factors than variance minimization about which a hedger should consider. For example, if a hedger assumes that a price process is sub-martingale and wants to take advantage of positive expected return, he or she will try to afford some risk by a non-perfect hedging. In this situation, variance reduction cannot be the perfect measure for hedging effectiveness.

70.5.1 Definition of Utility-Based Hedging

Given these restrictions, we can adopt a multivariable function as the alternative and consider additional factors other than variance to measure the hedging effectiveness. In particular, we can consider how large the expected return will be after hedging cost is offset, how much risk a hedger can afford to retain a certain amount of expected return, as well as how large the variance will be. Many previous studies, e.g., Kroner and Sultan (1993), Gagnon et al. (1998), Follmer and Leukert (1999, 2000), and Monoyios (2004), introduce the idea of utility function to construct a framework for this multivariate relationship.

A basic framework of utility-based hedging effectiveness measure is provided in Lien (2012a). Consider a two-date one-period model. The expected utility of an unhedged portfolio can be defined as

$$E\left[U\left(w_{1,u}\right)\right] = E[U(w_0 + s_1 - s_0)], \tag{70.31}$$

where w_0 is the initial wealth (i.e., the wealth at time 0), s_1 is the random value of the spot asset at time 1, s_0 is the value of spot asset at time 0, and $w_{1,u}$ is the random value of wealth at time 1 when there is no hedging conducted. If we adopt a hedging strategy, the expected utility of the hedged portfolio is

$$E\left[U\left(w_{1,h}\right)\right] = E[U(w_0 + s_1 - s_0 - h(f_1 - f_0))], \tag{70.32}$$

where f_1 is the random value of the hedging instrument at time 1, f_0 is the value of the hedging instrument at time 0, and $w_{1,\ h}$ is the random value of wealth at time 1 when hedging is conducted. Hedging performance is measured by the certainty equivalent C:

$$E\left[U\left(w_{1,u}+C\right)\right]=E\left[U\left(w_{1,h}\right)\right]. \tag{70.33}$$

70.5.2 Utility Function and Risk Aversion

One of the simplest types of utility function, which can be used as a hedging effectiveness measure, is the expected mean-variance utility function. It is also quite popular since it can consider all the factors above within a simple framework; see, for example, Kroner and Sultan (1993), Gagnon et al. (1998), and Lafuente and Novales (2003). Suppose that a hedger is endowed with a strictly increasing and twice-differentiable concave utility function $U(x)$, such that $U'(x) > 0$ and $U''(x) < 0$. Then the expected utility of the hedged portfolio P at time $t-1$ can be defined as

$$E_{t-1}[U(P)]=E_{t-1}(p_t)-\lambda Var_{t-1}(p_t), \tag{70.34}$$

where p is the return of the portfolio P and λ is a positive risk-aversion parameter.

The existence of the risk aversion parameter is suggested by Merton (1973). Chou (1988) explains that there exists a linear relationship between the equity premium π and return variance in the inter-temporal CAPM model of Merton (1973), such that

$$\pi_t=\lambda_m Var(M_t), \tag{70.35}$$

where M_t is the instantaneous market return and λ_m is the harmonic mean of individual investor's risk-aversion parameter. Various studies, e.g., Grossman and Shiller (1981) and Pindyck (1986), show that the idea of premium can explain much of the stock price changes beyond changes in dividends and interest rates. Also, their estimation results show that λ ranges approximately from 3 to 4.5.

One thing we should note is that the estimation of λ relies on the variance estimation method. Poterba and Summers (1986) employ a two-stage OLS procedure to estimate the variance and conclude that shocks to the volatility decay rapidly so that it is skeptical to claim that fluctuations in risk premia account for much of the variation in prices. On the other hand, Chou (1988) introduces GARCH-M model and argues that the persistence of volatility shocks is significant such that fluctuations in risk premia can explain much of the price changes. Given that the other aspects of both researches are quite similar, this observation implies that different variance estimation method will lead to different estimates for λ.

70.5.3 Utility-Based Hedging Effectiveness

Similar to the EHE case, Lien (2012a) shows that the sample utility-based hedging effectiveness estimator is downward biased and therefore tends to underestimate the true hedging performance, even when the sample estimator for the optimal hedge ratio is unbiased. To explain in detail why this happens, Lien first assumes that a hedger is endowed with a mean-variance expected utility function, i.e., Eq. 70.34. Given Eqs. 70.33 and 70.34, we obtain

$$E\left[U\left(w_{1,u}+C\right)\right]=w_0+C+E(\Delta s)-\lambda Var(\Delta s), \tag{70.36}$$

$$E\left[U\left(w_{1,h}\right)\right]=w_0+E(\Delta s-h\Delta f)-\lambda Var(\Delta s-h\Delta f), \tag{70.37}$$

where $\Delta s=s_1-s_0$ and $\Delta f=f_1-f_0$. From the above two equations, we derive

$$C=-hE(\Delta f)+\lambda[Var(\Delta s)-Var(\Delta s-h\Delta f)]. \tag{70.38}$$

The sample estimator of C is then

$$\hat{C}=-\hat{h}E(\Delta f)+\lambda\left[Var(\Delta s)-Var\left(\Delta s-\hat{h}\Delta f\right)\right]. \tag{70.39}$$

From Eqs. 70.38 and 70.39, we obtain

$$\hat{C}=C-\left(h-\hat{h}\right)E(\Delta f)..+\lambda\left[Var(\Delta s-h\Delta f)-Var\left(\Delta s-\hat{h}\Delta f\right)\right]. \tag{70.40}$$

After algebraic manipulations, Eq. 70.40 becomes

$$\hat{C}=C-\left(h-\hat{h}\right)E(\Delta f)+2\lambda\left(\hat{h}-h\right)Cov(\Delta s,\Delta f)-\lambda\left(\hat{h}^2-h^2\right)Var(\Delta f). \tag{70.41}$$

Suppose that $\hat{h}$ is an unbiased estimator of h, i.e., $E\left(\hat{h}\right)=h$, then

$$E\left(\hat{C}\right)=C-\lambda E\left(\hat{h}^2-h^2\right)Var(\Delta f)=C-\lambda Var\left(\hat{h}\right)Var(\Delta f)<C. \tag{70.42}$$

That is, the expected value of $\hat{C}$ is downward biased. Lien (2012a) shows that the downward bias result can be extended to the case when a hedger is endowed with another type of strictly increasing concave utility function.

Because the certainty equivalent is not a strictly monotonically decreasing function of the portfolio variance (except when $E\ (\Delta f)=0$ or when λ is infinitely large), the solution to variance minimization is not the same as the solution to certainty equivalent maximization. Therefore, OLS hedge ratio is not necessarily favored by the utility-based performance measure.

70.6 Conclusions

This paper analyzes the properties of Ederington hedging effectiveness (EHE) in general and within different statistical framework. The most popular EHE is measured by the percentage reduction in the unconditional return variance of the hedged portfolio relative to the unconditional return variance of the unhedged portfolio. Because of this emphasis on "unconditional" statistics, the OLS hedge strategy (which does not take into account any other information except current spot and futures returns) is most likely to outperform the optimal conditional hedge strategy.

This superiority of the OLS hedge ratio is challenged by the concern of the appropriateness to evaluate conditional hedging strategies by EHE. Nonetheless, the regime switching (RS) hedge ratio seems to be an exception. Under specific assumptions, the RS-OLS hedge ratio will outperform the conventional OLS hedge ratio.

Utility-based hedging effectiveness is another popular measure examined in the literature. The sample estimator of this effectiveness is, similar to the sample EHE estimator, biased. On the other hand, the measure does not necessarily favor the OLS hedge ratio except when the futures price is a martingale or when the hedger is extremely risk averse.

Recently there have been several alternative effectiveness measures related to tail risk such as lower partial moment, value at risk, and conditional value at risk. Similar problems prevail. That is, a hedge strategy may be chosen to minimize the conditional value at risk. However, when in the evaluation stage, it is the unconditional value at risk that counts. We do not address these issues in the current paper. It will be left for future research.

References

Alizadeh, A. H., & Nomikos, N. (2004). A Markov regime switching approach for hedging stock indices. *Journal of Futures Markets, 24*, 649–674.

Alizadeh, A. H., Nomikos, N., & Pouliasis, P. K. (2008). A Markov regime switching approach for hedging energy commodities. *Journal of Banking and Finance, 32*, 1970–1983.

Baillie, R. T., & Myers, R. (1991). Bivariate GARCH estimation of the optimal commodity futures hedge. *Journal of Applied Econometrics, 6*, 109–124.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics, 31*, 307–327.

Chen, F., & Sutcliffe, C. (2007). *Better cross hedge with composite hedging? Hedging equity portfolios using financial and commodity futures*. ICMA discussion papers in finance DP2007-04. University of Reading.

Chou, R. Y. (1988). Volatility persistence and stock valuations: Some empirical evidence using GARCH. *Journal of Applied Econometrics, 3*, 279–294.

Dawson, P. J., Tiffin, A. L., & White, B. (2000). Optimal hedging ratios for wheat and barley at the LIFFE: A GARCH approach. *Journal of Agricultural Economics, 51*, 47–161.

Ederington, L. H. (1979). The hedging performance of the new futures markets. *Journal of Finance, 34*, 157–170.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of variance of United Kingdom inflation. *Econometrica, 50*, 987–1008.
Follmer, H., & Leukert, P. (1999). Quantile hedging. *Finance and Stochastics, 3*, 251–273.
Follmer, H., & Leukert, P. (2000). Efficient hedging: Cost versus shortfall risk. *Finance and Stochastics, 4*, 117–146.
Gagnon, L., Lypny, G. J., & McCurdy, T. H. (1998). Hedging foreign currency portfolios. *Journal of Empirical Finance, 5*, 197–220.
Grossman, S. J., & Shiller, R. J. (1981). The determinants of the variability of stock market prices. *American Economic Review, 71*, 222–227.
Hamilton, J. D. (1988). Rational-expectations econometric analysis of changes in regime: An investigation of the term structure of interest rates. *Journal of Economic Dynamics and Control, 12*, 385–423.
Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica, 57*, 357–384.
Johnson, L. L. (1960). The theory of hedging and speculation in commodity futures. *Review of Economic Studies, 27*, 139–151.
Kavussanos, M. G., & Nomikos, N. K. (2000). Futures hedging when the structure of the underlying asset changes: The case of BIFFEX contracts. *Journal of Futures Markets, 20*, 775–801.
Kavussanos, M. G., & Visvikis, I. D. (2008). Hedging effectiveness of the Athens stock index futures contracts. *European Journal of Finance, 14*, 243–170.
Kroner, K. F., & Sultan, J. (1993). Time-varying distributions and dynamic hedging with foreign currency futures. *Journal of Financial and Quantitative Analysis, 28*, 535–551.
Lafuente, J., & Novales, A. (2003). Optimal hedging under departures from the cost-of-carry valuation: Evidence from the Spanish stock index futures markets. *Journal of Banking and Finance, 27*, 1053–1078.
Lee, H.-T. (2010). Regime switching correlation hedging. *Journal of Banking and Finance, 34*, 2728–2741.
Lee, H.-T., & Yoder, J. K. (2007a). A bivariate Markov regime switching GARCH approach to estimate time varying minimum variance hedge ratios. *Applied Economics, 39*, 1253–1265.
Lee, H.-T., & Yoder, J. K. (2007b). Optimal hedging with a Markov regime switching time-varying correlation. *Journal of Futures Markets, 27*, 495–516.
Lien, D. (1996). The effect of cointegration relationship on futures hedging: A note. *Journal of Futures Markets, 16*, 773–780.
Lien, D. (2004). Cointegration and the optimal hedge ratio: The general case. *Quarterly Review of Economics and Finance, 44*, 654–658.
Lien, D. (2005a). A note on the superiority of the OLS hedge ratio. *Journal of Futures Markets, 25*, 1121–1126.
Lien, D. (2005b). The use and abuse of the hedging effectiveness measure. *International Review of Financial Analysis, 14*, 277–282.
Lien, D. (2006). Estimation bias of futures hedging performance: A note. *Journal of Futures Markets, 26*, 835–841.
Lien, D. (2008). A note on estimating the benefit of a composite hedge. *Journal of Futures Markets, 28*, 711–716.
Lien, D. (2009). A note on the hedging effectiveness of GARCH models. *International Review of Economics and Finance, 18*, 110–112.
Lien, D. (2010). A note on the relationship between the variability of the hedge ratio and hedging performance. *Journal of Futures Markets, 30*, 1100–1104.
Lien, D. (2012a). A note on the performance of regime switching hedge strategy. *Journal of Futures Markets, 32*, 389–396.
Lien, D. (2012b). A note on utility-based futures hedging performance measure. *Journal of Futures Markets, 32*, 92–97.

Lien, D., & Shrestha, K. (2008). Hedging effectiveness comparisons: a note. *International Review of Economics and Finance, 17*, 391–396.

Merton, R. C. (1973). An intertemporal capital asset pricing model. *Econometrica, 41*, 867–887.

Monoyios, M. (2004). Performance of utility-based strategies for hedging basis risk. *Quantitative Finance, 4*, 245–255.

Myers, R. J. (1991). Estimating time-varying optimal hedge ratios on futures markets. *Journal of Futures Markets, 11*, 39–53.

Park, S. Y., & Jei, S. Y. (2010). Estimation and hedging effectiveness of time-varying hedge ratio – Flexible GARCH approaches. *Journal of Futures Markets, 30*, 71–99.

Pindyck, R. S. (1986). *Risk aversion and determinants of stock market behavior*. NBER working paper, No. 1921.

Poterba, J., & Summers, L. (1986). The persistence of volatility and stock market fluctuations. *American Economic Review, 76*, 1142–1151.

Sarno, L., & Valente, G. (2000). The cost of carry model and regime shifts in stock index futures markets: An empirical investigation. *Journal of Futures Markets, 20*, 603–624.

Stein, J. L. (1961). The simultaneous determination of spot and futures prices. *American Economic Review, 51*, 1012–1025.

Analytical Bounds for Treasury Bond Futures Prices

71

Ren-Raw Chen and Shih-Kuo Yeh

Contents

R.-R. Chen (✉)
Graduate School of Business Administration, Fordham University, New York, NY, USA
e-mail: rchen@fordham.edu

S.-K. Yeh
Department of Finance, National Chung Hsing University, Taichung 402, Taiwan, Republic of China
e-mail: seiko@nchu.edu.tw

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_71,

Abstract

The pricing of delivery options, particularly timing options, in Treasury bond futures is prohibitively expensive. Recursive use of the lattice model is unavoidable for valuing such options, as Boyle (1989) demonstrates. As a result, the main purpose of this study is to derive upper bounds and lower bounds for Treasury bond futures prices.

This study employs a maximum likelihood estimation technique presented by Chen and Scott (1993) to estimate the parameters for two-factor Cox-Ingersoll-Ross models of the term structure. Following the estimation, the factor values are solved for by matching the short rate with the cheapest-to-deliver bond price. Then, upper bounds and lower bounds for Treasury bond futures prices can be calculated.

This study first shows that the popular preference-free, closed-form cost of carry model is an upper bound for the Treasury bond futures price. Then, the next step is to derive analytical lower bounds for the futures price under one- and two-factor Cox-Ingersoll-Ross models of the term structure. The bound under the two-factor Cox-Ingersoll-Ross model is then tested empirically using weekly futures prices from January 1987 to December 2000.

Keywords

Treasury bond futures • Delivery options • Cox-Ingersoll-Ross models • Bounds • Maximum likelihood estimation • Term structure • Cheapest-to-deliver bond • Timing options • Quality options • Chicago board of trade

71.1 Introduction

Delivery options in Treasury bond futures are generally known as the quality option and three timing options. The *quality* option gives the short the right to deliver any eligible bond (no less than 15 years to maturity or first call) and various timing options give the short the flexibility to make the delivery decision at any time in the delivery month. The *end-of-month timing* option refers to the deliveries occurring during the last 7 business days in the delivery month when the futures market is closed to trading. During the remaining approximately 15 business days of the delivery month, the *wildcard timing* option describes the period from 2:00 p.m. to 8:00 p.m. (Chicago time) every day when the futures market is closed but the bond market is open, while the *accrued interest timing* option refers to the period from 7:20 a.m. to 2:00 p.m. when both the futures and its underlying bond markets are open.

Delivery options in T bond futures are difficult to price. Recursive use of the lattice model is unavoidable for valuing such options, as Boyle (1989) demonstrates, the futures price is effectively a forward price. Furthermore, as demonstrated later, the wildcard timing option is actually a compound forward price – one on top of the other – which cannot be accurately calculated without a multi-recursive system. Consequently, accurate valuation of these delivery options is very costly. This study thus derives fast bounds for the T bond futures price. These bounds can be quickly computed and can provide a crude conservative estimate for the T bond futures price.

Early discussion of the valuation of the quality option appears in Cox et al. (1981), who state that their valuations can be applied to futures with the quality option when the single spot bond price is replaced with the minimum price from the deliverable set. Hemler (1990) uses the exchange option formula of Margrabe's (1978) to price the quality option, but the pricing formula becomes intractable as the number of deliverable bonds increases. Carr (1988) was the first to use factor models to price the quality option, and Carr and Chen (1996) extend the Carr model to include a second factor. Ritchken and Sankarasubramanian (1992) use the Heath et al. (1992) framework to identify the quality option value. Finally, Livingston (1987) analyzes the quality option on the forward contract.

Timing options generally have no closed-form solutions and therefore are studied with lattice methods. Kane and Marcus (1986) lay out a general framework for analyzing the wildcard option. In their analysis, discounting is not considered in the wildcard period. Broadie and Sundaresan (1987) develop a lattice model to value the end-of-month option. They focus strictly on the futures price during the end-of-month period. Boyle (1989) uses a two-period model to demonstrate that the timing option could have a significant impact. His analysis assumes constant interest rates and does not apply directly to T bond futures.

Empiricists generally agree that the quality option has a nontrivial value.[1] However, unlike the evidence for the quality option, the evidence for the timing option is not so clear. This is because most studies do not distinguish between the quality option value and the value from the other timing options, let alone distinguish values among various timing options.[2]

Treasury bond futures contracts are one of the most liquid and widely traded interest rate derivative contracts worldwide and consequently have tight bid-ask spread and high volume. Practitioners thus typically use the market to calibrate the models they use to price other less liquid contracts. Hence, a pricing model that accurately prices both the quality and timing options is necessary to perform this task. However, as demonstrated later, such a model is too expensive since it involves a recursive search for the futures price at the beginning of the delivery month. To have a rough feel for the cost of computation of directly modeling the quality and timing options, this study uses a similar two-factor Cox-Ingersoll-Ross model to the one we use in this paper to compute six futures contract prices. Using a Dell Dimension 2400 with an Intel Celeron processor 2.4 GHz CPU, an average of 9,719.52 seconds (or 2.7 hours) per calculation is required under 102 steps. Clearly such high computational costs are too expensive for real-world applications.

This study derives several results regarding the lower and upper bounds for the futures price. First, this study derives the upper bounds in a model-free format and

[1]See, for example, Carr and Chen (1996), Kilcollin (1982), Benninga and Smirlock (1985), Kane and Marcus (1986), and Hedge (1990).

[2]See, for example, Arak and Goodman (1987), Hedge (1988), and Gay and Manaster (1986).

the lower bound in a semi-model-dependent format. This study proves that the model-free upper bond is the cost of carry model, which is closed form. The lower bond is in the format of an expectation. Since the bounds are almost model-free, violating them implies arbitrage profits. Secondly, this study derives an analytical lower bound for the Treasury bond futures price under the Cox-Ingersoll-Ross model. The study then provides empirical results to show that these bounds are reasonably tight – about 2–3 % above and below the futures price.[3]

The remainder of this paper is organized as follows. The next section studies the quality option. We first study the quality option under continuous marking to market or MTM (i.e., both futures and bond markets are open all the time). Next, the futures price with the quality option is demonstrated effectively to be a forward price when the futures market is closed but the bond market is open. Section 71.3 then provides the theoretical analysis and derives lower and upper bounds for the futures price. Lower bounds are obtained for the futures price under both the quality option and the timing options. We then show that the preference-free cost of carry formula is an upper bound for the futures price. Section 71.4 derives analytical formulas for the lower bound of the futures price (note that the cost of carry formula is model-free) under one- and two-factor Cox-Ingersoll-Ross models. Section 71.5 presents an empirical study where a two-factor equilibrium term structure model is estimated using the Chen and Scott (1993) technique. Finally, Section 71.6 gives a conclusion.

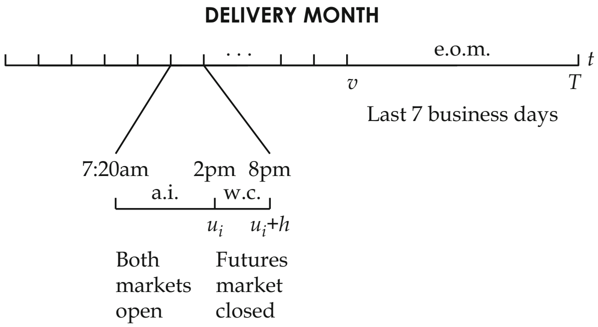
71.2 The Quality Option and the Futures Price

The delivery option that has the highest economic value is the quality option, which gives the short of the futures contract the right to choose the cheapest bond to deliver on the delivery date. Other delivery options that are embedded in T bond futures are known as the three timing options. The short can make a delivery at any time during the delivery month. The short can make a delivery even when the futures market is closed. At the end of the delivery month, for 7 business days, the futures market is closed but the short can still make a delivery. This is understood as the end-of-month timing option. For the remaining approximately 15 business days of the delivery month, the short can deliver either between 7:20 a.m. and 2:00 p.m. (Chicago time) when both the futures market and the underlying bond market are open or after 2:00 p.m. when the futures market is closed.[4]

[3]These bounds are not to be violated, or arbitrage profits should take place. As it will become clear (in Sect. 71.4), in the case of the upper bound that is model-free, a simple trading strategy can be formed to arbitrage against the violation (under perfect markets). In the case of the semi-model-dependent lower bound, arbitrage profits exist only if the assumed model is correct.

[4]T bond market is an over the counter market that has no official closing time, even though market practice adopts 3:00 p.m. Eastern time as a symbolic closing time. The futures market allows the short up to 8:00 p.m. Eastern time to make the delivery announcement, and hence theoretically there is a 5-h window for the wild card.

The former timing option is called the accrued interest timing option and the latter timing option is also known as the daily wild card play. The following picture graphically explains various timing options.



The last 7 business days of the month comprises the end-of-month period. This study uses v to denote the starting time and T to represent the ending time of this period. For the rest of the delivery month, there are two sections of each day, the accrued interest period and the wildcard period. For a regular futures trading day i between 7:20 a.m. and 2 p.m. Chicago time, both the bond and futures markets are open simultaneously. The futures market closes at 2 p.m., but there is no official closing time for the bond market (while conventionally 3 p.m. Eastern time is marked as a symbolic closing time for the bond market). Since the short has till 8 p.m. to make the delivery decision, the wildcard period is defined over 2 p.m. (u_i) to 8 p.m. ($u_i + h$).

The notation and symbols used in the paper are also summarized as follows:
$\Phi(t)$ = "quoted" futures price with all delivery options
$\Phi^*(t)$ = futures price with the quality option and continuous marking to market
$\Phi^{**}(t)$ = futures price with the quality option at the absence of continuous MTM
$\overline{\Phi(t)}$ = upper bound
$\Phi(t)$ = lower bound
$\overline{\Phi_i(t)}$ = futures price of the ith quoted bond price
$\Psi_i(t)$ = forward price of the ith quoted bond price
$a_i(t)$ = accrued interest of the ith bond
$P(t, T)$ = discount bond price at time t of \$1 at time T
$Q_i(t)$ = "quoted" coupon bond price of the ith bond
q_i = conversion factor of the ith bond
$\delta(t, T)$ = random discount factor between t and T

Note that under a specific model for the term structure (e.g., Vasicek or Cox-Ingersoll-Ross), the futures price of a specific bond can be priced in an analytical form (see Sect. 71.4). Before we start our analysis, we need Jamshidian's separation theorem (1987) and his definition of the forward measure.[5]

[5]Also see Hull (2009).

Theorem 1 (Forward Measure) *Let P(t, T) be the price of a pure discount bond delivering* \$1 *at some future date and it follows the dynamics as:*

$$\frac{dP(t,T)}{P(t,T)} = r(t)dt + b(t,T)dW^{Q}(t)$$

where r is the instantaneous risk-free rate, b is maturity dependent bond volatility, and $dW^{Q}(t)$ is the standard Wiener process defined under the risk-neutral space. Then the forward measure is defined as:

$$\frac{dP(t,T)}{P(t,T)} = \left(r(t) - b(t,T)^2\right)dt + b(t,T)dW^{F(T)}(t)$$

where $dW^{F(T)}(t) = dW^{Q}(t) + b(t, T)dt$. Under this forward measure, all expected values taken will be forward prices, that is:

$$\begin{aligned} E_t^Q[\delta(t,T)X(T)] &= E_t^Q[\delta(t,T)]E_t^{F(T)}[X(T)] \\ &= P(t,T)E_t^{F(T)}[X(T)] \end{aligned}$$

where $\delta(t,T) = \exp(-\int_t^T r(u)du)$ and $E_t^{F(T)}[X(T)]$ computes the forward price of X.

A simple proof of this theorem is given in an appendix although the original proof is available in Jamshidian (1987).

71.2.1 The Quality Option with Continuous Marking to Market

In the absence of all timing options, the quality option gives the short the right to deliver the cheapest bond only at maturity, T, and the short receives the following payoff:

$$\max\{q_i\Phi(T) - Q_i(T)\} \quad (71.1)$$

Note that the accrued interests of both bond and futures contracts are equal and canceled. Since the delivery value of Eq. 71.1 has to be identically 0 for all states, we can solve for the futures price at maturity as:

$$\Phi(T) = \min\left\{\frac{Q_i(T)}{q_i}\right\} \quad (71.2)$$

and today's futures price is merely a risk-neutral expectation of this payoff:

$$\begin{aligned}\Phi^*(t) &= E_t^Q\left[\min\left\{\frac{Q_i(T)}{q_i}\right\}\right]\\ &= \frac{E_t^Q[Q_1(T)]}{q_1} - E_t^Q\left[\max\left\{\frac{Q_1(T)}{q_1} - \frac{Q_i(T)}{q_i}\right\}\right]\\ &= \frac{\Phi_1(t)}{q_1} - E_t^Q\left[\max\left\{\frac{Q_1(T)}{q_1} - \frac{Q_i(T)}{q_i}\right\}\right]\end{aligned} \tag{71.3}$$

Note $\Phi_1(t) = E_t^Q[Q_1(T)]$ is the futures price of the first bond with no option and $\Phi^*(t)$ is that of the cheapest bond at maturity. This result has previously been shown by Carr (1988) and others. This equation says that the futures contract with the quality option is equivalent to a futures contract without the quality option (only bond 1 is eligible for delivery) with an exchange option held by the short. With a specific term structure model, Eq. 71.3 becomes an analytical solution.[6]

71.2.2 The Quality Option with No Marking to Market When the Futures Market Is Closed

Equation 71.3 is correct only if marking to market is applied continuously throughout the life of the futures contract. Unfortunately, during the last 7 business days of the delivery month, the futures market is not open and the futures contract is not marked to market. The futures price used for settlement in this period is the last settlement price at the beginning of the 7-day period. Since the futures price is already determined, the actual payoff at the last delivery day, T, is not necessarily 0. The short thus can actually gain or lose. To avoid arbitrage, the futures price at the beginning of the 7-day period should be set so that the expected present value of payoffs at maturity is 0. Under this circumstance, the futures price at the beginning of the 7-day period is a *forward price*, not a futures price. Formally, labeling the futures price as $\Phi^{**}(v)$ to represent the futures price at the beginning of the end-of-month period, v, should be so set that

$$E_v^Q[\delta(v,T)\max\{\Phi^{**}(v)q_i - Q_i(T)\}] = 0 \tag{71.4}$$

where δ is the stochastic discount factor assumed to be strictly less than 1. Using Theorem 1, we can then rewrite Eq. 71.4 as:

$$E_v^{F(T)}[\max\{\Phi^{**}(v)q_i - Q_i(T)\}] = 0 \tag{71.5}$$

which can be expanded as follows:

[6]For example, the closed-form solution under the one-factor Cox-Ingersoll-Ross model can be found in Carr (1988).

$$
\begin{aligned}
0 &= E_v^{F(T)}[\max\{\Phi^{**}(v)q_i - Q_i(T)\}] \\
0 &= E_v^{F(T)}[\Phi^{**}(v)q_1 - Q_1(T) + \max\{\Phi^{**}(v)(q_i - q_1) - (Q_i(T) - Q_1(T)), 0\}] \\
0 &= \Phi^{**}(v)q_1 - \Psi_1(v) + E_v^{F(T)}[\max\{Q_1(T) - Q_i(T) - \Phi^{**}(v)(q_1 - q_i), 0\}]
\end{aligned}
\tag{71.6}
$$

and the futures price at time v can be written as:

$$
\Phi^{**}(v) = \frac{\Psi_1(v)}{q_1} - \frac{1}{q_1} E_v^{F(T)}\left[\max\left\{Q_1(T) - Q_i(T) - K_i^{**}, 0\right\}\right] \tag{71.7}
$$

where $K_i^{**} = (q_1 - q_i)\Phi^{**}(v)$. Note that $\Psi_1(v) = E_v^{F(T)}[Q_1(T)]$ is the forward price of the first bond. The interpretation of this result is similar to that of Eq. 71.3, except that the risk-neutral measure is replaced by the forward measure defined in Theorem 1 and the futures price becomes the forward price. However, unlike Eq. 71.3, the futures price at time v has no easy solution, because it appears on both sides of the equation. This futures price has to be solved recursively using a numerical method. In a lattice framework suggested by Boyle (1989), we first choose an initial value for the futures price at time v, calculate payoffs at various states at maturity T, and then work backwards along the lattice. We adjust the futures price until the discounted payoff computed from the lattice is 0. Once the futures price at time v is set, we can then travel back along the lattice and use the risk-neutral probabilities till the end of the last wildcard period, $u_n + h$. Then the similar procedure for the end-of-month period is repeated for the last wildcard period to arrive at the futures price at the beginning of the wildcard period u_n. Again, the risk-neutral expectation is taken at $u_{n-1} + h$ and a recursive search is to compute the futures price at u_{n-1}. The process is repeated until the delivery month is over. Since the futures price becomes a forward price which cannot be obtained without a recursive search. The search for the "forward price" takes place at every node at all the times (i.e., $u_1, u_2, \ldots, u_n, v$). As a result, to compute the futures price with the quality option is prohibitively expensive.

With the presence of the end-of-month timing option, the futures price computed by Eq. 71.7 is an overestimate because the short has additional flexibility of choosing the best timing. If the short is allowed to deliver at any time in this 7-day period, then we need to compare the expected present value of future payoffs with the current delivery value. Higher current delivery value will trigger early deliveries. This is very similar to the American option pricing methodology where the intrinsic value is compared with the expected present value of future payoffs.

71.3 The Timing Options and Futures Price Bounds

In the previous section, we see that under the end-of-month and a series of wildcard periods, even the quality option alone is very complex to compute, let alone those timing options. In this section, we derive upper and lower bounds for these options in a general framework and analytical formulas are derived in the next section when a specific term structure model is chosen.

71.3.1 The Accrued Interest Timing Option

The accrued interest timing option refers to the flexibility for the short to deliver the cheapest bond any time in the delivery month when both futures and spot markets are open. This is every day from 7:20 a.m. to 2:00 p.m. (Chicago time) from the first day of the delivery month to right before the end-of-month period. Since the futures market is open, the futures contract is marked to market and deliveries can take place any time. As a result, the futures price can never be greater than the cheapest-to-deliver bond price. If the futures price were greater than the cheapest bond price, then deliveries would take place instantly. The short will sell the futures, buy the cheapest bond, make the delivery, and earn an arbitrage profit. Formally, for $t < v$, if the futures price is greater than the delivery value,

$$\begin{aligned} &\Phi(t) > \min\left\{\frac{Q_i(t)}{q_i}\right\} \\ &\text{iff} \\ &\max\{\Phi(t)q_i - Q_i(t)\} > 0 \end{aligned} \tag{71.8}$$

which represents arbitrage profit. Therefore, the futures price in the period where both markets are open must be less than the cheapest-to-deliver bond price to avoid arbitrage. On the other hand, if the futures price is lower, one can long futures and short spot, but the delivery will not occur because the short position of the futures contract will lose money if he makes a delivery. Consequently, the delivery will be postponed and there is no arbitrage profit to be made. If the futures price is always less than the cheapest-to-deliver bond price (adjusted by its conversion factor), the delivery payoff now is negative as opposed to 0 at the end. As a result, the short will never deliver until the last day. Consequently, the accrued interest timing option has no value. We restate this result in the following proposition:

Proposition 1 *The accrued interest timing option without the wildcard and end-of-month options has no value.*[7] □

The existence of the other timing options will lower the current futures price, further reducing the incentive for the short to deliver early. We state this result in the following Corollary:

Corollary 1-1 *The accrued interest timing option with the wildcard and end-of-month options has no value.* □

While the accrued interest timing option is worthless, the timing options at the end-of-month and the wildcard periods are not. When the futures market is closed, there is no marking to market in the futures market and the futures contract becomes a forward contract. Boyle (1989) demonstrates that in a case of forward contracts,

[7]The name "accrued interest" comes in because in the delivery month, the bond price increases due to accrued interests. Here, Q is a traded price that included accrued interests.

timing options will have value. We shall extend Boyle's analysis to stochastic interest rates so that we can evaluate T bond futures timing options.

71.3.2 The End-of-Month Timing Option

Without the end-of-month timing option, we know that the futures price should be set according to Eq. 71.7. With the end-of-month timing option, deliveries can occur any time in the end-of-month period as long as the current delivery payoff is more than the present value of the expected payoff.

When both quality and timing options exist, the short makes a rational delivery decision when the immediate delivery value is higher than the expected discounted value should delivery take place later. This is like the early exercise of an American option. There is no closed-form solution to compute American option prices. Precisely as Boyle (1989) observes, the pricing of quality and timing options do need a lattice model.

To avoid arbitrage, today's futures price needs to be set so that the expected discounted payoff is nil. As a result, if we can identify a function that is always greater than both the delivery payoff and the discounted present value, this function is guaranteed to have a positive present value at time v. This is in spirit similar to the application in Chen and Yeh (2002). The trick is to identify a function that is always greater than the delivery value and the continuation value (ft: Continuation value is the value if it is not optimal to exercise (i.e., delivery). In the binomial model, the continuation value is the value at the node that reflects all possible exercises.)

We guess the function of the following, for $v < t < T$:

$$\begin{aligned} E_t^Q\left[\max\left\{\frac{1}{\delta(t,T)}\Phi(v)q_i - Q_i(T)\right\}\right] &> E_t^Q[\max\{\Phi(v)q_i - \delta(t,T)Q_i(T)\}] \\ &> \max\{\Phi(v)q_i - Q_i(t)\} \end{aligned} \tag{71.9}$$

where δ is the stochastic discount factor which is assumed to be strictly less than 1. This value is greater than the present value of the delivery payoff at any time $t \in [v, T]$. Equation 71.9 states that the upper bound is always greater than the exercise value of the futures contract. The last line is obtained as follows. Note that the martingale result states that: $E_t^Q[\delta(t, T)(Q_i(T) + a_i(T))] = Q_i(t) + a_i(t)$, in other words, discounted market price of a bond should equal its current value, assuming there is no coupon in between t and T.[8] Since the accrued interest is linear but discounting is not (i.e., $P(t, T)a_i(T) > a_i(t)$) it follows that $E_t^Q[\delta(t, T)Q_i(T)] < Q_i(t)$ but the difference is small.

[8]If there is a coupon in between t and T, we simply subtract the coupon value from the expected value.

Equation 71.9 shows that the proposed function is greater than the delivery value at any time. We can also show that the function has a higher value at an earlier time than at a later time. That is:

$$
\begin{aligned}
&E_t^Q\left[\max\left\{\frac{1}{\delta(t,T)}\Phi(v)q_i - Q_i(T)\right\}\right]\\
&> E_t^Q\left[\max\left\{\frac{1}{\delta(t+\Delta t,T)}\Phi(v)q_i - Q_i(T)\right\}\right]\\
&> E_t^Q\left[\delta(t,t+\Delta t)E_{t+\Delta t}^Q\left(\max\left\{\frac{1}{\delta(t+\Delta t,T)}\Phi(v)q_i - Q_i(T)\right\}\right)\right]
\end{aligned}
\tag{71.10}
$$

It is seen that the proposed function is always greater than the delivery value and the discounted continuation value. It must be the case that it is an upper bound for the end-of-month period timing option value. Hence, at time v, the payoff should be positive:

$$
E_v^Q\left[\max\left\{\frac{1}{\delta(v,T)}\Phi(v)q_i - Q_i(T)\right\}\right] > 0 \tag{71.11}
$$

which can be expanded as follows:

$$
\begin{aligned}
&E_v^Q\left[\frac{1}{\delta(v,T)}\Phi(v)q_1 - Q_1(T) + \max\left\{\frac{1}{\delta(v,T)}\Phi(v)(q_i-q_1) - (Q_i(T)-Q_1(T))\right\}\right] > 0\\
&E_v^Q\left[\frac{1}{\delta(v,T)}\right]\Phi(v)q_1 - \Phi_1(v) + E_v^Q\left[\max\left\{Q_1(T) - Q_i(T) - \frac{1}{\delta(v,T)}\Phi(v)(q_1-q_i)\right\}\right] > 0
\end{aligned}
\tag{71.12}
$$

This implies that the futures price should be bounded from below as follows:

$$
\begin{aligned}
\Phi(v) &> \frac{\Phi_1(v)\Delta(v,T)}{q_1} - \frac{\Delta(v,T)}{q_1}E_v^Q\left[\max\left\{Q_1(T) - Q_i(T) - \frac{1}{\delta(v,T)}K_i, 0\right\}\right]\\
&> \frac{\Phi_1(v)\Delta(v,T)}{q_1} - \frac{\Delta(v,T)}{q_1}E_v^Q[\max\{Q_1(T) - Q_i(T) - K_i, 0\}]
\end{aligned}
\tag{71.13}
$$

where

$$
K_i = (q_1 - q_i)\Phi(v) \text{ and}
$$

$$
\Delta(v,T) = \frac{1}{E_v^Q[1/\delta(v,T)]}
$$

Note that the second inequality holds because δ is strictly less than 1. Therefore, the right hand side of the above equation is a lower bound. The lower bound for any time t, $\Phi(t)$, is the risk-neutral expectation of the above lower bound at time v:

$$
\begin{aligned}
\Phi(t) &= E_t^Q\left\{\frac{\Phi_1(v)\Delta(v,T)}{q_1} - \frac{\Delta(v,T)}{q_1}E_v^Q[\max\{Q_1(T) - Q_i(T) - K_i, 0\}]\right\} \\
&= \frac{\Phi_1(t)\Delta(v,T)}{q_1} - \frac{\Delta(v,T)}{q_1}E_t^Q[\max\{Q_1(T) - Q_i(T) - K_i, 0\}]
\end{aligned}
\tag{71.14}
$$

Note that K_i is a function of $\Phi(v)$ which cannot be solved without a recursive search procedure. To arrive at an analytical lower bound, we replace this value with a closed-form futures price $\Phi^*(v)$. We state this result in a following proposition:

Proposition 2 *The futures price under only the end-of-month timing option is bounded from below by the following risk-neutral expectation*:

$$
\frac{\Phi_1(t)\Delta(v,T)}{q_1} - \frac{\Delta(v,T)}{q_1}E_t^Q\left[\max\left\{Q_1(T) - Q_i(T) - K_i^*, 0\right\}\right] \tag{71.15}
$$

where $K_i^* = (q_1 - q_i)\Phi^*(v)$ *and* $\Phi^*(v)$ *is the futures price with only the quality option defined in* Eq. 71.3. □

It is interesting to note that the end-of-month option has a value even if the quality option does not exist. When there is no quality option but the timing option is allowed, the delivery may occur early. The short always compares the delivery payoff $\Phi(v)q - Q(t)$ where $v < t < T$ with the expected present value of the delivery payoff at maturity. We can show that:

$$
E_t^Q[\delta(t,T)(\Phi(v)q - Q(T))] > P(t,T)\Phi(v)q - Q(t) < \Phi(v)q - Q(t) \tag{71.16}
$$

Since the direction of the inequality can go either way, it is likely that early deliveries can take place. This demonstrates that the timing option, even in the absence of the quality option, does have value. The difference between the first two terms in Eq. 71.16 is $P(t,T)a(T) - a(t)$ where a is the accrued interest and the difference of the last two terms is $(1 - P(t,T))\Phi(v)$. As a result, whether or not deliveries will occur early depends upon which effect is larger. This result should not be confused with the result from Boyle (1989) where the timing option is defined differently.

71.3.3 The Wild Card Timing Option

In addition to the end-of-month period where the futures market is closed but the bond market is open, there is a 6-h period every day for about 15 days where the futures market is also closed. This is called the daily wild card timing option. The wild card option is different from the end-of-month option in that the futures market will reopen after each wild card period and the futures contract will be marked to market. If bond prices drop in the wild card period, given that the futures price is fixed, the short can benefit from delivering a cheaper bond. However, the short can equally benefit from the marking to market when the futures market reopens. As a result, the incentive for the short to deliver in the wild card period is minimal. Delivery can take place in a wildcard period only when the payoff from immediate delivery exceeds the expected present value of marking to market on the next day.

We now proceed to derive the bound of the wild card option. For each daily wild card period, $(u, u + h)$, we define the following function as the upper bound of the delivery payoff (for $u < t < u + h$):

$$E_t^Q[\max\{\Phi(u)q_i - \delta(t, u+h)Q_i(u+h)\}] \tag{71.17}$$

This is an upper bound of the payoff because it is greater than (i) the payoff from immediate delivery:

$$\begin{aligned} E_t^Q[\max\{\Phi(u)q_i - \delta(t, u+h)Q_i(u+h)\}] &\geq \max\{\Phi(u)q_i - E_t^Q[\delta(t, u+h)Q_i(u+h)]\} \\ &> \max\{\Phi(u)q_i - Q_i(t)]\} \end{aligned} \tag{71.18}$$

where the second line is obtained by the fact that $E_t^Q[\delta(t,T)Q_i(T)] < Q_i(t)$ proved earlier and (ii) the discounted expected payoff from delivering at the end of the wild card period:

$$\begin{aligned} &E_t^Q[\max\{\Phi(u)q_i - \delta(t, u+h)Q_i(u+h)] \\ &\quad > E_t^Q[\delta(t, u+h)\max\{[\Phi(u)q_i - Q_i(u+h)]\}] \end{aligned} \tag{71.19}$$

Hence, Eq. 71.19 is indeed an upper bound for the wild card option value, which is greater than 0:

$$\begin{aligned} &E_t^Q[\max\{\Phi(u)q_i - \delta(t, u+h)Q_i(u+h)\}] > 0 \\ &E_t^Q\left[\max\left\{\Phi(u) - \delta(t, u+h)\frac{Q_i(u+h)}{q_i}\right\}\right] > 0 \\ &\Phi(u) - E_t^Q\left[\delta(t, u+h)\min\left\{\frac{Q_i(u+h)}{q_i}\right\}\right] > 0 \\ &\Phi(u) > P(t, u+h)E_t^{F(u+h)}\left[\min\left\{\frac{Q_i(u+h)}{q_i}\right\}\right] \end{aligned} \tag{71.20}$$

Note that $\min\left\{\frac{Q_i(u+h)}{q_i}\right\} \geq \Phi(u+h)$ when both markets are open from Sect 71.3.1.[9] Therefore, $\Phi(u) > P(t, u+h)E_t^{F(u+h)}[\Phi(u+h)]$. This is no surprise because the end-of-month option will reduce the futures price prior to time v, which in turn will reduce the futures price at time $u + h$. Hence, it is

[9]Note that in the second line of Eq. 71.17 where q_i is divided through is due to the fact that there exists a bond i such that $\max\{\Phi(u)q_i - \delta(t, u+h)Q_i(u+h)\} > 0$ in all states.

Proposition 3 *Given the futures price next morning (i.e., $\Phi(u_i + h)$) when the futures market reopens at day $i + 1$ (assuming continuous marking to market), the futures price prior to each wildcard period (i.e., $\Phi(u_i)$) is bounded from below by*

$$P(u_i, u_i + h)E_t^{F(u_i+h)}[\Phi(u_i + h)] \tag{71.21}$$

where u_i is the beginning of a wild card period depicted on page 4 and $u_i + h$ is end of the wild card period (which is assumed to be the same as the time when the futures market reopens the next morning). □

71.3.4 Putting It All Together for the Lower Bound

So far, we have derived the lower bound for the futures price of the end-of-month period, $\Phi(v)$, and each of the wildcard period, $\Phi(u)$, where u represents the beginning time of any wildcard period. The futures price of any given time is a recursive calculation of Eq. 71.21. The easiest way to understand the calculation is to picture a univariate lattice model. The lower bound for the futures price at time v is calculated by Eq. 71.15. We shall label it $\underline{\Phi(v)}$ for the lower bound at time v. Then, the regular risk-neutral expectation is taken until the end of the last wildcard period, $u_n + h$ where u_n represents the beginning of the nth (last) wildcard period, is reached. The correct futures price, $\Phi(u_n + h)$, at this moment is unknown since it requires a repeated recursive process described in Sect. 71.3. But we can replace it with the lower bound $\underline{\Phi(u_n + h)} = E^Q_{u_n+h}\left[\underline{\Phi(v)}\right]$. Then, we apply Eq. 71.21 to compute the lower bound at time u_n to get $\underline{\Phi(u_n)} = P(u_n, u_n + h)E^{F(u_n+h)}_{u_n}\left[\underline{\Phi(u_n + h)}\right]$. Repeat this process through all the wildcard periods, u_{n-1}, u_{n-2}, $\ldots$, u_1 to get $\underline{\Phi(u_1)} = P(u_1, u_1 + h)$ $E^{F(u_1+h)}_{u_1}\left[\underline{\Phi(u_1 + h)}\right]$. Then the regular risk-neutral expectation is taken to the current time: $\underline{\Phi(t)} = E^Q_t\left[\underline{\Phi(u_1)}\right]$. Repeated substitutions yield the following general result for the lower bound at the current time $t < u_1$,

$$\begin{aligned}
\underline{\Phi(t)} &= E^Q_t\left[\underline{\Phi(u_1)}\right] \\
&= E^Q_t\left[\delta(u_1, u_1 + h)E^Q_{u_1+h}\left[\underline{\Phi(u_2)}\right]\right] \\
&= E^Q_t\left[\delta(u_1, u_1 + h)E^Q_{u_2}\left[\delta(u_2, u_2 + h)\left[\underline{\Phi(u_2 + h)}\right]\right]\right] \\
&= \cdots \\
&= E^Q_t\left[\prod_{j=1}^{n}\delta(u_j, u_j + h)\underline{\Phi(v)}\right] \\
&= E^Q_t\left[\prod_{j=1}^{n}\delta(u_j, u_j + h)\left(\frac{\Phi_1(v)\Delta(v,T)}{q_1} - \frac{\Delta(v,T)}{q_1}\max_i\{Q_1(T) - Q_i(T) - K_i^*\}\right)\right]
\end{aligned} \tag{71.22}$$

The second line of the above equation is obtained by substituting the lower bound for $\underline{\Phi(u_1)}$ (i.e., $\underline{\Phi(u_1)} = E^Q_{u_1}\left[\delta(u_1, u_1 + h)\left[\underline{\Phi(u_1 + h)}\right]\right.$) and the law of iterative expectations under the risk-neutral measure. We summarize in a proposition:

Proposition 4 *The futures price is bounded from below by the following risk-neutral expectation*:

$$\underline{\Phi(t)} = E^Q_t\left[\prod_{j=1}^{n}\delta\left(u_j, u_j + h\right)\left(\frac{\Phi_1(v)\Delta(v,T)}{q_1} - \frac{\Delta(v,T)}{q_1}\max_i\left\{Q_1(T) - Q_i(T) - K_i^*\right\}\right)\right] \tag{71.23}$$

□

71.3.5 The Cost of Carry Model: The Upper Bound

After deriving the lower bound of the futures price, in the next proposition, we show that the cost of carry model provides an upper bound for the futures price. The well-known cost of carry formula is the following:

$$\Phi_*(t) = \frac{\frac{Q_*(t)+a_*(t)}{P(t,T)} - a_*(T)}{q_*} \tag{71.24}$$

where Q_*, q_*, and a_* are quoted price, conversion factor, and accrued interest of the cheapest bond at time t. Rearranging terms to get:

$$\begin{aligned}\Phi_*(t) &= \frac{\frac{Q_*(t)+a_*(t)}{P(t,T)} - a_*(T)}{q_*} \\ &= E_t^{F(T)}\left[\frac{Q_*(T)}{q_*}\right] \\ &= E_t^{F(T)}\left[\min\left\{\frac{Q_i(T)}{q_i}\right\}\right] \\ &> E_t^{Q}\left[\min\left\{\frac{Q_i(T)}{q_i}\right\}\right]\end{aligned} \tag{71.25}$$

As we can see, the cost of carry model is equal to a forward expectation of the payoff. The futures price without the timing options is a risk-neutral expectation of the payoff (see Eq. 71.3). The last inequality is obtained due to the following:

$$\operatorname{cov}\left[\delta(t,T), \min\left\{\frac{Q_i(T)}{q_i}\right\}\right] > 0 \tag{71.26}$$

This is easy to see because when r increases (decreases), both discount factor, δ, and all quoted bond prices, Q_i's, decrease (increase), and the sign of the covariance is therefore positive. Note that the futures price without timing options is already an upper bound; the cost of carry model used by practitioners is a more conservative upper bound of the futures price. We state the result in the following proposition:

Proposition 5 *The futures price is bounded from above by the cost of carry model.*

$$\overline{\Phi(t)} = \Phi_*(t) \tag{71.27}$$

□

It is generally believed that the futures price with the quality option (Eq. 71.3) is the upper bound of the futures price, since it ignores the timing options. Indeed, if Eq. 71.3 can be evaluated accurately, it is a much tighter lower bound than the cost of carry model shown above. However, note that the cost of carry model is a "model-free" result, while Eq. 71.3 relies upon a specific term structure model. As a result, if the term structure model is not correctly specified, Eq. 71.3 may not serve the role of upper bound well. As we shall see in the empirical section, under a two-factor Cox-Ingersoll-Ross model, Eq. 71.3 does not always provide an upper bound. On the other hand, the violation of the cost of carry upper bound implies arbitrage opportunities.

71.4 Analytical Bounds for Explicit Term Structure Models

The study of the bounds of option prices starts as early as the beginning of modern option pricing theory. There are different approaches to find bounds for option prices. Like Lo (1987) develops the semi-parametric upper bounds of the expected payoffs of options. Zhang (1994) extends the methodology to obtain tighter upper and lower bounds for option prices. In this section, we use the one- and two-factor Cox-Ingersoll-Ross (1985) models to demonstrate how one can calculate the upper bounds of the delivery options and the lower bound of the futures price analytically. Quoted coupon bond price should be equal to:

$$Q(t) = \sum_{j=1}^{m} P\left(t, T_j\right) c_j - a(t), \tag{71.28}$$

Define additional notation $\Phi(t, T_i, T_j) = E_t^Q[P(T_i, T_j)]$ to be the futures price of a pure discount bond delivered at time T_i and $\Psi(t, T_i, T_j) = P(t, T_i)/P(t, T_j)$ to be the forward price of a pure discount bond. These general results are independent of model assumption and of the number of factors.

71.4.1 Single-Factor Model

For the sake of easy exposition and no loss of generality, we shall derive analytical lower bound for the futures price at time v (beginning of end-of-month period).

The lower bound at an arbitrary time t can be derived similarly. Assume $Q_1 > Q_i$ for $i \neq 1$. We follow Carr (1988) that in a single-factor model, the whole distribution of r can be partitioned into n disjoint segments, denoted by $\Omega_i \equiv [r^*_{k(i)-1},\ r^*_{k(i)}]$ where $r_0{}^* = 0$ and $r_n{}^* = \infty$, each of which represents a segment where Q_i maximizes the payoff function: $\max\{Q_1 - Q_i - K_i^*,\}$. The analytic result of the expected value (taken at time v) of Eq. 71.15 is then derived as follows:

$$\begin{aligned} W_v &= E_v^Q\left[\max\left\{Q_1 - Q_i - K_i^*, 0\right\}\right] \\ &= \sum\nolimits_{i=2}^{n} \int\limits_{\Omega_i} \left[\sum\nolimits_{j=1}^{b(1)} c_{1j} P\left(T, T_{1j}\right) - \sum\nolimits_{j=1}^{b(i)} c_{ij} P\left(T, T_{ij}\right) - K_i^*\right] \varphi(r) dr \\ &= \sum\nolimits_{j=1}^{b(1)} c_{1j} \int\limits_{\Sigma_{i=2}^{n}\Omega_i} P\left(T, T_{1j}\right) \varphi(r) dr - \sum\nolimits_{i=2}^{n} \sum\nolimits_{j=1}^{b(i)} c_{ij} \int\limits_{\Omega_i} P\left(T, T_{ij}\right) \varphi(r) dr - K_i^* \end{aligned} \tag{71.29}$$

where $b(i)$ is the last coupon payment time for bond i, $K_i^* = (q_1 - q_i)\Phi^*(v)$, and $\Phi^*(v)$ is the futures price under continuous marking to market (defined by Eq. 71.3). Note that in each region Ω_i, bond i maximizes the payoff $\max\{Q_1 - Q_i - K_i, 0\}$ and $\varphi(r)$ is the risk-neutral density of the interest rate.

Without the consideration of any wild card, the lower bound for the futures price at any arbitrary time t is a risk-neutral expectation of Eq. 71.29:

$$\begin{aligned} \underline{\Phi(t)} &= E_t^Q\left[\prod\nolimits_{j=1}^{n} \delta\left(u_j, u_j + h\right)\left(\frac{\Phi_1(v)\Delta(v,T)}{q_1} - \frac{\Delta(v,T)}{q_1}\max_i\left\{Q_1(T) - Q_i(T) - K_i^*\right\}\right)\right] \\ &= \int_{-\infty}^{\infty} dr(u_1) \int_{-\infty}^{\infty} dr(u_2) \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dr(u_n) \prod\nolimits_{j=1}^{n} \delta\left(u_j, u_j + h\right) \int_{-\infty}^{\infty} dr(v) \frac{\Phi_1(v)\Delta(v,T)}{q_1} \\ &\quad - \frac{\Delta(v,T)}{q_1} \int_{-\infty}^{\infty} dr(T) \max_i\left\{Q_1(T) - Q_i(T) - K_i^*\right\} \varphi(r(u_1), \cdots, r(T)) \\ &= \int_{-\infty}^{\infty} dr(u_1) \int_{-\infty}^{\infty} dr(u_2) \int_{-\infty}^{\infty} \cdots \int_{-\infty}^{\infty} dr(u_n) \prod\nolimits_{j=1}^{n} \delta\left(u_j, u_j + h\right) \int_{-\infty}^{\infty} dr(v) \frac{\Phi_1(v)\Delta(v,T)}{q_1} \\ &\quad - \frac{\Delta(v,T)}{q_1} \sum\nolimits_{i=2}^{n} \int_{\Omega_i} \left[\sum\nolimits_{j=1}^{b(1)} c_1 P\left(T, T_{1j}\right) - \sum\nolimits_{j=1}^{b(i)} c_i P\left(T, T_{ij}\right) - K_i^*\right] \varphi(r(u_1), \cdots, r(T)) \\ &= \int_{-\infty}^{\infty} dr(u_1) \cdots \int_{-\infty}^{\infty} dr(u_n) \prod\nolimits_{j=1}^{n} \delta\left(u_j, u_j + h\right) \int_{-\infty}^{\infty} dr(v) \frac{\Phi_1(v)\Delta(v,T)}{q_1} - \frac{\Delta(v,T)}{q_1} W_v \varphi(r(u_1), \cdots, r(v)) \end{aligned} \tag{71.30}$$

where W_v is defined in Eq. 71.29.

In the case of CIR, the interest rate process follows the square root process:

$$dr = (\alpha\mu - (\alpha + \varsigma)r)dt + \sigma\sqrt{r}dW^Q \tag{71.31}$$

where a is the reverting speed, μ is the reverting level, σ is the volatility parameter, and ς is the market price of risk which is constant under log utility. The futures price with only the quality option is in Carr (1988) as:

$$\Phi^*(t) = \sum_{i=1}^{n}\sum_{j=1}^{b(i)}\frac{c_{ij}}{q_i}\Phi\left(t,T,T_{ij}\right)\sum\nolimits_{k=1}^{m(i)} I_{ik}\left[\chi_j^2\left(r_{ik}^*\right) - \chi_j^2\left(r_{ik-1}^*\right)\right] \tag{71.32}$$

where

$\Phi(t, u, v) = C(t, u, v)e^{-\,r(t)D(t,u,v)}$, futures price under the CIR model without any options

$$C(t,u,v) = A(u,v)\left(\frac{\eta(t,u)}{\eta(t,u)+B(u,v)}\right)^{\frac{2\alpha\mu}{\sigma^2}}$$

$$D(t,u,v) = \frac{B(u,v)\eta(t,u)e^{-(\alpha+\varsigma)(u-t)}}{\eta(t,u)+B(u,v)}$$

$$A(u,v) = \left(\frac{2\gamma e^{(\alpha+\varsigma+\gamma)(v-u)/2}}{(\alpha+\varsigma+\gamma)(e^{\gamma(v-u)}-1)+2\gamma}\right)^{\frac{2\alpha\mu}{\sigma^2}}$$

$$B(u,v) = \frac{2\left(e^{\gamma(v-u)}-1\right)}{(\alpha+\varsigma+\gamma)(e^{\gamma(v-u)}-1)+2\gamma}$$

$$\eta(t,u) = \frac{2(\alpha+\varsigma)}{\sigma^2(1-e^{-(\alpha+\varsigma)(u-t)})}$$

$$\chi_j^2(r^*) = \chi^2\left[2\eta(t,u)r^*;\frac{4\alpha\mu}{\sigma^2},2\eta(t,u)re^{-(\alpha+\varsigma)\left(T_{ij}-t\right)}\right]$$

$m(i)$ represents the domain where bond i can be the cheapest to deliver, similar to Ω_i in Eq. 71.29. T_{ij} represents the time of the jth coupon of bond i. c_{ij} represents the jth cash flow (coupon or coupon and principal) of bond i. For example, when $j < b(i)$, c_{ij} is coupon, and when $j = b(i)$, c_{ij} is coupon plus principal. In addition, note that I_{ik} is the indicator function equal to 1 for the ith bond and between the critical values of r^*_{ik-1} and r^*_{ik} and $\chi^2(x, y, z)$ is a noncentral chi-square probability function with limit x, degrees of freedom y, and degrees of noncentrality z.

Under the CIR model, Eq. 71.29 becomes

$$\begin{aligned} W_v^{\mathrm{CIR}} &= E_v^Q\left[\max\left\{Q_1 - Q_i - K_i^*, 0\right\}\right] \\ &= \left\{\sum\nolimits_{i=1}^{n-1}\left[\sum\nolimits_{j=1}^{b(i)} c_1\Phi\left(v,T,T_{1j}\right)\chi^2(\bar{r}_1) - \sum\nolimits_{j=1}^{n} c_i\Phi\left(v,T,T_{ij}\right)\chi^2(\bar{r}_i) - K_i^*\chi^2(\bar{r}_1)\right]\right\} \end{aligned} \tag{71.33}$$

$\bar{r}_1$ and $\bar{r}_i$ denote crossover rates that are over a plausible range of interest rates; there exists only one such rate that determine regions where one bond is cheaper and regions where the other is cheaper. Then the lower bound under the CIR model of the term structure is still Eq. 71.30 but with W_v^{CIR} replacing W_v.

71.4.2 Two-Factor Model

We use the two-factor model of the following kind:[10]

$$r = y_1 + y_2 \tag{71.34}$$

where each factor follows a square root process as in Eq. 71.31:

$$dy_i = (\alpha_i\mu_i - (\alpha_i + \varsigma_i)y_i)dt + \sigma_i\sqrt{y_i}dW_i^Q \tag{71.35}$$

where $i = 1, 2$ and $dW_1^Q dW_2^Q = 0$. Under this framework, the two-factor model works the same way as the one-factor models. The difference is that the univariate integrals in the one-factor models are replaced with two-dimensional integrals:

$$\begin{aligned}\underline{\Phi(t)} =& \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dr(u_1)\cdots\int\int_{-\infty}^{\infty} dr(u_n)\prod\nolimits_{j=1}^{n}\delta\left(u_j, u_j + h\right)\\ & \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} dr(v)\frac{\Phi_1(v)\Delta(v,T)}{q_1} - \frac{\Delta(v,T)}{q_1}W_v^{\mathrm{CIR}(2)}\hat{\varphi}(r(u_1),\cdots,r(v))\end{aligned} \tag{71.36}$$

where

$$\begin{aligned}W_v^{\mathrm{CIR}(2)} =& \left\{\sum\nolimits_{i=1}^{n-1}\left[\sum\nolimits_{j=1}^{b(1)} c_1\Phi\left(v,T,T_{1j}\right)\iint\limits_{\Omega_i}\hat{\varphi}(y_1,y_2)dy_1dy_2\right.\right.\\ & \left.\left. -\sum\nolimits_{j=1}^{b(i)} c_i\Phi\left(v,T,T_{ij}\right)\iint\limits_{\Omega_i}\widehat{\varphi}(y_1,y_2)dy_1dy_2 - K_i^*\iint\limits_{\Sigma_{i=1}^{n-1}\Omega_i}\hat{\varphi}(y_1,y_2)dy_1dy_2\right]\right\}\end{aligned}$$

The bivariate integrals may become quadruple integrals as we move backwards in time. The lattice approach proposed by Longstaff and Schwartz (1992) can be efficiently implemented to calculate the result. Since the lower bound requires only risk-neutral expectations, it can be computed without recursive loops and be extremely fast.

[10] This two-factor model is adopted by a number of authors. See Chen and Scott (1993), Turnbull and Milne (1991), Langetieg (1980), and Hull and White (1990).

Table 71.1 Parameter estimates of the two-factor Cox-Ingersoll-Ross model

	Chen-Scott estimation					New estimation			
	Factor 1	Std. err.	Factor 2	Std. err.		Factor 1	Std. err.	Factor 2	Std. err.
α	1.834100	0.222800	0.005212	0.115600	α	0.879967	0.001014	0.004423	0.000014
μ	0.051480	0.005321	0.030830	0.683300	μ	0.043822	0.000009	0.029555	0.000097
σ	0.154300	0.005529	0.066890	0.002110	σ	0.097855	0.001429	0.095974	0.000018
ς	−0.125300	0.180600	−0.066500	0.115400	ς	−0.146140	0.000151	−0.178846	0.000361
	Likelihood function = 7750.82					Likelihood function = 11722.81			
	# of obs. 470					# of obs. 416			

Chen-Scott estimates are taken from Exhibit 2, Panel B on page 21 of Chen and Scott (1993) who take Thursday weekly prices of 13-week, 26-week, 5-year, and longest maturity Treasuries. The period of study is from January 1980 to December 1988. The new estimates use Friday weekly T-Bill rates of 3 months and 6 months and CMT rates of 5 years and 30 years. The period of study is from January 1991 to December 1998. The new estimates are estimated with RATS where the number of usable observations in the estimation is 387

71.5 Empirical Study

This section empirically examines the magnitude of each bound using a two-factor CIR model. Evidence is presented for two non-overlapping periods: 1987–1991 and 1992–2000. Both in-sample and out-of-sample tests are performed during each period. Results are similar for both periods, implying the model is robust. Furthermore, for both periods, out-of-sample performance is pleasantly satisfactory.

71.5.1 Term Structure Model Estimation

In estimating the two-factor CIR term structure model, this study uses weekly (Friday) four Treasury interest rate series: the 3-month and 6-month Treasury bills and the 5-year and 30-year Constant Maturity Treasury (CMT) interest rates to estimate the parameters for the two-factor CIR model. Weekly data is obtained from January 4, 1991 to December 29, 1998, comprising 416 observations in total. Data is from the Aremos USFIN databank.

Parameter estimation is important because without good estimates, the two-factor CIR model cannot work properly. Several estimation techniques regarding term structure models have recently been developed. Duan and Simonato (1999) use the state space model with the Kalman filter to estimate exponential-affine term structure models. The estimation procedure is identical to that described in Chen and Scott (1993). Its methodology employs maximum likelihood estimation for a time-series bond price data. Hamilton (1994) provides a detailed introduction for maximum likelihood estimation applied in time-series data. Furthermore, Yeh and Lin (2003) use curve fitting techniques and also used cross-sectional bond price data to estimate the Vasicek (1977) and the CIR models. In addition to our estimates, as a robustness comparison, we also use the results from Chen and Scott (1993) who used a weekly dataset from 1980 to 1988, and Table 71.1 lists the estimates obtained using both estimations. The estimates change little from one period to another, while the new estimates show slightly lower reverting level and slower mean reversion. The first factor remains strong mean reversion while the second remains close to a random walk.

The term structure estimation must also estimate the factor values. Chen and Scott (1993) compute factor values by fitting the long and short rates of the yield curve. For the purpose of this study (correctly pricing the cheapest-to-deliver bond correctly),[11] the factor values are solved for by matching the short rate with the cheapest-to-deliver bond price. In reality, the delivery options are priced based on the cheapest-to-delivery bond and a series of exchange options to the next cheapest, the third cheapest, and so on. Calibrating the term structure model of Chen-Scott to the cheapest-to-delivery bond yields the most accurate valuation of the delivery options using the two-factor CIR model. The two-factor CIR model is generally understood not to closely fit the yield curve.[12] To mitigate the concerns of Jagannathan et al. (2003), this study examines how our term structure best fits the set of deliverable bonds. Unlike Ho et al. (1992), this study is not particularly concerned with the whole yield curve fit because most of the risk of the delivery options resides in the set of deliverable bonds. Furthermore, as a practical concern, we present the fitting performance of the three most relevant bonds – the cheapest, second cheapest, and third cheapest. The probability of other bonds becoming the cheapest is small and the impact of other deliverable bonds is believed to be negligible.

Theoretically, the cheapest bond at any time should be fitted perfectly by tweaking the second factor, since there is one equation and one unknown. However, no solution exists for the second factor at the following dates when trying to fit the cheapest bond: 980903, 980910, 980917, 980924, 981001, 981015, 981029, 981203, 981210, and 981217. Figure 71.1 plots the yield curves for a subperiod (January 2, 1998, to December 28, 2000) from the CMT dataset. The above dates, where the second factor fails to coincide (CTD bond fails to fit) with the period, occur when the yield curve is steeply sloped and the short rates are small. This problem has already been described in Chen and Scott (1993). Chen and Scott recommend a three-factor model to improve the fit. However, because this problem only exists for 10 out of 722 cases (252 observations in the first subperiod, 1987–1991, and 470 observations in the second subperiod, 1992–2000)[13] and because of the complexity of estimating a three-factor model, the two-factor model is retained.[14] Alternatively, the first factor can be left flexible until the CTD bond can be fitted. However, to maintain consistency, the CTD bond is allowed to

[11]T bond futures prices are affected by all bonds underlying the yield curve, and yet doubtlessly the cheapest-to-deliver bond has the most influence.

[12]See, for example, Chen and Scott (1993) and Jagannathan et al. (2003).

[13]All 10 cases are in the second subperiod: 1992–2000.

[14]Chen and Scott (1993) argue that the three-factor model does not necessarily dominate the two-factor model; in that, the three-factor model, although fits better the term structure, generates extra volatility. See Chen and Scott for details.

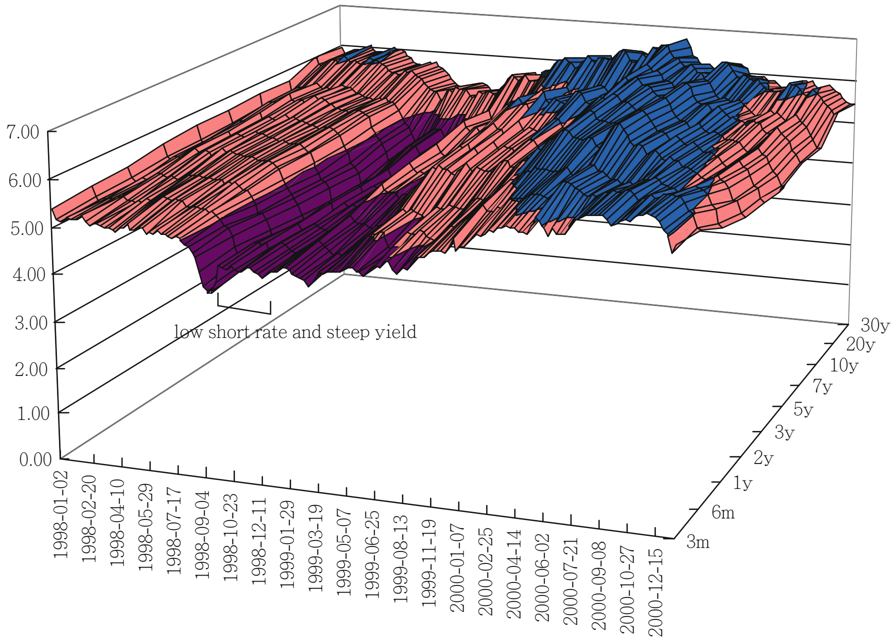


Fig. 71.1 Yield curves for the selected period

imperfectly fit those 10 dates.[15] The ill-fitted CTD bonds for those 10 dates thus hurt the tightness of the bounds. The following summary illustrates the cheapest bond that fails to be fitted and the difference between the market and the model prices.

Date	Coupon	Maturity	Market price	Model price	% diff
980903	11.250	150215	164.6250	159.2477	3.38
980910	11.250	150215	167.9063	158.6218	5.85
980917	11.250	150215	167.2500	163.3068	2.41
980924	11.250	150215	168.3438	163.2337	3.13
981001	11.250	150215	171.7188	163.7861	4.84
981015	11.250	150215	169.3438	165.2365	2.49
981029	11.250	150215	168.2813	164.6084	2.23
981203	11.250	150215	168.7813	162.1672	4.08
981210	11.250	150215	169.3438	161.4943	4.86
981217	11.250	150215	167.9688	161.1123	4.26

Notably, other than these ten dates, the CTD bond fits perfectly. To mitigate the criticism of Jagannathan et al. (2003), the fitting performance of the second cheapest and the third cheapest must also be examined. Figure 71.2 presents the

[15]The result of the alternative fitting is available upon request.

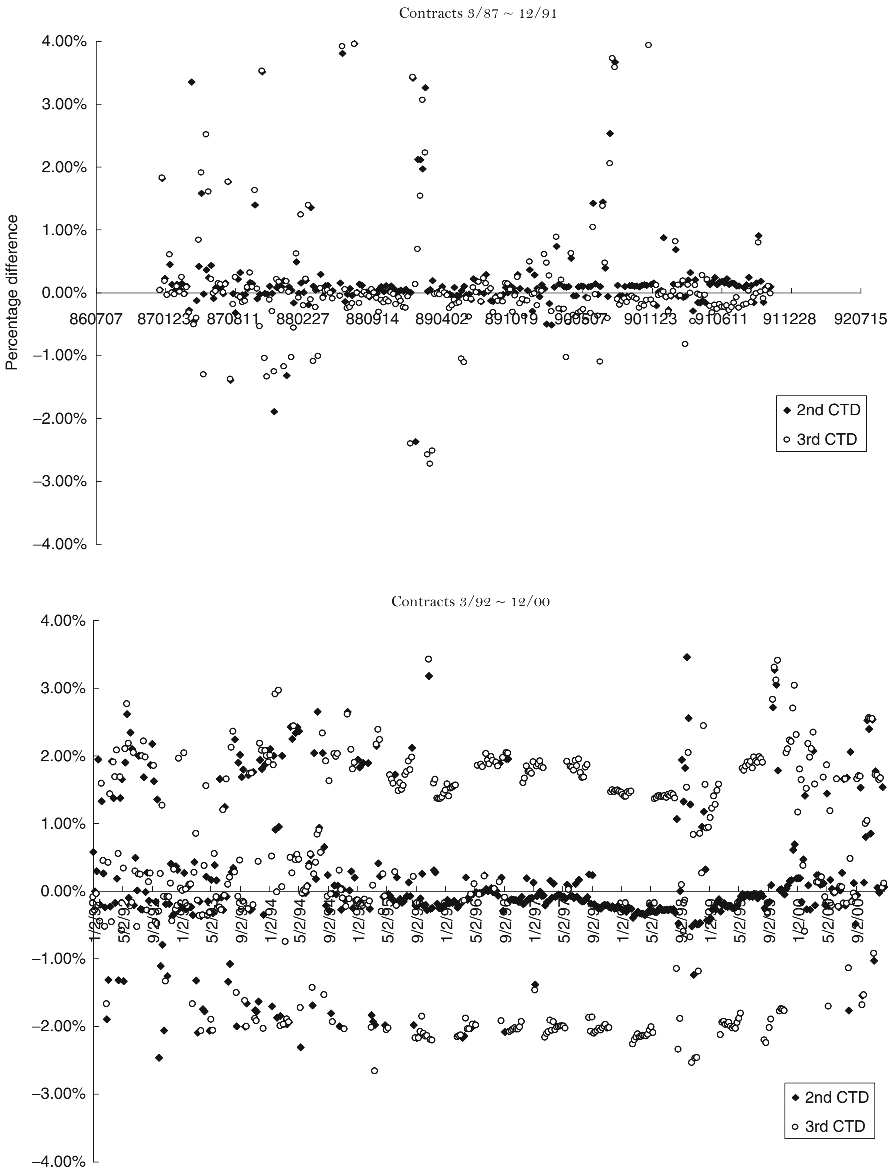


Fig. 71.2 Fitting performance of the second and third cheapest-to-deliver bonds
Note: The pricing error is measured as percentage error of the market price: model price ÷ market price – 1. The average percentage errors are 30 basis points and 14 basis points for the 2nd CTD and 3rd CTD, respectively. The root mean square errors are 1.07 % and 1.20 %, respectively
Note: The pricing error is measured as percentage error of the market price: model price ÷ market price – 1. The average percentage errors are 10 basis points and 26 basis points for the 2nd CTD and 3rd CTD, respectively. The root mean square errors are 1.04 % and 1.61 %, respectively

fitting performance of the two-factor model (with the 3-month short rate and the CTD bond perfectly fitted). The percentage fitting error (theoretical price ÷ market price – 1) is plotted. The second CTD bonds are fitted very well. The average percentage error (APE) is 30 basis points in the first period (1987–1991) and 10 basis points in the second subperiod (1992–2000). The root mean square errors (RMSE) for both periods are 1.07 % and 1.04 %, respectively. The numbers may seem to suggest that the second sample period provides better fit, but eyeballing the graphs reveals good fit during most of the first sample period; the second CTD is well fitted and only half of the time in the second period is well fitted.

The fitting performance of the third CTD bonds displays a very different profile. The third CTD bonds presents equal fit to the second CTD bonds during the first sample period (1987–1991), but the former have worse fit than the later during the second subperiod (1992–2000). As opposed to 30 basis points APE and 1.07 % RMSE for the second CTD bond, the APE and the RMSE for the third CTD bond are 14 basis points and 1.2 % in 1987–1991. However, during 1992–2000, the APE and RMSE increase from 10 basis points and 1.04 %, respectively, for the second CTD bond to 26 basis points and 1.61 %, respectively, for the third CTD bond. The worse fit of the third CTD bond and the 10 cases of unsuccessful fit of the CTD during the second subperiod might explain the slightly worse bound performance (show later) during the second period.

71.5.2 Futures Data

Daily futures prices are obtained from the Chicago Board of Trade (CBOT) between January 1987 and December 2000. Table 71.2 lists the summarized statistics. Notably, the decline in the futures price for March 2000 contract results from the change of the discount rate in the conversion factor (from 8 % to 6 %). However, this study collects the futures prices weekly (Thursday) for two different (non-overlapping) periods. One period is from January 8, 1987 through October 31, 1991 (252 observations) and the other runs from November 7, 1991 through November 2, 2000 (470 observations). The first period, which covers the quarterly contracts during March 1987 through December 1991, uses the Chen-Scott estimates, and the second period, which covers contracts during March 1992 through December 2000, uses the new estimates. Weekly futures prices with maturity ranging from 6 weeks to 4½ months are selected from the CBOT daily price dataset.

The cost of carry model requires the knowledge of all deliverable bonds at the trade date. All deliverable bonds are collected from the Wall Street Journal for all the trade dates and the average of the bid and ask is taken as the bond price. This study also uses the 3-month T bill rates for the cost of carry model. There are about 26 deliverable bonds for any given trade date. Conversion factors are computed using the CBOT formula.[16]

[16]Hull (2009) has an excellent demonstration of such a computation.

Table 71.2 Summary statistics of daily futures prices

Contract month	N	Mean	Std. dev	Min	Max
All maturities	3537	103.69	11.4274	77.78	134.66
8703	21	100.62	0.6833	99.47	101.59
8706	63	97.23	3.2328	88.56	101.38
8709	64	90.72	1.5606	86.84	93.19
8712	65	84.59	3.3369	77.78	90.09
8803	63	88.23	2.147	83.72	93.91
8806	64	91.19	1.9662	97.34	94.16
8809	64	86.68	1.2191	84.44	89.56
8812	65	87.53	2.1673	83.94	91.41
8903	63	89.09	1.1461	86.97	91.44
8906	64	88.73	1.0879	86.5	91.28
8909	64	95.23	3.1389	88.34	100.38
8912	65	97.42	1.1772	95.25	99.84
9003	63	98.29	1.9	93.22	100.28
9006	64	92.26	1.6765	88.59	94.72
9009	64	93.15	1.1783	89.78	95.19
9012	65	89.6	1.2881	87.16	93.09
9103	63	94.91	1.6509	91.09	97.56
9106	64	95.95	1.1031	93.44	97.94
9109	63	93.83	0.925	92.28	95.94
9112	64	98.17	1.4522	95.25	100.41
9203	62	101.25	2.106	97.78	105.25
9206	62	98.94	0.8243	97.28	100.31
9209	64	100.79	2.0015	97.31	105.16
9212	64	104.46	1.1772	102.31	106.91
9303	61	103.9	1.7221	100.28	107.22
9306	64	109.79	1.9499	105.69	112.66
9309	64	112.3	2.3068	108.44	115.97
9312	64	118	2.83	102.63	121.94
9403	62	110.33	0.9456	113.34	117.44
9406	64	108.69	3.4739	103.25	115.34
9409	64	103.02	1.2617	100.31	105.44
9412	64	100.08	1.9172	97.06	103.81
9503	60	98.64	1.5588	95.44	101.47
9506	64	103.57	1.4609	100.5	106.31
9509	64	112.34	2.2275	106.97	115.75
9512	61	113.51	2.6347	108.69	117.44
9603	63	119.33	1.4683	116.75	121.56
9606	65	112.94	3.5877	106.75	120.22
9609	62	108.1	1.0822	105.88	111.84
9612	62	109.72	1.7013	106.41	113
9703	59	113.06	1.9931	109.78	120.06
9706	61	109.45	1.9678	106.63	113.44

(*continued*)

Table 71.2 (continued)

Contract month	N	Mean	Std. dev	Min	Max
9709	62	111.83	0.3393	108.31	116.75
9712	62	114.62	1.7645	112.06	118.47
9803	59	120.13	1.8508	117.03	123.72
9806	61	120.56	0.8548	118.66	122.44
9809	63	122.1	1.3907	118.88	124.16
9812	62	127.7	2.8354	122.97	134.66
9903	58	127.75	1.4533	124.72	130.63
9906	64	121.89	1.4661	119.47	126.19
9909	63	116.3	1.4175	113.63	119.38
9912	62	113.38	1.2839	110.84	116.16
0003	60	92.38	1.9398	89.22	95.66
0006	63	95.8	1.7785	92.47	99.34
0009	63	96.46	1.8667	92.66	99.38
0012	62	99.54	0.8819	97.63	101.22

Daily futures prices are taken with maturity between 6 weeks and 4½ months for each contract. Such a selection enjoys high liquidity and rare overlapping between contracts

71.6 Results

This study assumes no gap between the close of the bond market for any given day and the open of the futures market in the next morning. As a result, correctly dating all the timing periods in the lattice requires counting the number of trading days. As has been pointed out previously, each month contains about 22 trading days. The last 7 days attribute to the end-of-month period. For each of the remaining 15 days, there are about 6 h for the day period where both bond and futures markets are open and another about 6 h for the night period where only the bond market is open. To accurately calculate various timing option values, the time to maturity in this study is not measured using calendar days but business days.[17] Accurate day counting is necessary because we need to calculate expectations at various times.

First, all deliverable bonds are ranked by their conversion factors. The bond with the largest conversion factor is then chosen as the primary bond for delivery and its futures price is calculated using the two-factor version of the Cox-Ingersoll-Ross model (1981). Various timing options provide the short additional flexibility of choosing the best timing.

Calculating the upper bound value for the end-of-month option for any given time prior to v requires calculating Eq. 71.36 and then using Eq. 71.14. As noted earlier, the wildcard value can be ignored if the lower bound of the futures price at the beginning time of the end-of-month period, v, is already sufficiently low. That is, if the lower bound for the end-of-month option is employed to substitute for $\Phi(v)$, the

[17] That is, we do the business day count between trade day and the last day of the delivery month and assume 252 trading days for a given year.

loss of the wildcard value is translated into the end-of-month option. Restated, the wildcard value can be efficiently incorporated into the lower bound for the end-of-month option. If the wildcard value is completely eliminated by this substitution, then the lower bound for the end-of-month option becomes a lower bound for both the end-of-month and wildcard options. As we shall see, this is indeed the case for the periods examined. Finally, the cost of carry model of Eq. 71.24 is computed and compared with the actual futures price.

The empirical examination of the upper and lower bounds is presented in two periods where the term structure model is estimated separately. The first period contains the futures contracts from March 1987 to December 1991. The empirical results during this period use the parameter estimates of Chen and Scott (1993), which use data from January 1980 to December 1988 to estimate the term structure. Hence, contracts from March 1987 to December 1988 are considered in-sample and contracts from March 1989 to December 1991 are considered out of sample. The second period contains the futures contracts from March 1992 to December 2000. The parameters are reestimated using the Treasury data from January 1991 through December 1998 to estimate the term structure, and in-sample and out-of-sample tests are also performed. The parameters are reestimated because significantly lower interest rates are observed during the later period.

The first part of Table 71.3 lists results in averages for the 20 contracts (8703–9112) studied in this investigation. The first three columns of Table 71.3 present actual futures prices, lower bound futures prices using Eq. 71.36 which considers only the end-of-month option and upper bound futures prices using which is the cost of carry model. The average for the whole period is listed at the bottom of the table. The estimates obtained using the cost of carry model are on average 2 % higher than the actual futures price, while the lower bound is 2 % lower than the actual futures price. Weekly prices of these three series are plotted in Fig. 71.3.

The end-of-month option bound values are listed in column 4. This value includes both the quality option and the timing option values. Separating these two values is difficult because no consistent method exists for measuring the quality option. Figure 71.3 shows that the lower bound for the futures price provided by this upper bound is sufficiently conservative to include all daily wildcard values. Furthermore, the bound is as tight as the cost of carry model, approximately 2 % on average lower than the actual futures price.

This study also estimates the two-factor Cox-Ingersoll-Ross term structure model for a more recent dataset (weekly, from January 4, 1991 through December 29, 1998). The in-sample test is for the contracts from March 1991 to December 1998 and the out-of-sample test is for the contracts from March 1999 to December 2000. Somewhat different and yet very interesting results can be observed. Similar to the first half of Table 71.3, the second half of the table presents the results from the second period in a parallel fashion. First, the model-free upper bound, the cost of carry model, performs as well as in the first period, remaining roughly 2 % above the actual price, a very robust result.

The most surprising result involves the lower bound. Using the same term structure model, the lower bound, on average, remains within about 2 % below

Table 71.3 Empirical performance of upper and lower bounds

Contract month	# of obs.	(1)	(2)	(3)	(4)
March 1987–December 1991					
8703	4	100.703	99.486	100.585	1.505
8706	13	97.875	96.136	100.132	2.092
8709	13	90.719	89.204	93.853	2.031
8712	13	84.546	83.357	86.323	1.887
8803	13	88.269	87.545	89.407	2.208
8806	13	91.267	89.973	95.513	1.971
8809	13	86.628	84.728	87.053	1.956
8812	13	87.269	85.947	87.530	2.367
8903	13	89.123	87.205	89.912	2.496
8906	13	88.712	86.533	92.655	2.491
8909	14	94.980	92.763	96.251	2.594
8912	13	97.471	95.300	100.400	2.264
9003	12	98.430	96.106	100.170	2.491
9006	14	92.252	90.549	95.804	2.421
9009	13	93.238	91.168	94.006	2.302
9012	13	89.572	88.004	92.559	2.146
9103	13	95.195	93.275	95.982	2.003
9106	13	96.003	94.693	97.561	1.855
9109	13	93.902	92.723	94.322	1.856
9112	13	98.152	96.792	100.638	1.603
All maturities	252	92.414	90.758	94.306	2.151
March 1992–December 2000					
9203	13	101.216	99.800	102.820	1.392
9206	13	98.875	95.497	101.650	3.315
9209	13	100.930	98.520	101.607	2.366
9212	13	104.577	100.955	106.022	3.563
9303	13	103.926	101.612	107.472	2.267
9306	13	110.099	107.530	110.817	2.531
9309	13	112.274	109.783	113.701	2.463
9312	13	118.125	115.264	120.051	2.843
9403	13	115.250	114.138	117.530	1.096
9406	13	108.777	107.669	110.239	1.090
9409	13	103.132	100.401	104.535	2.712
9412	13	100.277	97.763	100.997	2.473
9503	13	98.438	94.237	100.829	4.153
9506	13	103.394	100.743	104.408	2.616
9509	14	112.212	108.228	113.331	3.958
9512	12	113.485	112.215	114.642	1.240
9603	14	119.299	115.368	123.136	3.905
9606	13	112.681	111.661	113.476	0.993
9609	13	108.375	104.510	108.601	3.840
9612	13	109.630	108.310	111.131	1.294

(*continued*)

Table 71.3 (continued)

Contract month	# of obs.	(1)	(2)	(3)	(4)
9703	13	112.834	109.111	116.659	3.700
9706	13	109.301	108.105	110.123	1.172
9709	13	111.875	107.909	114.585	3.944
9712	13	114.690	113.795	115.964	0.871
9803	13	120.329	116.394	124.512	3.911
9806	13	120.625	120.158	121.157	0.443
9809	13	122.120	118.071	123.111	4.026
9812	13	127.772	124.326	129.021	3.054
9903	13	127.916	121.848	131.740	5.584
9906	13	121.709	121.580	122.289	0.104
9909	13	116.298	112.651	116.150	3.629
9912	13	113.397	111.580	114.369	1.797
0003	13	92.378	89.690	93.742	3.746
0006	13	95.856	93.482	97.851	2.848
0009	14	96.574	95.796	97.963	2.662
0012	13	99.606	97.489	101.067	3.729
All maturities	470	109.940	107.378	111.584	2.651

(1) Is actual futures price
(2) Is lower bound (Eq. 71.36)
(3) Is cost of carry price, also upper bound (Eq. 71.24)
(4) Is average of (1)−(2), a measure of bound tightness
The theoretical values are computed using the Chen-Scott estimates (left panel of Table 71.1)
The theoretical values are computed using the new estimates (right panel of Table 71.1)

the actual futures price. Notably, the term structure estimation fitting performance result during the first period (March 1987 to December 1991) is different from that during the second period (March 1992 to December 2000). The term structure model performs poorly during the second period but it performs well during the first period. However, the lower bound still performs well during both periods. This is due to the fact that the lower bound, on the other hand, is only "semi-model-dependent" in that its theory is model-free and only the implementation requires a term structure model. Hence, in theory, the lower bound is a "looser" bound and thus unsurprisingly can perform well no matter what the term structure model fits. Furthermore, from Table 71.3 (second part), throughout all contracts, the lower bound constantly falls below the actual futures price. The weekly lower bound performance can be seen in the second part of Fig. 71.3. This observation raises an interesting issue, namely, that using an ill-fitted term structure model to estimate contract value, the performance of the estimate relies extremely on the performance of the underlying model. However, when estimating a range of values for the contract, the accuracy of the underlying model becomes less sensitive. In reality, no trader is seeking "the price," since model assumptions are always inconsistent with reality. However, robust models (models that are robust to parameter changes) are useful in that they provide useful implications traders can use to gain insights. What we have learned from this empirical test precisely demonstrates this point.

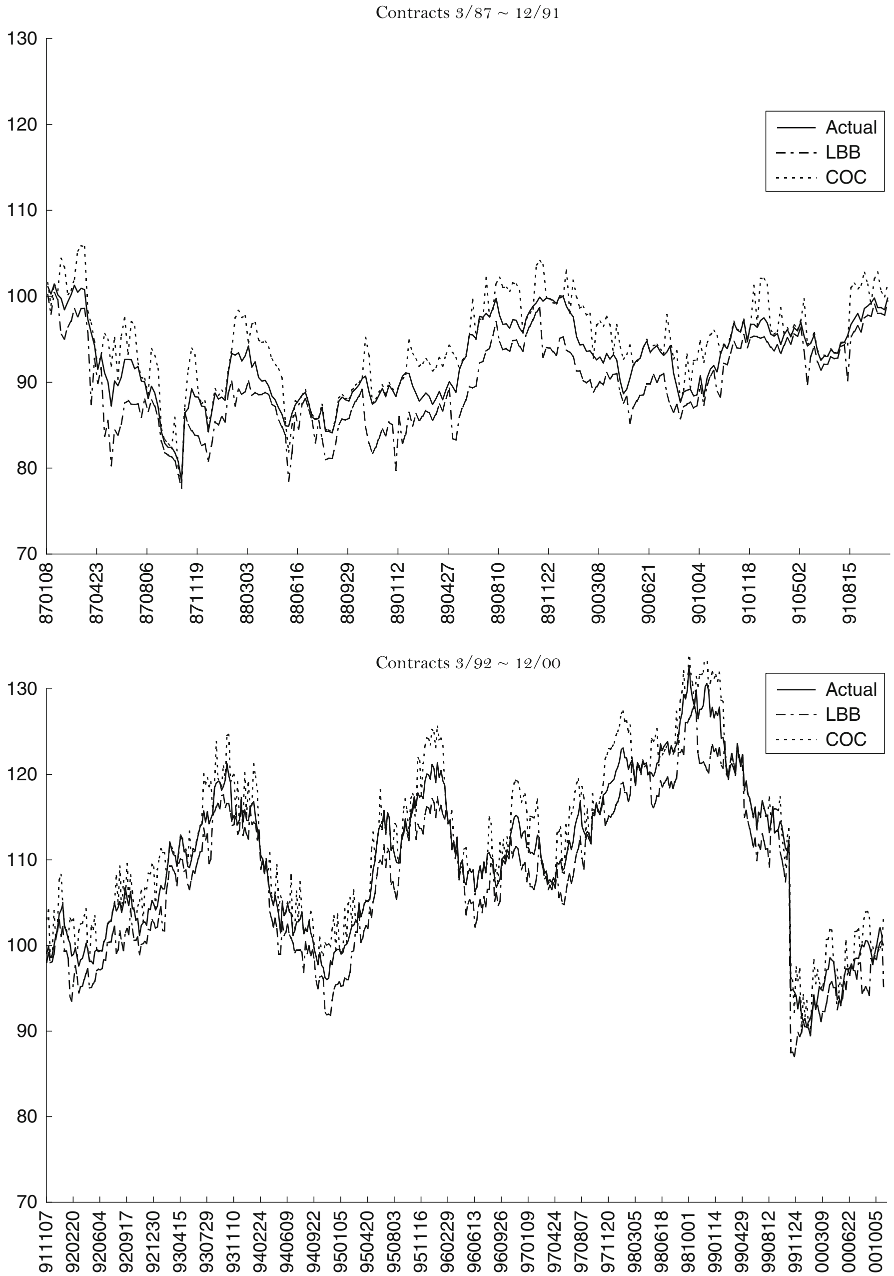


Fig. 71.3 Weekly time-series plot of actual futures prices (actual), their upper (COC) and lower (LBB) bounds

71.7 Discussions

The importance of the bounds becomes clear if one realizes that it is nearly impossible to compute the delivery options accurately. However, it is almost equally important to recognize that violating such bounds implies arbitrage opportunities. This is particularly interesting for the upper bound because the upper bound of the futures price – the cost of carry – is model-free. The calculation results show that the futures price exceeded its upper bound in 164 out of 722 (or 22.71 %) weeks. The magnitude of violation is on average 28.6 basis points (annualized) or half a basis point a week in the event that violation occurs. At such times, investors can sell futures and buy the CTD bond and then dynamically switch to the new CTD bond if necessary. Such a strategy, as suggested by the forward measure, should yield an arbitrage profit of half a basis point each time. This profit must exceed the transaction costs to be profitable.

The violation of the lower bound only generates arbitrage profits when the adopted model is true. The calculation results show that the lower bound is violated in 32 out of 722 cases (or 4.4 %). Notably, the fewer violations of the lower bound than upper bounds result mainly from the more conservative estimate of the lower bound (i.e., we compound upper bound values of the embedded delivery options). The results suggest that a more efficient lower bound remains to be discovered. Future research can examine this question.

From Fig. 71.3, we note that there are periods where bounds are tight and others where bounds are loose. To examine any potential systematic biases, this study runs regressions of the "tightness" of the bound against a number of possible factors that affect the futures price. For consistency, this study runs the following regressions[18] for the period of January 2, 1992, to November 2, 2000, containing 462 weekly observations:

$$\begin{aligned}
\Phi_t - \underline{\Phi}_t &= a_0 + a_1(\mathrm{CTD}_t - \mathrm{SCTD}_t) + a_2(3\mathrm{MTB}_t - 3\mathrm{MTB}_{t-1}) \\
&\quad + a_3(30\mathrm{YTB}_t - 30\mathrm{YTB}_{t-1}) + a_4(\mathrm{CF}_t) + e_t \\
\overline{\Phi}_t - \Phi_t &= b_0 + b_1(\mathrm{CTD}_t - \mathrm{SCTD}_t) + b_2(3\mathrm{MTB}_t - 3\mathrm{MTB}_{t-1}) \\
&\quad + b_3(30\mathrm{YTB}_t - 30\mathrm{YTB}_{t-1}) + b_4(\mathrm{CF}_t) + u_t
\end{aligned}$$

where
Φ = market futures price
$\overline{\Phi}$ = upper bound (COC)
$\underline{\Phi}$ = lower bound
CTD = cheapest to deliver bond
SCTD = second cheapest to deliver bond
3MTB = 3-month T bill rate
30YTB = 30-year T bond rate
CF = conversion factor of the cheapest bond

[18] Lee et al. (2000) provide excellent introductions about how to conduct multiple regressions.

Table 71.4 Regression results

	Lower bound			Upper bound		
	Coefficient	Std. err.	t	Coefficient	Std. err.	t
Intercept	2.620156	0.080332	32.61647	1.62892	0.097364	16.73019
CTD_t–$SCTD_t$	0.615754	0.240687	2.558313	0.043733	0.291717	0.149915
$3MTB_t$–$3MTB_{t-1}$	−0.85906	1.056233	−0.81332	−1.35587	1.280171	−1.05913
$30YTB_t$–$30YTB_{t-1}$	−1.55869	0.919818	−1.69456	−3.24591	1.114835	−2.91156
CF_t	−0.79337	0.970453	−0.81753	−0.42436	1.176205	−0.36079
Adjusted R^2	2.28 %			1.87 %		
# of obs.	462			462		

Regression period is from January 2, 1992 till November 2, 2000, total of 462 weekly observations. Regression equations are

$$\Phi_t - \underline{\Phi}_t = a_0 + a_1(\mathrm{CTD}_t - \mathrm{SCTD}_t) + a_2(3\mathrm{MTB}_t - 3\mathrm{MTB}_{t-1}) + a_3(30\mathrm{YTB}_t - 30\mathrm{YTB}_{t-1}) + a_4(\mathrm{CF}_t) + e_t$$

$$\overline{\Phi}_t - \Phi_t = b_0 + b_1(\mathrm{CTD}_t - \mathrm{SCTD}_t) + b_2(3\mathrm{MTB}_t - 3\mathrm{MTB}_{t-1}) + b_3(30\mathrm{YTB}_t - 30\mathrm{YTB}_{t-1}) + b_4(\mathrm{CF}_t) + u_t$$

where
Φ = market futures price
$\overline{\Phi}$ = upper bound (COC)
$\underline{\Phi}$ = lower bound
CTD = cheapest to deliver bond
SCTD = second cheapest to deliver bond
3MTB = 3-month T bill rate
30YTB = 30-year T bond rate
CF = conversion factor of the cheapest bond

and the results are reported in Table 71.4. Interestingly, the lower bound performance is more sensitive to the fitting of the second cheapest bond and the upper bound performance is more sensitive to the long rate. This result is unsurprising because the lower bound is a model-driven result while the upper bound is model-free and hence relies on the long rate.

Finally, this study argues that the timing options are more valuable during the first period than the second period. Notably, the timing options are negatively related to the interest rates. Lower interest rates during the second period reduce the value of the timing options.

71.8 Conclusion

This study derives lower and upper bound formulas for the Treasury bond futures prices. The lower bound of the futures price is obtained by integrating all upper bounds for the delivery options. The cost of carry model is found to be an upper bound of the futures price. These bounds are model free and can be used with any choice of the term structure model. Analytical results are obtained when using a two-factor Cox-Ingersoll-Ross model. The results provide investors with an efficient range of how far futures prices can move. During the two sample periods of 1987–1991 and 1992–2000, the cost of carry model is found to be about 2 % above the actual futures price and the lower bound is found to be about 2 % below.

As opposed to recursively using the lattice model to iteratively obtain an accurate estimate of the futures price, which is prohibitively expensive, as Boyle (1989) demonstrates, the bounds provided in this paper can be computed quickly and accurately. Thus, these bounds can then provide traders with a useful guide to the true futures price.

Appendix

From Theorem 1, we have:

$$\begin{aligned} E_t^Q[\delta(t,T)X(T)] &= E_t^Q[\delta(t,T)]E_t^{F(T)}[X(T)] \\ &= P(t,T)E_t^{F(T)}[X(T)] \end{aligned} \tag{71.37}$$

where δ is strictly less than 1. Due to the risk-neutral pricing result we have, the LHS must equal $X(t)$, and hence:

$$X(t) = \frac{E_t^{F(T)}[X(T)]}{P(t,T)} \tag{71.38}$$

Note that the forward measure is maturity dependent. Clearly, the Radon-Nikodym derivative (RND) is

$$\eta(t,T) = \frac{\delta(t,T)}{P(t,T)} \tag{71.39}$$

Since the measure is T-dependent, so should be the RND (usually, RND is just $\eta(t)$). Let the interest rate process be

$$dr(t) = \hat{\mu}(r,t)dt + \sigma(r,t)dW^Q(t) \tag{71.40}$$

Applying Ito's lemma,

$$
\begin{aligned}
0 &= \ln\ P(T,T) = \ln P(t,T) + \int_t^T \frac{1}{P(u,T)}\Bigg[P_u(u,T)du + P_r(u,T)dr \\
&\quad + \frac{1}{2}P_{rr}(u,T)(dr)^2\Bigg].d\hat{W}(u) - \int_t^T \frac{1}{2}\left[\frac{\sigma(r,u)P_r(u,T)}{P(u,T)}\right]^2 du \\
&= \ln\ P(t,T) + \int_t^T \frac{1}{P(u,T)}\left[P_u(u,T)du + P_r(u,T)\hat{\mu}(r,u) + \frac{1}{2}P_{rr}(u,T)\sigma(r,u)^2\right]du \\
&\quad + \int_t^T \frac{1}{P(u,T)}P_r(u,T)\sigma(r,u)d\hat{W}(u) - \int_t^T \frac{1}{2}\left[\frac{\sigma(r,u)P_r(u,T)}{P(u,T)}\right]^2 du \\
&= \ln\ P(t,T) + \int_t^T r(u)du + \int_t^T \frac{1}{P(u,T)}P_r(u,T)\sigma(r,u)d\hat{W}(u) - \int_t^T \frac{1}{2}\left[\frac{\sigma(r,u)P_r(u,T)}{P(u,T)}\right]^2 du
\end{aligned}
\tag{71.41}
$$

Letting:

$$
\theta(t,T) = -\frac{\sigma(r,t)P_r(t,T)}{P(t,T)} \tag{71.42}
$$

and moving the first two terms to the left:

$$
\begin{aligned}
-\int_t^T r(u)du - \ln\ P(t,T) &= \int_t^T -\theta(u,T)d\hat{W}(u)\ - \int_t^T \frac{1}{2}\theta(u,T)^2 du \\
\frac{\delta(t,T)}{P(t,T)} = \eta(t,T) &= \exp\left(\int_t^T -\theta(u,T)d\hat{W}(u) - \int_t^T \frac{1}{2}\theta(u,T)^2 du\right)
\end{aligned}
\tag{71.43}
$$

This implies the Girsanov transformation of the following:

$$
W^{F(T)}(t) = W^Q(t) + \int_t^T \theta(u)dt\ = W^Q(t) - \int_t^T \sigma(r,u)\frac{P_r(u,T)}{P(u,T)}du \tag{71.44}
$$

The interest rate process under the forward measure henceforth becomes:

$$
dr(t) = \left[\hat{\mu}(r,t) + \sigma(r,t)^2\frac{P_r(t,T)}{P(t,T)}\right]dt + \sigma(r,t)dW^{F(T)}(t) \tag{71.45}
$$

Note that the forward measure is quite general. It does not depend on any specific assumption on the interest rate process.

References

Arak, M., & Goodman, L. (1987). Treasury bond futures: Valuing the delivery option. *Journal of Futures Markets, 7*(3), 269–286.

Benninga, S., & Smirlock, M. (1985). An empirical analysis of the delivery option, marking to market, and the pricing of Treasury bond futures. *Journal of Futures Markets, 5*(3), 361–374.

Boyle, P. (1989). The quality option and timing option in futures contracts. *Journal of Finance, 44*(1), 101–113.
Broadie, M., & Sundaresan, S. (1987) The pricing of timing and quality options: An application to Treasury bond futures markets. Working paper.
Carr, P. (1988) Valuing bond futures and the quality option. Working paper, University of California at Los Angeles.
Carr, P., & Chen, R. (1996). Valuing bond futures and the quality option. Working paper, Rutgers University, November.
Chen, R., & Scott, L. (1993). Maximum likelihood estimation of a multi-factor equilibrium model of the term structure of interest rates. *Journal of Fixed Income, 3*(3), 14–32.
Chen, R., & Yeh, S. K. (2002). Analytical bounds for American option prices. *Journal of Financial and Quantitative Analysis, 37*(01), 117–135.
Cox, J., Ingersoll, J., & Ross, S. (1981). The relation between forward prices and futures prices. *Journal of Financial Economics, 9*(4), 321–346.
Cox, J., Ingersoll, J., & Ross, S. (1985). A theory of the term structure of interest rates. *Econometrica, 53*(2), 385–407.
Duan, J., & Simonato, J. (1999). Estimating and testing exponential-affine term structure models by Kalman Filter. *Review of Quantitative Finance and Accounting, 13*, 111–135.
Gay, G., & Manaster, S. (1986). Implicit delivery options and optimal delivery strategies for financial futures contracts. *Journal of Financial Economics, 16*(1), 41–72.
Hamiltion, J. (1994). *Time series analysis*. Princeton: Princeton University Press.
Heath, D., Jarrow, R., & Morton, A. (1992). Bond pricing and the term structure of interest rates: A new methodology for contingent claims valuation. *Econometrica, 60*(1), 77–105.
Hedge, S. (1988). An empirical analysis of implicit delivery options in the Treasury bond futures contract. *Journal of Banking and Finance, 12*(2), 469–492.
Hedge, S. (1990). An ex post valuation of the quality option in Treasury bond futures contract. *Journal of Banking and Finance, 14*(4), 741–760.
Hemler, M. (1990). The quality delivery option in Treasury bond futures contracts. *Journal of Finance, 45*(5), 1565–1586.
Ho, T., Lee, S., & Son, K. (1992). Term structure estimation and pricing of callable Treasury bonds. *Review of Quantitative Finance and Accounting, 2*(2), 127–143.
Hull, J. (2009) *Options, futures, and other derivative contracts* (7th ed.) Prentice Hall.
Hull, J., & White, A. (1990). Pricing interest-rate-derivative securities. *The Review of Financial Studies, 3*(4), 573–592.
Jagannathan, R., Kaplin, A., & Sun, S. (2003). An evaluation of multi-factor CIR models using LIBOR, swap rates, and cap and swaption prices. *Journal of Econometrics, 116*, 113–146.
Jamshidian, F. (1989). An exact bond option formula. *Journal of Finance, 4*(1), 205–209.
Kane, A., & Marcus, A. (1984). Conversion factor risk and hedging in the Treasury bond futures market. *Journal of Futures Markets, 4*(1), 55–64.
Kane, A., & Marcus, A. (1986). Valuation and optimal exercise of the wild card option in the Treasury bond futures market. *Journal of Finance, 41*(1), 195–207.
Kilcollin, T. (1982). Difference systems in financial futures markets. *Journal of Finance, 37*(5), 1883–1897.
Langetieg, T. (1980). A multivariate model of the term structure. *Journal of Finance, 35*(1), 71–97.
Lee C., Lee A., & Lee J. (2000). *Statistics for business and financial economics* (2nd ed). World Scientific Publishers.
Livingston, M. (1987). The delivery option on forward contracts. *Journal of Financial and Quantitative Analysis, 22*(01), 79–87.
Lo, A. W. (1987). Semi-parametric upper bounds for options prices and expected payoffs. *Journal of Financial Economics, 19*(2), 373–387.
Longstaff, F., & Schwartz, E. (1992). Interest rate volatility and the term structure: A two-factor general equilibrium model. *Journal of Finance, 47*(4), 1259–1282.

Margrabe, W. (1978). The value of an option to exchange one asset for another. *Journal of Finance, 33*, 177–186.

Ritchken, P., & Sankarasubramanian, L. (1992). Pricing the quality option in Treasury bond futures. *Mathematical Finance, 2*(3), 197–214.

Turnbull, S., & Milne, F. (1991). A simple approach to interest-rate option pricing. *The Review of Financial Studies, 4*(1), 87–120.

Vasicek, O. (1977). An equilibrium characterization of the term structure. *Journal of Financial Economics, 5*(2), 177–188.

Yeh, S. K., & Lin, B. H. (2003). Term structure fitting and information content: An empirical study on the Taiwanese government bond market. *Review of Pacific Basin Financial Markets and Policies, 6*(3), 1–44.

Zhang, P. G. (1994). Bounds for option prices and expected payoffs. *Review of Quantitative Finance and Accounting, 4*(2), 179–197.

Rating Dynamics of Fallen Angels and Their Speculative Grade-Rated Peers: Static vs. Dynamic Approach

72

Huong Dang

Contents

Abstract

This study adopts the survival analysis framework (Allison, P. D. (1984). *Event history analysis*. Beverly Hills: Sage) to examine issuer-heterogeneity and time-heterogeneity in the rating migrations of fallen angels (FAs) and their speculative grade-rated peers (FA peers). Cox's hazard model is considered the

H. Dang
University of Canterbury, Christchurch, New Zealand
e-mail: huong.dang@canterbury.ac.nz

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_72,

preeminent method to estimate the probability that an issuer survives in its current rating grade at any point in time t over the time horizon T. In this study, estimation is based on two Cox's hazard models, including a proportional hazard model (Cox, Journal of Royal Statistical Society Series B (Methodological) 34:187–220, 1972) and a dynamic hazard model. The first model employs a static estimation approach and time-independent covariates, whereas the second uses a dynamic estimation approach and time-dependent covariates. To allow for any dependence among rating states of the same issuer, the marginal event-specific method (Wei et al., Journal of The American Statistical Association 84:1065–1073, 1989) was used to obtain robust variance estimates. For validation purpose, the Brier score (Brier, Monthly Weather Review 78(1):1–3, 1950) and its covariance decomposition (Yates, Organizational Behaviour and Human Performance 30:132–156, 1982) were applied to assess the forecast performance of estimated models in forming time-varying survival probability estimates for issuers out of sample.

It was found that FAs and their peers exhibit strong but markedly different dependences on rating history, industry sectors, and macroeconomic conditions. These factors jointly, and in several cases separately, are more important than the current rating in determining future rating changes. A key finding is that past rating behaviors persist even after controlling for the industry sector and the evolution of the macroeconomic environment over the time for which the current rating persists. Switching from a static to a dynamic estimation framework markedly improves the forecast performance of the upgrade model for FAs. The results suggest that rating history provides important diagnostic information and different rating paths require different dynamic migration models.

Keywords

Survival analysis • Hazard model • Time-varying covariate • Recurrent event • Brier score • Covariance decomposition • Rating migration • Fallen angel • Markov property • Issuer-heterogeneity • Time-heterogeneity

72.1 Introduction

Market participants are interested in the rating migration propensity of fallen angels (FAs) which were initially rated as investment grades but experienced deterioration in credit quality and fell to speculative grades. Such dramatic downgrades are consistently associated with a statistically significant negative return in the stock market (Holthausen and Leftwich 1986; Ederington and Goh 1999) and in the bond market (Hite and Warga 1997). Furthermore, bad news associated with a rating downgrade is transferred from the downgraded company to its rivals. The rivals of downgraded firms with speculative grade-rated debts experience significant downward revisions of earnings forecasts (Caton and Goh 2003).

Institutional investors such as investment grade bond funds, due to restricted risk levels, may not be able, or may be limited, to hold a small percentage of speculative grade-rated bonds (Altman and Kao 1992b, p. 73; Cantor and Packer 1997). They may be forced to sell bonds that drop into the speculative grade ratings (fallen angels) and may buy back bonds that regain the investment grade status (rising stars). Accurately estimating the migration probabilities of FAs is important in the context of portfolio construction and credit risk management. Tsaig et al. (2011) showed that credit migration can explain as much as 51 % of volatility and 35 % of economic risk capital for a typical loan portfolio.

In practice, the discrete time cohort Markov approach has been widely used by credit rating agencies to model rating migration processes. This approach assumes that the departure probability out of the current rating depends entirely on the current rating (Markov process) and remains constant over time (time-homogeneous). The time-homogeneous Markov property, however, is not strongly supported by empirical studies. Frydman and Schuermann (2008) found that obligors of the same rating grade migrate at different rates and the heterogeneity persists after controlling for the state of the business cycle or the industry sector. The sources of issuer-heterogeneity in the rating process can be attributed to several aspects of rating history such as the direction of the prior rating change,[1] the duration of the previous rating states,[2] the first rating received,[3] and the period of time since first rated.[4] Failing to consider issuer-heterogeneity can result in inaccurate estimates of value at risk and a misleading picture of the economic risk capital (Kadam and Lenk 2008).

An interesting question then arises: Does one migration model fit all, or do different rating histories require different models? Specifically, do FAs exhibit the same issuer dependence and time dependence as their speculative grade-rated peers? Literature confirms some established perceptions about FAs and supports the notion that FAs exhibit different rating dynamics compared to their peers.[5]

This study adopts the survival analysis framework (Allison 1984) to examine a wide variety of issuer-heterogeneity and time-heterogeneity in the rating dynamics of US nonfinancial FAs and their speculative grade-rated peers. I estimate rating migration models of FAs, after their fall dates. Comparator speculative grade-rated issuers that experienced a downgrade but not a FA event at lag-one rating state (FA peers) are identified, and rating migration models are estimated for these peers. FAs that are further downgraded are compared to speculative grade-rated peers experiencing a downgrade, and FAs that regain investment grade status are compared to their peers experiencing an upgrade.

Cox's hazard model (Cox 1972) is considered a preeminent method to estimate the probability that an issuer survives in its current rating grade (*start rating*) at any

[1]See, for example, Altman and Kao (1992b), Carty and Fons (1994), Lando and Skodeberg (2002), Hamilton and Cantor (2004), and Figlewski et al. (2012).

[2]See Carty and Fons (1994) and Lando and Skodeberg (2002).

[3]See, for example, Altman and Kao (1991, 1992a, b).

[4]See Altman (1998) and Figlewski et al. (2012).

[5]See Mann et al. (2003), Vazza et al. (2005a), and Figlewski et al. (2012).

point in time t over the time horizon T. In this study, estimation is based on two hazard models including a Cox's proportional hazard model (Cox 1972) and a Cox's dynamic hazard model. The Cox's proportional hazard model (static model) employs a static estimation framework and incorporates time-independent covariates. The Cox's dynamic hazard model (dynamic model) departs from the conventional Cox's proportional hazard model by adopting a dynamic estimation framework and including both time-independent and time-varying covariates. The dynamic model captures the evolution of the macroeconomic environment over the time during which the current rating persists. Time-varying survival probability estimates were generated by the static and the dynamic hazard models for FAs and FA peers out of sample. The Brier score (Brier 1950) was used to assess the predictive accuracy of survival probability estimates. The covariance decomposition of the Brier score (Yates 1982) provides insights into the sources of forecast errors.

The results of the proportional and dynamic hazard models for upgrades and downgrades of FAs and their peers offer an improved understanding of the following questions: First, how can rating history, after controlling for the industry sector and the economic environment, explain and predict subsequent rating changes for FAs and their peers? Second, does the development of economic factors over rating durations matter, and if so, how does it affect the migration hazards of FAs and FA peers? Third, does the impact of rating history on the migration process persist in a dynamic estimation framework? Fourth, does the dynamic estimation framework improve the forecast performance of rating history?

Comparing the proportional and the dynamic hazard models for FA issuers with the respective models for FA peers, the results show that the significant variables differ. In a dynamic estimation framework, the downgrade process of FAs exhibits strong dependence on macroeconomic factors, whereas the upgrade process is entirely determined by the current rating and rating history. For FA peers, downgrades are substantially influenced by past rating behaviors, whereas upgrades are strongly impacted by the industry sector. It is clear that issuer-heterogeneity (rating history, industry sectors) and time-heterogeneity (macroeconomic environment) jointly, and in several cases separately, are more important than the current rating in determining future rating changes. A key finding is that past rating behaviors persist even after controlling for the industry sector and the evolution of the macroeconomic climate over rating durations. Switching from a static to a dynamic estimation framework markedly improves the forecast performance of the upgrade hazard model for FAs.

The contribution of this study is threefold. First, it enriches and strengthens the evidence of issuer-heterogeneity and time-heterogeneity in the rating dynamics of FAs and their peers observed in both static and dynamic estimation frameworks. The study incorporates well-documented empirical properties and presents new evidence on additional aspects of rating history which has received little attention in previous studies. Second, the study contributes to the framework for estimating rating migration models. Cox's hazard model with time-varying covariates is well suited to model recurrent rating migration events. The study overcomes computational challenges in forming dynamic survival estimates when the standard proportionality assumption of the Cox's hazard model does not hold. The dynamic predictions allow the survival probability estimates of holdout issuers to vary over the forecast horizons. Finally, the study enriches the literature on the framework for evaluating probability forecasts. The Brier score has received little attention in finance studies[6] though it is a popular measure for ex post evaluation of probability forecasts in meteorology (Murphy and Winkler 1977; Winkler 1996). The covariance decomposition (Yates 1982) of the Brier score offers the possibility to evaluate forecast characteristics in terms of discrimination, calibration, and variance.

This study is structured as follows: Sect. 72.2 provides a discussion of the literature on rating migration dynamics. Section 72.3 presents the method used, followed by a description of the data in Sect. 72.4. Section 72.5 summarizes the results of the models for FAs and their peers. Section 72.6 presents the Brier scores and Yates's covariance decompositions. Section 72.7 summarizes the main findings.

72.2 Literature Review

This section provides a summary of the literature on corporate rating dynamics of FAs and their speculative graded-rated peers, with a focus on issuer-heterogeneity (rating history, industry sector) and time-heterogeneity (business cycle).

The literature suggests that various aspects of rating history such as rating momentum, duration dependence, aging effect, and the first rating impact on future rating distribution. Issuers downgraded to a given rating, compared with those upgraded to the same rating grade, are more likely to drop to lower rating categories (Altman and Kao 1992b; Carty and Fons 1994; Bangia et al. 2002; Lando and Skodeberg 2002; Hamilton and Cantor 2004; Mah and Verde 2004; Figlewski et al. 2012). Issuers of different lagged rating durations have different departure probabilities out of the current rating (Carty and Fons 1994; Lando and Skodeberg 2002). Newly rated firms, compared with seasoned firms of the same rating class, exhibit a smaller probability of rating migrations within a few years (Altman and Kao 1991; Altman 1992, 1998). The longer since a firm was first rated, the more likely it will default (Figlewski et al. 2012). Furthermore, new issues of different ratings retain their original ratings in a different manner and their rating stability varies over time. For example, originally BB-rated bonds, which were in the speculative grade barrier, exhibit the least stability in retaining the initial rating and do not show a clear propensity to migrate in either direction (Altman and Kao 1991, pp. 19–20, 1992b, pp. 65–67).

[6]Samuelson and Rosenthal (1986), Bessler and Ruffley (2004), Yao et al. (2005), Grunert et al. (2005), and Dang (2010) are among the few studies in finance that applied this scoring rule to assess the predictive accuracy of estimated models. Johnstone (2002) suggested that the Brier score performs better than categorical measures in accurately assessing forecast performance.

For a FA, the rating it received on the fall date strongly impacts on its future rating changes (Mann et al. 2003). FAs which plunge to the speculative grade barrier (Moody's Ba or Standard & Poor's BB class) exhibit a greater propensity to rise to the investment grade territory and are less vulnerable to default. The higher the rating a FA received prior to and after the fall date, the more likely it will become a "prodigal son" and regain investment grade status. FAs, during the initial years of financial distress, experience strong downward momentum (Mann et al. 2003) and tend to travel multiple notches downward in a drastic and quick succession rather than through gradual, mild steps (Johnson 2004). This pattern is understandable as once FAs lose their investment grade status, their operations are impaired due to regulations or private contracts (Cantor and Packer 1997; Standard & Poor's 2001).

Industry risk is an important determinant of rating distribution. In general, ratings from the utility sector are more stable than those from the industrial sector due to a smaller volatility of future revenues (Kadam and Lenk 2008). For FAs, utilities are more likely to be upgraded for every horizon, whereas nonutilities do not exhibit exceptional performance or a favorable tendency towards upgrades (Altman and Kao 1992a, p. 19). The magnitude of subsequent rating changes for nonutility FAs is dramatic. Public utility FAs exhibit strong negative serial correlation and are more likely to rise up the rating scales after the fall date. In contrast, nonutility FAs display a positive serial correlation and tend to continue the downward journey after losing the investment grade status (Altman and Kao 1992a). FAs in high-velocity sectors such as telecommunications exhibit rapid deterioration in their credit profile. Distressed FAs tend to be weeded out during shakeouts while surviving FAs cling to life and show strong recovery within 5 years from their fall date. In contrast, FAs in low-velocity sectors such as leisure and media experience slow decline in creditworthiness, but they do not make a strong rebound in the subsequent years (Vazza et al. 2005a, p. 17).

In terms of time-heterogeneity, speculative grade-rated issuers exhibit different migration patterns as time extends (Carty and Fons 1994). Issuers of low credit ratings such as B and Caa are vulnerable to a downgrade in the short term and have a migration probability that decreases with time. If they survive but fail to substantially improve their credit quality, they tend to retain their existing ratings over a long time. Issuers in the middle rating grades such as Ba can either go up or go down the rating scale and have a migration probability that varies constantly with time.

FAs, following the fall date, exhibit a faster rate of migration, a greater likelihood to default, and a shorter median time to default compared with their peers[7] (Mann et al. 2003; Vazza et al. 2005a). However, as time passes FAs gained a robust franchise value, enhanced business strength, and improved profitability (Vazza et al. 2005a).

[7]Vazza et al. (2005a) defined FA peers as those originally rated in speculative grades and have identical rating distribution characteristics as FAs. Mann et al. (2003) defined FA peers as speculative grade-rated issuers that were of the same ratings as FAs at the time they lost investment grade status and never rated in the investment grade spectrum.

In the long term, FAs are less risky, display a greater tendency to survive, and are more likely to rise back to the investment grade territory. Non-defaulted FAs appear to cling to life for many more years than their peers.

The time-heterogeneity in the rating migration dynamics of FAs and their peers can partly be attributed to macroeconomic conditions. The state of the economy is a major driver of systematic credit risk (Blume et al. 1991), and low ratings are more sensitive to business cycles than high ratings. The occurrence of downgrades and the generation of FAs soar as market conditions deteriorate (Nickell et al. 2000; Vazza et al. 2005a). Furthermore, rating volatility increases during business cycle troughs and decreases during peaks. Thus, failure to account for the state of the economy may result in an underestimation of "downward potential of high-yield portfolio" in contraction periods or "suboptimal capital allocation in lending business" (Bangia et al. 2002, p. 469).

The above evidences emphasize the need to account for rating history, industry sector, and the business cycle in modelling the rating migrations of FAs and their peers.

72.3 Models

72.3.1 Rating States

The substantive processes that govern the occurrence and timing of downgrade and upgrade events to FAs/FA peers were examined using the survival analysis framework (Allison 1984; Blossfeld et al. 1989; Yamaguchi 1991; Hosmer et al. 2008; Dang 2010; Figlewski et al. 2012). A rating state starts from the time an issuer is downgraded to a rating class (*start rating*) subsequent to the commencement date of the study. For a FA, the *start rating* is also the rating it receives on the fall date upon its entrance to speculative grade territory. The rating state ends at the time the issuer migrates to another rating class (*end rating*), withdraws from being rated, or the observation period terminates. The time a firm keeps the same rating grade is the survival time (survival duration). If a firm exits from a rating class due to any reason other than the migration event of interest, the survival time is treated as censored. In the upgrade model, down states (rating states with the *start rating* better than the *end rating*) are censored and vice versa for the downgrade model. Rating states finishing after the end of the model estimation period are also treated as censored. Survival analysis addresses this type of censoring by examining only that part of the duration in the estimation period (Yamaguchi 1991; Blossfeld and Rohwer 1995).

Rating states are pooled across issuers and time. Rating states that pass the screening test of having experienced at least two migrations (i.e., non-censored lag-one and lag-two rating states[8]) will remain in the sample and be incorporated in

[8]To get enough observations to make meaningful inference of the effect of rating history, the study considers lag-one and lag-two rating states. By definition, all FAs and FA peers in the study experienced a downgrade at lag-one rating state.

the estimation process. The estimation procedure makes use of event time risk sets, which are composed of all the firm ratings that are at risk of a rating change at event time t. In the process of estimating the model, a new risk set is formed at each event time t when a rating transition occurs. Firm ratings leave the risk set once they experience a rating transition or when they are censored.

An issuer may contribute several rating transitions to the dataset. The presence of repeated rating transitions for the same firm is likely to introduce dependence among the observations. This problem is reduced to the extent that covariates in the models control for dependence. To allow for any dependence, the marginal event-specific method (Wei et al. 1989) is used to obtain robust variance estimates. Under this approach, rating states contribute to an event time risk set as long as they are under observation at the event time t the risk set is formed. The event time risk set arrangement "resets the clock" after a migration event occurs and time is measured from the last downgrade event. The emphasis is on the duration between the last downgrade event and the subsequent migration event and "each migration event is analyzed as a separate process" (Hosmer et al. 2008, pp. 290–294).

72.3.2 Cox's Hazard Models

Cox's proportional and dynamic hazard models were developed for two generic migration outcomes (upgrade and downgrade) and were estimated separately for FAs and FA peers over the period 1984–2000. Survival analysis in general and Cox's hazard model in particular offer several advantages over the discrete time cohort Markov approach widely used by credit rating agencies. The hazard model uses both completed transitions and censored observations in the estimation process (Yamaguchi 1991), resulting in consistent parameter estimates (Allison 1995). It does not make any assumptions about the distribution of survival times (Allison 1995) and provides descriptive information of the survivor function of event times (Lee 1980). It is possible to incorporate time-varying covariates into the model to capture the evolution of risk factors over time (Allison 1995, p. 183). It also allows a rigorous testing of issuer-heterogeneity and time-heterogeneity in rating dynamics (Lando and Skodeberg 2002).

72.3.2.1 Cox's Proportional Hazard Model

The Cox's proportional hazard model (Cox 1972) can be expressed as follows:

$$h_m(t, Z) = h(0, t)\exp\left[Z_j^m \beta_{\mathrm{j}}\right] \tag{72.1}$$

where $h_m(t, Z)$ is the migration hazard of rating state m at time t given its time-fixed covariate vector Z_j^m. $h(0,t)$ is the unspecified nonnegative baseline hazard, which is the hazard with the covariate vector set to zero, at time t. β_j is the vector of estimated coefficients for time-fixed covariate vector Z_j^m.

The likelihood $L_{t_m}^m$ that rating state m experiences a rating migration of interest at ordered time t_m is calculated as follows:

$$L_{t_m}^m = \frac{\exp\left(\beta_j Z_j^m\right)}{\sum_{i \in R(t_m)} \exp\left(\beta_j Z_j^i\right)} \tag{72.2}$$

where i represents a rating state in the risk set formed at ordered event time $t_{m,}$ $R(t_m)$.

The baseline hazard $h(0,t)$ cancels out in the numerator and denominator when forming the likelihood function and is not required in the estimation process. Taking the product of the likelihoods, for all states that migrated, across N ordered migration times t_m gives the full partial likelihood, PL, as follows:

$$PL = \prod_{m=1}^{N} L_{t_m}^m = \prod_{m=1}^{N} \left[\frac{\exp\left(\beta_j Z_j^m\right)}{\sum_{i \in R(t_m)} \exp\left(\beta_j Z_j^i\right)} \right] \tag{72.3}$$

The vectors of the estimated coefficients $\hat{\beta}_j$ can be obtained in the absence of knowledge of the baseline hazard $h(0,\ t)$ by maximizing the full partial likelihood in Eq. 72.3 (Kalbfleisch and Prentice 1980; Namboodiri and Suchindran 1987; Lawless 2003). Appendix 1 presents further details of the estimation approach.

The Cox's proportional hazard model has the property that the hazards for any two firms m and n in the risk set $R(t_m)$ does not vary between observed event times (Lawless 2003). Taking the ratio of the hazards for firms m and n and applying Eq. 72.1,

$$\frac{h_m(t,Z)}{h_n(t,Z)} = \frac{h(0,t)\exp\left[Z_j^m \beta_{\mathrm{j}}\right]}{h(0,t)\exp\left[Z_j^n \beta_{\mathrm{j}}\right]} = \exp\left[\beta_j\left(Z_j^m - Z_j^n\right)\right] \tag{72.4}$$

In the presence of time-fixed covariates Z_j, this property can be used to derive the estimated baseline hazard $\hat{h}(0,t)$ (see Lawless 2003).

The estimated hazard function of holdout state q at time t, $\hat{h}_q[t,Z]$, can be obtained by substituting in Eq. 72.1 the estimated baseline hazard $\hat{h}(0,t)$ derived from in-sample analysis, the estimated coefficient vector $\hat{\beta}_j$ obtained from Eq. 72.3, and state q's actual covariate vector Z_j^q. The estimated survival function of holdout state q at time t is

$$\hat{S}_q[t,Z] = \exp - \left(\sum \hat{h}_q[t,Z]\right) \tag{72.5}$$

72.3.2.2 Dynamic Cox's Hazard Model

The dynamic Cox's hazard model can be expressed as follows:

$$h_m[t, Z, Z(t)] = h(0, t)\exp\left[Z_j^m \beta_\mathrm{j} + Z_p^m(t)\beta_\mathrm{p}\right] \tag{72.6}$$

where $h_m(t, Z, Z(t))$ is the migration hazard of rating state m at time t given its time-fixed covariate vector Z_j^m and its time-varying covariate vector $Z_p^m(t)$. $h(0, t)$ is the unspecified nonnegative baseline hazard at time t. β_p is the vector of estimated coefficients for time-varying covariate vector $Z_p^m(t)$, and β_j is the vector of estimated coefficients for time-fixed covariate vector Z_j^m.

The likelihood $L_{t_m}^m$ is constructed similarly as in Eq. 72.2:

$$L_{t_m}^m = \frac{\exp\left[\beta_j Z_j^m + \beta_p Z_p^m(t_m)\right]}{\sum_{i \in R(t_m)} \exp\left[\beta_j Z_j^i + \beta_p Z_p^i(t_m)\right]} \tag{72.7}$$

The full partial likelihood PL is constructed similarly as in Eq. 72.3:

$$PL = \prod_{m=1}^{N} \left[\frac{\exp\left[\beta_j Z_j^m + \beta_p Z_p^m(t_m)\right]}{\sum_{i \in R(t_m)} \exp\left[\beta_j Z_j^i + \beta_p Z_p^i(t_m)\right]} \right] \tag{72.8}$$

The estimation process requires the updated values of time-varying covariates $Z_p(t)$ at each event time for all states i in the risk set formed at that event time (Andersen 1992). The vectors of the estimated coefficients $\hat{\beta}_j$ and $\hat{\beta}_p$ can be obtained by maximizing the full partial likelihood in Eq. 72.8.

In the presence of the time-varying covariates $Z_p(t)$, the ratio of the hazards for firms m and n is not constant between observed event times:

$$\frac{h_m[t, Z, Z(t)]}{h_n[t, Z, Z(t)]} = \frac{h(0, t)\exp\left[Z_j^m \beta_\mathrm{j} + Z_p^m(t_m)\beta_\mathrm{p}\right]}{h(0, t)\exp\left[Z_j^n \beta_\mathrm{j} + Z_p^n(t_m)\beta_\mathrm{p}\right]} \tag{72.9}$$

$$\frac{h_m[t, Z, Z(t)]}{h_n[t, Z, Z(t)]} = \exp\left\{\beta_j\left(Z_j^m - Z_j^n\right) + \beta_p\left[Z_p^m(t_m) - Z_p^n(t_m)\right]\right\}$$

It is therefore not possible to extract the baseline hazard $h(0, t)$ from the dynamic Cox's hazard model regression results. To resolve this issue, this study takes an approach proposed by Andersen (1992) and adapts the SAS codes published by Chen et al. (2005). The integrated baseline hazard function $H(0, t)$ can be estimated given the vectors of the estimated coefficients $\hat{\beta}_p$ and $\hat{\beta}_j$ obtained from Eq. 72.8:

$$\hat{H}(0,t) = \sum_{t_m \le t} \frac{D_m}{\sum_{i \in R(t_m)} \exp\left(\hat{\beta}_j Z_j^i + \hat{\beta}_p Z_p^i(t_m)\right)} \tag{72.10}$$

where D_m is the indicator for whether the migration event occurred to state m at ordered time t_m within the interval $[0, t]$.

The integrated baseline hazard function $H(0, t)$ can also be expressed as a step function discontinued at ordered event time t_m (Chen et al. 2005):

$$H(0,t) = \sum_{t_m \in t} [h(0, t_{m-1})(t_m - t_{m-1})] \tag{72.11}$$

The estimated baseline hazard function $\hat{h}(0,t)$ can be derived from Eqs. 72.10 and 72.11. The use of a step function is well suited to speculative grade-rated issuers given their volatile ratings and rapid migration propensity. The narrow gaps between successive event times allow relatively accurate estimation of the baseline hazards.

The estimated hazard function of holdout state q at time t, $\hat{h}_q[t, Z, Z(t)]$, can be obtained by substituting in Eq. 72.6 the estimated baseline hazard function $\hat{h}(0,t)$ derived from Eqs. 72.10 and 72.11, the estimated coefficient vector $\hat{\beta}_p$ and $\hat{\beta}_j$ obtained from Eq. 72.8, and state q's actual covariate vector Z_j^q and $Z_p^q(t)$. The estimated survival function of holdout state q at time t is

$$\hat{S}_q[t, Z, Z(t)] = \exp - \left(\sum \hat{h}_q[t, Z, Z(t)]\right) \tag{72.12}$$

72.3.3 Forecast Evaluation

72.3.3.1 Forecast Horizons

The conventional forecast horizon is 1 year, and credit rating agencies, in practice, publish annual transition matrices. In this study, survival probability estimates for holdout issuers were formed at 1- and 2-year horizons. These forecast horizons were selected for the following reasons: First, there are noticeable concentrations of estimation rating states in survival durations of 1–2 years, resulting in consistent estimates of the baseline hazard at time $t = 1$ year and $t = 2$ years. The availability of holdout observations at 1- and 2-year horizons allows an unbiased accuracy assessment of the models' forecast performance. Second, as implied by Mann et al. (2003), the 2-year time frame highlights the relative risk of FAs as compared to their peers. The 1- and 2-year forecast horizons also provide a short-term and near-intermediate-term prospective which is relevant to speculative grade-rated issuers. Third, it is of particular interest to examine the predictive accuracy of rating history at 1- and 2-year horizons given the evidence that the Markov property adequately holds within 1 or 2 years (Kiefer and Larson 2007; Frydman and Schuermann 2008).

72.3.3.2 Brier Score

The actual survival status of each issuer in the holdout sample was recorded and mapped against the survival probability estimate generated by the static/dynamic hazard models at forecast time t. The Brier score (Brier 1950) at forecast time t, B_t, is defined as follows:

$$B_t = \frac{\sum_{q=1}^{N_t} \left(f_t^{state\_q} - a^{state\_q} \right)^2}{N_t} \tag{72.13}$$

where

N_t is the number of survival forecasts at time t (the number of holdout rating states at time t)

$f_t^{state\_q}$ indicates the survival probability forecast that the holdout state q will survive at forecast time t[9]

a^{state_q} is the known survival outcome of holdout state q. If holdout state q survives, $a^{state\_q} = 1$, and if holdout state q experienced the migration event of interest (i.e., a downgrade in the downgrade model or an upgrade in the upgrade model), $a^{state\_q} = 0$.

72.3.3.3 Covariance Decomposition

Using the covariance decomposition (Yates 1982), the Brier score can be broken down into skill components including slope (discrimination), bias (calibration), and scatter (variance).[10] The most basic form of the covariance decomposition of Brier score at the forecast time t, B_t, is given as

$$B_t = \underbrace{\overline{d}_t\left(1 - \overline{d}_t\right)}_{\text{Uncertainty}} + \underbrace{\left(\overline{f}_t - \overline{d}_t\right)}_{\text{Bias square}}{}^2 + \underbrace{S_{f_t}^2}_{\text{Scatter}} \underbrace{-2S_{f_t d_t}}_{\text{Covariance}} \tag{72.14}$$

where

$\overline{d}_t$, or $\overline{d}$ for short, is the overall mean survival index, or the survival base rate at time t

$\overline{d}_t\left(1 - \overline{d}_t\right)$, or $\overline{d}\left(1 - \overline{d}\right)$ for short, is the variance of the outcome index at the forecast time t

$\overline{f}_t$, or $\overline{f}$ for ease of notation, is the overall mean survival forecast at the forecast time t

$\left(\overline{f}_t - \overline{d}_t\right)$, or $\left(\overline{f} - \overline{d}\right)$ for short, is the bias of forecasts at the forecast time t

[9]The notation was changed from $\hat{S}_q[t, Z]$ Eq. 72.5 or $\hat{S}_q[t, Z, Z(t)]$ Eq. 72.12 to $f_t^{state\_q}$ to provide a compact presentation of the formula in a form consistent with the literature review on the Brier score.

[10]The covariance decomposition proposed by Yates (1982) provides components of forecast accuracy that are more basic than the Sander (1963) and Murphy (1973) decompositions (Yates 1982, p. 141).

$S^2_{f_t}$, or S^2_f for short, is the variance of the forecasts, or scatter, at the forecast time t
$S_{f_t d_t}$, or S_{fd} for short, is the covariance of the survival outcome index and the survival probability forecasts at the forecast time t.

The component, $\bar{d}(1-\bar{d})$, namely, the outcome index variance, is determined by "natural forces." It reflects an aspect of forecast accuracy that does not depend on the predictive power of the estimated model (Yates 1982, p. 139).

The first skill component, $(\bar{f}-\bar{d})$, or bias, indicates the ability of the estimated model to match the overall mean survival probability forecast $\bar{f}$ to the mean survival outcome index $\bar{d}$ (Yates 1982). Bias reflects the overall pessimism or optimism of the forecaster in assigning probability survival forecasts to holdout observations. This term can be either positive or negative. The smaller the absolute value of bias, the lower the Brier score.

The second skill component, scatter or S^2_f, is the pooled variance of the forecasts. This term is derived from the distribution of the probability survival forecasts assigned to survived states f_1 and forecasts assigned to non-survived states f_0 (Arkes et al. 1995, p. 121). Scatter represents the "noisiness" of survival probability forecasts and reflects the sensitivity of the model to information that is not related to the event occurrence. The smaller the scatter, the lower the Brier score. The scatter, or variance of forecasts, S^2_f, takes a minimum value of zero when the model produces constant forecasts.

The third term of the skill components, covariance S_{fd}, can be expressed as

$$S_{fd} = (\bar{f}_1 - \bar{f}_0)\left[\bar{d}(1-\bar{d})\right] \tag{72.15}$$

where
$\bar{f}_1$ is the mean survival forecasts assigned to rating states that actually survived
$\bar{f}_0$ is the mean survival forecasts assigned to non-survived states, i.e., states that experienced the migration event of interest.

Since the outcome index variance $\bar{d}(1-\bar{d})$ is determined by "natural forces," the covariance is determined by the term $(\bar{f}_1 - \bar{f}_0)$ (Yates 1982, p. 138). This term is equal to the slope of the regression line for a survival forecast on the outcome index. Given a base rate $\bar{d}$, the larger the term $(\bar{f}_1 - \bar{f}_0)$ or the steeper the slope, the lower the Brier score. The slope indicates the ability of the model to distinguish between the group of survived states f_1 and the group of non-survived states f_0. A steeper slope reflects the model's ability to assign higher probability survival forecasts to survived states than to non-survived states. It also shows the model's sensitivity to the information that is related to the event occurrence.

72.3.4 Variables

72.3.4.1 Measurements of Variables

Rating Scales

Corporate issuer ratings were obtained from Standard & Poor's CreditPro 2005 dataset. The data does not include rating outlooks and credit watch listings.

Standard & Poor's alphabetical rating scale includes 10 major rating categories varying from excellent credit quality (AAA) to default (D) as follows: AAA, AA, A, BBB, BB, B, CCC, CC, C, and D. A plus (+) or a minus (−) can be added to ratings from AA to CCC to capture the relative ranking within each rating class.[11] Ratings from AAA to BBB− are investment grade and ratings from BB+ to C are speculative grade.

This study employs a numerical rating scale to represent Standard & Poor's alphabetical rating scale. For the rating classes above AA+ and below CCC−, the plus (+) and minus (−) notches are not employed by Standard & Poor's. The omission of notches for AAA, CC, and C classes may lead to different gap lengths in rating scales. It is suggested that a one notch rating change in low rating grades implies a larger increase in default risk (Jorion and Zhang 2007). An issue arising here is the treatment of the rating gap between AAA and AA+ and the gap between ratings below CCC−. Similar to Dang (2010), this study assumed that the gap between AAA and AA+ is not one notch but two, implicitly including AAA− in the rating scale. A similar approach has been applied to rating classes below CCC−. As a result, the numeric rating scales were coded from 0 to 26 with 0 indicating the default class (D) and 26 indicating the AAA class. This numerical conversion maintains the rank order of the letter ratings, captures fine revisions intra-ratings, and reduces the effect of nonlinearity in the top and bottom rating classes. In addition, a dummy variable (*dummy junk boundary*) was also created to capture any nonlinearity between speculative grade rating boundary (BB−, BB, BB+) and adjacent rating grades.

Macroeconomic Time Series

Seven macroeconomic variables[12] were included in the static and dynamic hazard models. The *dummy recession* indicates the state of the business cycle. The Chicago Fed National Activity Index (*CFNAI*), the output growth gap (*RealGDPg actual minus potential*), and the *industrial production change* capture general economic activity. The *S&P500 quarterly return* and *S&P500 annualized standard deviation* represent the performance of the stock market, while the *term structure slope* reflects credit conditions.

As macroeconomic conditions are likely to affect rating migrations with a lag, this study uses an exponentially weighted average of lagged observations computed quarterly to construct five macroeconomic covariates (except *dummy recession* and *CFNAI*). The construction of macroeconomic lagged values is similar to the approach applied by Dang (2010) and Figlewski et al. (2012). The exponentially weighted average value X_t for the quarterly series x for a given macroeconomic variable in quarter t is calculated using data up to the previous quarter as

[11] http://www.standardandpoors.com/ratings/definitions-and-faqs/en/us/ (Accessed 17 August 2012)

[12] Seventeen macroeconomic candidate variables were considered and those that exhibited strong multicollinearity were eliminated.

$$X_t = \frac{\sum_{k=1}^{K} \delta^{k-1} x_{t-k}}{\sum_{k=1}^{K} \delta^{k-1}} \tag{72.16}$$

where $K = 6$ is the length of the lagged window and $\delta = 0.6815$ is the decay factor.

The *CFNAI* is published as a 3-month moving average and is used without further transformation.

Of seven macroeconomic covariates, one (*dummy recession*) is constructed as time-fixed in the static and dynamic hazard models, and six were constructed as either time-fixed in the static models or time-varying in the dynamic models. The value of the time-fixed covariate vector Z_j was measured at the beginning of each estimation and holdout rating state. The value of the time-varying macroeconomic covariate vector $Z_p^m(t)$ for estimation state m was taken quarterly over its rating duration, and the value used in Eq. 72.8 was updated to the most recent quarterly value as each risk set was formed. For holdout state q, the known information of the macroeconomic environment is limited to its commencement. Thus, the value of the macroeconomic covariate vector $Z_p^q(t)$ used to form survival forecasts for holdout state q was measured at state q's beginning and entered in Eqs. 72.6 and 72.12 without being subsequently updated. So, $Z_p^q(t) = Z_p^q(t = 0)$. The disadvantage is that the macro data for holdout observations become static in relation to time.

72.3.4.2 Definitions of Variables

The candidate variables that capture issuer-heterogeneity (rating history, industry sector) and time-heterogeneity (business cycle) in rating migration dynamics were identified from the literature.

Rating Variables

Start rating: The rating at the beginning of the current rating state. For a FA experiencing a plunge from an investment grade rating to a speculative grade rating at lag-one state, the *start rating* is also the rating on the fall date.

Dummy junk boundary: This dummy takes the value of one if the *start rating* falls within the speculative (junk) boundary BB−, BB, BB+, and zero otherwise.

Original rating: The rating of the firm when it was first rated.

Rate prior change: This indicates the average number of rating changes per year over the rating history of an issuer. It is calculated as the number of prior migrations (downgrades and upgrades) observed between the entry of the firm to the study and the beginning of the current rating state divided by the period from the time of entry till the start of the current rating state.

Rate prior down: This equals the average number of downgrades per year over the rating history of an issuer. It is calculated similar to *rate prior change* except that the numerator is the number of downgrades observed between the entry of the firm to the study and the beginning of the current state.

Lag one: The duration (in years) of the rating state that ends with a downgrade[13] and immediately precedes the current rating.

Lag two: The duration (in years) of the rating state that ends with either a downgrade or an upgrade (i.e., is not censored) and immediately precedes the lag-one rating.

Dummy lag2 down: This variable captures the direction of the lag-two regrade and takes the value of one if the lag-two rating ends with a downgrade and zero otherwise.

Number NR (*not rated*): This indicates the number of times a firm underwent a break in rating history (a rating withdrawal[14]) from the entry of the firm to the study until the beginning of the lag-one rating state.

Number prior fallen angel (FA): This indicates the number of times an issuer experienced a FA event (a downgrade from an investment grade rating to a speculative grade rating) from the entry of the firm to the study until the beginning of the lag-one rating state.[15]

Number rising star (RS): This indicates the number of times a firm experienced a RS event (an upgrade from a speculative grade rating to an investment grade rating) from the entry of the firm to the study until the start of the lag-one rating state.

Number big down: This indicates the number of times a firm experienced a substantial downgrade jump, defined as a jump of at least three rating notches,[16] from the entry of the firm to the study until the beginning of the current rating state.

Number big up: This variable indicates the number of times a firm experienced a substantial upgrade jump, defined as a jump of at least two rating notches, from the entry of the firm to the study until the beginning of the lag-one rating state.

Age since first rated: The rating age of the firm, which is equal to the length in years from the time the firm was first rated until the beginning of the current state.

---

[13]By definition, all FAs and FA peers in this study experienced a downgrade at lag-one rating state.

[14]Rating withdrawals bear negative credit implications if there is a lack of information to accurately assess debt issues (Carty 1997, p. 10). Issuers are likely to withdraw from being rated when they expect a downgrade. In this case, being unrated (censored) substitutes for being downgraded. The characteristics of issuers lost to non-independent (informative) censoring are often associated with the [migration] process under study (Blossfeld and Rohwer 1995; Kalbfleisch and Prentice 1980). There is no statistical test to check for and no standard methods for handling informative censoring (Allison 1995, p. 14). In this study, two sensitivity tests suggested by Allison (1995, pp. 249–252) to examine the effect of informative censoring on the main results have been applied to the proportional hazard models for FAs and their peers. It is found that being unrated is not informative.

[15]*Number prior FA* does not take into account the FA event FAs experienced at lag-one rating state. By definition, none of FA peers in this study experienced a FA event at lag-one rating state.

[16]Substantial rating changes of more than one letter grade (i.e., three rating notches) were more frequently observed in the ratings B through C (Lucas and Lonski 1992) and were less frequent than rating revisions of small magnitude (Carty and Fons 1994; Carty 1997). Downgrades involved a much bigger change in credit rating than upgrades (Jorion and Zhang 2007).

Control Variables

Industry dummies: Firms' industry sectors, as given by Standard & Poor's, were used to control for industry effects. The industry dummy took a value of one if the firm was in an industry sector and zero otherwise. Due to their unique business nature and credit risk exposure, firms in the *financial institution* sector[17] were excluded from the study. The data include 11 sectors, which resulted in ten dummy variables with the *insurance* sector left uncoded.

Industrial production change: As published by the US Federal Reserve Board.

RealGDPg actual minus potential (output growth gap): This variable measures the deviation of the actual real GDP growth (as published by the US Bureau of Economic Analysis) from the potential real GDP growth (as published by the St. Louis Federal Reserve).

CFNAI (Chicago Fed National Activity Index): This composite index is published by the Chicago Federal Reserve and is computed as a 3-month moving average of 85 monthly economic series.

Dummy recession: This variable takes the value of one if the rating state starts at the time of a recession, defined by the National Bureau of Economic Research, as 1 August 1990 to 31 March 1991 or 1 April 2001 to 30 November 2001, and zero otherwise.

S&P500 quarterly return: Definition sourced from Datastream.

S&P500 annualized standard deviation: Daily returns for the quarter are used to compute the standard deviation and this is expressed as an annual standard deviation.

Term structure slope: The slope is measured as the spread between 3-month and 10-year US Treasury Constant Maturity rates as published by the US Federal Reserve.

72.4 Data

72.4.1 Estimation and Holdout Periods

The rating behaviors of FAs and their peers in the USA were examined over the period 1984–2000.[18] This period covers different phases of the business cycle in the USA including the economic recession from July 1990 to March 1991.

[17] According to Lando and Skodeberg (2002), most financial institutions were assigned investment rating grades. As confidence- and capital-sensitive entities, it is difficult for financial institutions to run business with a poor credit profile or low credit rating. Lando and Skodeberg (2002) found that the duration dependence and the downward momentum are less pronounced for issuers in the financial institution sector than for issuers in other sectors. As this study examines the question of rating history dependence in the rating dynamics of speculative grade-rated issuers, financial institution sector was excluded from this study.

[18] The year 1984 was selected as the starting point for several reasons. 1982 is as far back as all macro data are available, and Standard & Poor's rating scales were changed in 1983. The growth of the US high-yield bond market and rating migrations from 1984 also constitute a significant source of events to this study.

The estimation period also captures major market downturns such as the stock market crash in 1987 and the collapse of the Long-Term Capital Market Hedge Fund in 1998. The period subsequent to the estimation period, 2001–2005, was used to construct holdout samples for FAs and their peers. The holdout period observed the 9/11 terrorist attack, the internet bubble burst, and the economic recession from March 2001 to November 2001. This period also witnessed altered business dynamics stemming from deregulations, the shakeouts that weeded out vulnerable FAs in high-velocity sectors, and the notorious collapse of tainted FAs such as WorldCom and Enron.

The descriptive statistics of the time series for the exponentially weighted average (except *CFNAI*) of macroeconomic variables are presented in Table 72.1. Additional analysis (not reported) indicates that macroeconomic time series in the estimation and holdout periods have statistically different mean and median values (except *term structure slope*).

72.4.2 Estimation and Holdout Samples

The estimation population includes 276 FAs and 1,102 FA-peer candidates. From the pool of 1,102 peer candidates, 276 observations were randomly chosen to form the FA-peer estimation sample. The holdout population includes 141 FAs and 937 FA-peer candidates. The FA-peer holdout sample with 141 observations was created randomly from the universe of 937 FA-peer candidates.[19]

The frequency of downgrades and upgrades for FA and FA peers in the estimation and the holdout samples are presented in Table 72.2. The migration propensity of estimation and holdout FAs differed markedly. Compared to estimation FAs, hold-out FAs are more vulnerable to downgrades and are less likely to regain investment grade status. This reinforces the notions that trends in corporate credit quality change over time (Carty and Fons 1994), with downgrades outnumbering upgrades (Altman and Kao 1991; Lucas and Lonski 1992; Lando and Skodeberg 2002).

72.4.3 Survival Time Distribution

The descriptive statistics of the survival time (survival duration) of FA/FA-peer down states and up states in the estimation and holdout periods are summarized in Panel A/Panel B of Table 72.3, respectively. Additional analysis (not reported) shows that FA and FA-peer down states (up states) in the estimation (holdout) period have statistically different mean survival time values.

[19]The FA sample and the universe of FA-peer candidates in the estimation/holdout period have markedly different distribution of issuers in the upper speculative rating classes (BB, BB+). Thus, it is impossible to construct from the candidate pool a FA-peer sample with the same current rating distribution and the same sample size as the FA sample for either the estimation or the holdout period.

Table 72.1 Descriptive statistics of macroeconomic variables

	Sample	Mean	Median	Standard deviation	Minimum	Maximum	Skewness	Kurtosis
Chicago Fed National Activity Index (CFNAI)	1984–2000	0.113	0.155	0.447	−1.287	1.156	−0.677	1.711
	2001–2005	−0.301	−0.353	0.466	−1.109	0.352	−0.197	−1.087
RealGDPg actual minus potential (%)	1984–2000	0.131	0.153	0.348	−0.927	1.126	0.045	2.623
	2001–2005	−0.16	−0.194	0.352	−0.795	0.32	−0.258	−1.286
Industrial production change (%)	1984–2000	0.939	1.011	0.662	−0.853	2.521	−0.108	0.071
	2001–2005	0.183	0.42	0.655	−1.293	1.085	−1.02	0.138
SP500 quarterly return (%)	1984–2000	3.609	3.294	2.99	−3.869	9.484	−0.077	−0.458
	2001–2005	−0.501	0.157	4.33	−9.267	6.385	−0.344	−0.71
SP500 annual standard deviation	1984–2000	1.769	1.678	0.612	0.649	3.789	1.014	1.303
	2001–2005	2.284	2.467	0.68	1.27	3.233	−0.279	−1.435
Term structure slope (%)	1984–2000	1.716	1.718	0.943	0.165	3.363	−0.018	−1.299
	2001–2005	2.14	2.621	1.029	−0.128	3.156	−1.2	0.248

Table 72.2 Downgrade and upgrade frequency

	FAs			FA peers		
	Total	Down states	Up states	Total	Down states	Up states
Estimation sample, 1984–2000	276	102 (37 %)	104 (37.7 %)	276	163 (59 %)	46 (16.7 %)
Holdout sample, 2001–2005	141	83 (58.9 %)	13 (9.2 %)	141	71 (50.4 %)	22 (15.6 %)

The survival duration histograms of FA/FA-peer down states and up states in the estimation sample are depicted in Fig. 72.1. The time-to-event distributions are quite different between down states and up states. For down states, particularly FA peers, the histograms suggest that further decline in survival duration tends to be swift. There is a noticeable concentration of FA/FA-peer down states in durations within 1 year, whereas the durations for FA/FA-peer up states tend to be longer.

As shown in Table 72.3 and Fig. 72.1, estimation FA peers are riskier than estimation FAs as evidenced by a higher (lower) likelihood to downgrade (upgrade) and a shorter median time to downgrade. The reverse applies in the holdout period. Holdout FAs are more vulnerable to downgrades and experience a shorter (longer) median time to downgrade (upgrade) than holdout FA peers (Table 72.3). This is consistent with the notion that FAs display faster downward migration and greater rating velocity than their peers within few years since the fall date.

72.5 Estimation Results

The Cox's proportional hazard models and the Cox's dynamic hazard models were estimated for FAs/FA peers over the period 1984–2000. The results of the estimated hazard models are given in Panel A – Table 72.4. Panel B – Table 72.4 – is appended to Panel A and provides statistics on the fit of the models. The backward stepwise estimation procedure was employed. Significant variables were retained in the models according to the log-likelihood ratio test, at the 10 % level or better, derived from the maximum likelihood procedure used to estimate the models. In interpreting Panel A of Table 72.4, a negative coefficient reduces the hazard of the migration event being modelled. The reported hazard ratios represent the relative change in the hazard for a one-unit change in the independent variable. The discussion that follows focuses on the effect of past rating behaviors on the downgrade and upgrade hazard of FAs and their peers.

72.5.1 Cox's Proportional Hazard Models

72.5.1.1 Control Variables

With respect to macroeconomic variables and industry sector dummies, several key results are observed (Panel A of Table 72.4). FAs, particularly with respect to upgrades, are sensitive to the economic environment prevailing on the fall date.

Table 72.3 Descriptive statistics of survival time

Samples	Rating states	No. of obs (sample size)	Mean (years)	Median (years)	Standard deviation	Minimum (days)	Maximum (years)	Skewness	Kurtosis
Panel A: Descriptive statistics of FAs' survival time									
1984–2000	Down states	102 (276)	1.12	0.83	1.07	2	5.08	1.44	1.9
	Up states	104 (276)	2.19	2	1.62	31	10.44	1.75	5.74
2001–2005	Down states	83 (141)	0.46	0.24	0.73	1	4.03	2.92	9.15
	Up states	13 (141)	1.92	2.18	1.12	29	3.1	−0.58	−1.23
Panel B: Descriptive statistics of FA peers' survival time									
Samples	Rating states	No. of obs (sample size)	Mean (years)	Median (years)	Standard deviation	Minimum (days)	Maximum (years)	Skewness	Kurtosis
1984–2000	Down states	163 (276)	0.62	0.26	1.05	1	7.64	4.03	20.36
	Up states	46 (276)	1.95	1.71	1.29	18	6.21	1.09	1.47
2001–2005	Down states	71 (141)	0.53	0.29	0.57	1	2.36	1.66	2.43
	Up states	22 (141)	1.02	0.64	1.07	1	4.69	1.97	5.57

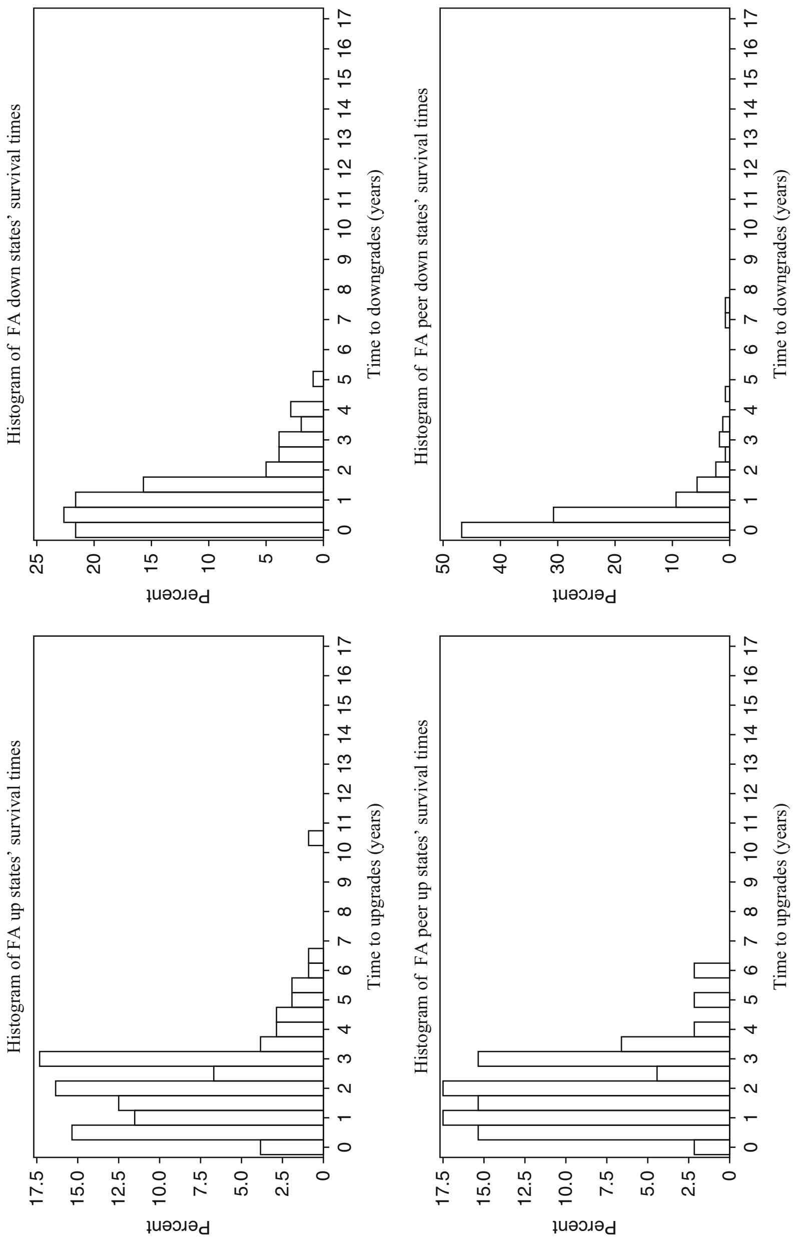


Fig. 72.1 Survival time histogram, 1984–2000

Table 72.4 Cox's hazard models, 1984–2000

Panel A: Coefficient estimates (backward stepwise selection)																
	Cox's proportional hazard models								Dynamic Cox's hazard models (with time-varying covariates)							
	FA downgrade		FA-peer downgrade		FA upgrade		FA-peer upgrade		FA downgrade		FA-peer downgrade		FA upgrade		FA-peer upgrade	
Variables	Parameter estimate	Hazard ratio	Parameter estimate	Hazard ratio	Parameter estimate	Hazard ratio	Parameter estimate	Hazard ratio	Parameter estimate	Hazard ratio	Parameter estimate	Hazard ratio	Parameter estimate	Hazard ratio	Parameter estimate	Hazard ratio
Rating																
Start_rating			−0.20504***	0.815	−0.37299***	0.689	−0.36534***	0.694			−0.21228***	0.809	−0.23684*	0.789	−0.37994***	0.684
Dummy_junk_boundary					1.77343***	5.891	1.15321***	3.168					1.34784**	3.849	1.18741***	3.279
Age_since_first_rated			−0.07108***	0.931							−0.05684**	0.945	−0.06323**	0.939		
Original_rating													0.08709**	1.091		
Lag_one	−0.16415*	0.849							−0.19625**	0.822						
Lag_two																
Dummy_lag2_down			1.30892***	3.702							1.24809***	3.484				
Rate_prior_change			2.31989***	10.175			−0.29792**	0.742			2.11083**	8.255				
Rate_prior_down			−2.51459***	0.081							−2.30957**	0.099			−0.32878**	0.72
Number_not rated					0.84484**	2.328							1.04557**	2.845		
Number_prior_fallen angel																
Number_rising star	−0.43198	0.649														
Number_big_down																
Number_big_up	0.90975***	2.484							0.55624**	1.744						
Macroeconomic																
Dummy recession									1.01336*	2.755						
CFNAI					2.19648***	8.993			−2.09952***	0.123	−1.33964***	0.262				
RealGDPg_actual_minus_potential					−1.58492***	0.205			1.48577**	4.418						
Industrial_production_change	0.27751	1.32			−1.06589***	0.344			0.7123**	2.039	0.62888**	1.876				
SP500_quarterly_return	−0.07292**	0.93									0.08373***	1.087				
SP500_annual_standard deviation											0.42558**	1.53				
Term_structure_slope	−0.64038***	0.527	−0.37832***	0.685					−0.49195***	0.611	−0.25417**	0.776				
Industry sector																
Aerospace/automotive/capital goods/metal					−0.69147**	0.501										
Consumer/service sector			0.33691	1.401												
Energy and natural resources																
Leisure time/media	−1.95492**	0.142							−1.96074**	0.141						
Real estate																
Utility			0.85778*	2.358												

(*continued*)

Table 72.4 (continued)

Panel A: Coefficient estimates (backward stepwise selection)																
	Cox's proportional hazard models								Dynamic Cox's hazard models (with time-varying covariates)							
	FA downgrade		FA-peer downgrade		FA upgrade		FA-peer upgrade		FA downgrade		FA-peer downgrade		FA upgrade		FA-peer upgrade	
Variables	Parameter estimate	Hazard ratio	Parameter estimate	Hazard ratio	Parameter estimate	Hazard ratio	Parameter estimate	Hazard ratio	Parameter estimate	Hazard ratio	Parameter estimate	Hazard ratio	Parameter estimate	Hazard ratio	Parameter estimate	Hazard ratio
Forest and building products/homebuilders							1.54429***	4.685							1.5883***	4.895
Health care/chemicals	−1.33992**	0.262			−0.87047	0.419			−1.15371**	0.315						
Transportation							1.54248***	4.676							1.62093***	5.058
High technology/computers/office equipment							1.65675***	5.242							1.65733***	5.245

Panel B: Model fit statistics								
	Cox's proportional hazard models				Dynamic Cox's hazard models (with time-varying covariates)			
	FA downgrade	FA-peer downgrade	FA upgrade	FA-peer upgrade	FA downgrade	FA-peer downgrade	FA upgrade	FA-peer upgrade
-2 log likelihood (without covariates)	958.76	1464.18	853.75	347.76	958.76	1464.18	853.75	347.76
-2 log likelihood (with covariates)	898.37	1369.6	817.35	324.92	898.23	1357.2	835.45	324.23
Likelihood ratio	60.39	94.58	36.4	22.84	60.53	106.98	18.33	23.53
Degree of freedom	8	8	8	6	9	10	5	6
Pr > ChiSq	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

***p-value ≤ 1 %

**1 % $< p$-value ≤ 5 %

*5 % $< p$-value ≤ 10 % based on Wald chi-square tests

The partial likelihood ratio test is calculated as twice the difference between the log partial likelihood of the model containing the covariate and the log partial likelihood for the model not containing the covariate. As shown in Panel B, the likelihood ratio statistics are significant at better than 1 % level, rejecting the global null hypothesis that all coefficients are equal to zero

For example, changes in *CFNAI* have no impact on the downward journey of FAs, whereas a one-unit increase in *CFNAI* makes a return to the investment grade universe 799.3 % more likely. On the other hand, FA peers are more vulnerable to industry risk. Those in the *utility s* ector are 135.8 % more vulnerable to downgrades. This is in contrast to the notion that issuers in *utility* sector have a stable rating process (Kadam and Lenk 2008). It is noticeable that macroeconomic conditions are not significant, whereas industry sectors are the key determinants of the upgrade probability of FA peers. Those in *high technology/computers/office equipment*, *forest and building products/homebuilders*, or *transportation* sectors are hundreds of percentage points more likely to become rising stars.

72.5.1.2 Rating Behaviors

FAs and their peers exhibit different dependence on the current rating and past rating behaviors (Panel A of Table 72.4). The *start rating* – the rating a FA received on the fall date – and *dummy lag2 down* are not significant in determining the downgrade hazard of FAs. The absence of *start rating* in the downgrade model is in direct contrast to the Markov property and is inconsistent with the "destination pattern" – the lower a rating a FA descends to, the more likely it will default (Mann et al. 2003, p. 5). On the other hand, the absence of *dummy lag2 down* reinforces the notions that the strong effect of a previous rating change become weaker with the passage of time (Hamilton and Cantor 2004, p. 10) and does not persist after 2 or 3 years (Fledelius et al. 2004).

Two rating history variables (*number big up* and *lag one*) are significant in the downgrade model for FAs. The effect of *number big up* seems to contradict credit rating agencies' policy to "limit rating reversal and dampen rating volatility" (Hamilton and Cantor 2004, p. 3). A substantial jump to higher rating grades (*number big up*) makes a rating bounce 148.4 % more likely. In contrast to *number big up*, *lag one* has a smaller impact. Extending the duration of *lag-one* rating state by 1 year merely increases the probability that the current rating persists by 15 %. The effect of *lag one* is consistent with the duration dependence phenomenon suggested by Lando and Skodeberg (2002).

The downgrade process of FA peers exhibit substantial dependence on past rating behaviors. Those experiencing a downgrade at lag-two rating state (*dummy lag2 down*) are 270 % more likely to descend to lower rating classes. The strong impact of *dummy lag2 down* is consistent with the downward momentum in corporate rating dynamics.[20] Prior rating volatility (*rate prior change*) and prior downgrade volatility (*rate prior down*) have reverse coefficient signs, of which the earlier dominates with a 10 times stronger effect. Increasing prior rating volatility

[20]The effect of a previous rating change decays as time passes (Hamilton and Cantor 2004, p. 10). Thus, the shorter the *lag-one* rating state, the more influential the rating change at lag-two state (*dummy lag2 down*). Additional analysis (not reported) indicates that FA peer down states have a shorter *lag one* than FAs down states. Consequently, the effect of *dummy lag2 down* on the probability of a subsequent downgrade persists on FA peers but does not hold on FAs (as discussed earlier).

(*rate prior change*) by one migration per year makes a further downgrade for FA peers 917.5 % more likely. In comparison to *dummy lag2 down* and *rate prior change*, the current raring (*start rating*) and rating age (*age since first rated*) have a modest effect. The better the current rating (*start rating*) and the longer the time since an issuer was first rated (*age since first rated*), the more likely a FA peer will retain its current rating. The effect of *age since first rated* contrasts with the notion that aging issuers are vulnerable to downgrades and default (Altman and Kao 1991; Altman 1992, 1998; Figlewski et al. 2012).

The upgrade models for FAs and FA peers share two common variables; both capture the current rating state (*start rating* and *dummy junk boundary*). FAs and their peers with a better *start rating* are more likely to stay in their current rating grade. Those in the speculative grade barrier, BB+, BB, or BB− (*dummy junk boundary*), have a favorable tendency towards upgrades and are several times more likely to be rising stars. The effect of *dummy junk boundary* is more pronounced than the effect of the *start rating* and is stronger for FAs than FA peers. The effect of *dummy junk boundary* is consistent with the "destination pattern" – a higher rating assigned to a FA upon its entrance to the speculative rating spectrum makes a return to the investment rating universe more likely. This is most apparent in the speculative grade boundary (Mann et al. 2003, p. 5).

Only one rating history variable is significant in the upgrade models for FAs and FA peers. FAs with a prior rating withdrawal (*number not rated*) are 133 % more likely to regain the investment grade status. In contrast, FA peers with a frequent migration history (*rate prior change*) are 25.8 % less likely to ascend to higher rating grades.

Overall, in a static estimation framework, several aspects of rating history are key determinants of the migration hazard of FAs/FA peers. The question is whether the impact of rating history persists in a dynamic estimation framework?

72.5.2 Dynamic Cox's Hazard Models

Some noticeable changes were observed when switching from a static to a dynamic estimation framework. The following discussion focuses on the distinguishing features between the respective proportional models and the dynamic models for FAs and their peers (Panel A of Table 72.4).

72.5.2.1 Control Variables

In a dynamic estimation framework, the downgrade process of FAs and their peers exhibit strong dependence on macroeconomic conditions, whereas their upgrades are entirely driven by the current rating, past rating behaviors, and industry sectors. This reinforces the concept that upgrades are more sensitive to firm-specific risk factors than macroeconomic shocks.

FAs, with respect to downgrades, are very sensitive to the economic environment prevailing at each migration time. During an economic recession (*dummy recession*) or a period of large output growth gap (*RealGDPg_actual_minus_potential*) and large *industrial production change*, FAs exhibit rapid deterioration and are hundreds of percentage points more vulnerable to downgrades.

72.5.2.2 Rating Behaviors

For downgrades, relative to the corresponding static model, the dynamic model for FAs/FA peers includes the same set of significant rating variables; all of them retain the same coefficient signs. The substantial effect of past rating volatility (*rate prior change*) on FA peers diminishes though it still dominates other rating behaviors. In both estimation frameworks, the current rating (*start rating*) is either not significant or has a minimal impact on the downgrade process of FAs and their peers.

For upgrades, the results of the proportional and dynamic models for FA peers are consistent, with one exception. *Rate prior down* replaces *rate prior change* and has a similar coefficient. FA peers with a frequent downgrade history (*rate prior down*) are less likely to become rising stars. In both estimation frameworks, the upgrade process of FA peers is entirely driven by endogenous risk factors, of which industry sectors are more influential than rating history and the current rating.

Noticeable changes were observed in the upgrade model for FAs. None of macroeconomic factors/industry sectors affects FAs' probabilities to regain the investment grade status. In other words, the upgrade process of FAs is entirely determined by the current rating and rating history, of which rating history dominates the current rating. Relative to the corresponding static model, the dynamic model for FAs features a stronger effect and a greater number of rating history variables. Variables significant in the proportional model (*start rating*, *dummy junk boundary*, *number not rated*) are also present in the dynamic model and retain the same coefficient signs. The current rating, specifically *dummy junk boundary*, is less influential; its hazard ratio is about 200 basis points smaller when switching to a dynamic estimation procedure. *Age since first rated* and the *original rating* become significant. Aging FAs (*age since first rated*) are less likely to make an uphill climb, whereas originally high-rated FAs (*original rating*) exhibit a favorable tendency towards upgrades.

In summary, past rating behaviors persist even after controlling for the industry sector and the development of macroeconomic conditions over the time for which the current rating persists. As suggested by Lando (2004, p. 97), the Markov chain assumptions do not hold if ratings vary across the business cycle or depend on the age of the bond. The results of this study therefore suggest that rating changes for FAs and FA peers are non-Markovian, and rating history variables can be used in estimating future rating changes. The question is how accurate are such forecasts?

72.6 Forecast Accuracy Assessment

The Brier scores (Brier 1950) of survival probability forecasts generated by the static and dynamic hazard models for FAs/FA peers were calculated as in Eq. 72.13

and decomposed into forecast components as in Eq. 72.14. A Brier score of zero indicates perfect predictive ability and a Brier score of one indicates no predictive ability. Variation in the Brier score through time is a consequence of both the passage of time and the changing holdout sample composition.[21]

Table 72.5 summarizes the covariance decompositions of the Brier scores at 1- and 2-year forecast horizons. Removing the outcome index variance $\overline{d}_t\left(1-\overline{d}_t\right)$ from the overall Brier score B_t results in the skill component score, which "levels the playing field" and enhances the validity of comparative forecast assessments (Yates 1982). The following discussion focuses on the skill components (calibration or bias, variance or scatter, and discrimination or slope) of the Brier score. In the interests of brevity, the covariance graphs which illustrate the covariance components of the Brier scores are not presented. However, an example of a covariance graph can be found in Appendix 2.

72.6.1 The Predictive Accuracy of Cox's Proportional Hazard Models

72.6.1.1 Upgrade Models for FAs/FA Peers

The proportional (static) upgrade models for FAs and FAs peers exhibit modest forecast performance as evidenced by relatively small skill component scores. A lower skill component score implies better forecast performance.

The upgrade model for FAs demonstrates excessive pessimism and underestimates survival estimates by a large margin. Additional analysis of the covariance graphs indicates that the model generally places 1- and 2-year survival forecasts in the middle and pessimistic probability categories that are substantially beneath the survival base rate ($\overline{d}_1 = 90.78\ \%$ and $\overline{d}_2 = 75\ \%$). This corresponds to a large bias of $-21.98\ \%$ (Appendix 2) and $-19.85\ \%$ at 1- and 2-year horizons, respectively.

In contrasts, the model for FA peers demonstrates well-calibrated survival estimates across forecast horizons. Additional analysis of the covariance graphs (not reported) indicates that the model for FA peers assigns 1- and 2-year survival estimates in the optimistic probability categories that are close to the mean survival index ($\overline{d}_1 = 84.4\ \%$ and $\overline{d}_2 = 72.2\ \%$). This tendency translates into a negligible bias at both forecast horizons.

Both the upgrade models for FAs and their peers do well in removing irrelevant information and achieve small variability (scatter) in estimates. However, survival forecasts for survived and upgraded issuers concentrate in few probability deciles, reflecting poor discrimination ability.

[21] In forming the survival forecasts for holdout FAs/FA peers, the approach of Chen et al. (2005) is followed. As the time horizon unfolds, Chen et al. (2005) deleted from the holdout sample at time t those cases which are censored, or have experienced the event, before time t. The approach of Chen et al. (2005) results in a holdout sample that reduces with the passage of time. The number of survival forecasts N_t Eq. 72.13 accordingly reduces as the forecast time t gets longer.

Table 72.5 Brier score of time-varying survival forecasts for FAs and their peers in the holdout period 2001–2005

	Cox's proportional hazard models				Dynamic Cox's hazard models			
	FA downgrade model		FA upgrade model		FA downgrade model		FA upgrade model	
Forecast horizon t (year)	1	2	1	2	1	2	1	2
Number of forecast, N_t	141	44	141	44	141	44	141	44
Mean survival index (survival base rate), $\overline{d}_t$	0.4113	0.7727	0.9078	0.75	0.4113	0.7727	0.9078	0.75
Mean survival forecast, $\overline{f}_t$	0.8029	0.7059	0.688	0.5515	0.8577	0.762	0.9521	0.867
(1) **Brier score B_t = (2) + (3) + (4) − (5)**	**0.4021**	**0.1927**	**0.1578**	**0.2299**	**0.4388**	**0.1863**	**0.0829**	**0.1897**
(2) Outcome index variance, $\overline{d}_t(1-\overline{d}_t)$	0.2421	0.1756	0.0837	0.1875	0.2421	0.1756	0.0837	0.1875
Skill component score=(1) − (2) = (3) + (4) − (5)	**0.16**	**0.0171**	**0.0741**	**0.0424**	**0.1967**	**0.0107**	**−0.0008**	**0.0022**
(3) Forecast variance (scatter), $S^2_{f_t}$	0.0072	0.0171	0.0372	0.026	0.0051	0.0073	0.001	0.0117
(4) Bias square, $(\overline{f}_t-\overline{d}_t)^2$	0.1533	0.0045	0.0483	0.0394	0.1992	0.0001	0.002	0.0137
(5) 2 forecast-outcome-covariance, or 2 $S_{f_t d_t}$	0.0005	0.0045	0.0113	0.023	0.0076	−0.0032	0.0038	0.0232
Slope $(\overline{f}_1-\overline{f}_0)=[(5)/(2)]/2$	0.0010	0.0128	0.0675	0.0613	0.0157	−0.0091	0.0227	0.0619
	FA downgrade model		**FA upgrade model**		**FA downgrade model**		**FA upgrade model**	
Forecast horizon t (year)	1	2	1	2	1	2	1	2
Number of forecast, N_t	141	36	141	36	141	36	141	36
Mean survival index (survival base rate), $\overline{d}_t$	0.4965	0.7222	0.844	0.7222	0.4965	0.7222	0.844	0.7222
Mean survival forecast, $\overline{f}_t$	0.5846	0.5057	0.8565	0.7046	0.7311	0.6604	0.8469	0.6864
(1) **Brier score B_t = (2) + (3) + (4) − (5)**	**0.2834**	**0.285**	**0.1421**	**0.2111**	**0.3119**	**0.2345**	**0.1461**	**0.1991**
(2) Outcome index variance, $\overline{d}_t(1-\overline{d}_t)$	0.25	0.2006	0.1317	0.2006	0.25	0.2006	0.1317	0.2006
Skill component score = (1) − (2) = (3) + (4) − (5)	**0.0334**	**0.0844**	**0.0104**	**0.0105**	**0.0619**	**0.0339**	**0.0144**	**−0.0015**
(3) Forecast variance (scatter), $S^2_{f_t}$	0.0287	0.0213	0.0128	0.033	0.0284	0.0288	0.0181	0.0404
(4) Bias square, $(\overline{f}_t-\overline{d}_t)^2$	0.0078	0.0469	0.0002	0.0003	0.055	0.0038	0	0.0013
(5) 2 forecast-outcome-covariance, or 2 $S_{f_t d_t}$	0.0031	−0.0162	0.0025	0.0228	0.0216	−0.0013	0.0037	0.0431
Slope $(\overline{f}_1-\overline{f}_0)=[(5)/(2)]/2$	0.0062	−0.0404	0.0095	0.0568	0.0432	−0.0032	0.0140	0.1074

72.6.1.2 Downgrade Models for FAs/FA Peers

The proportional downgrade models for FAs and FA peers perform poorly. Additional analysis of the covariance graphs (not reported) shows that most of the 1-year estimates for FAs were placed in the probability categories that are substantially higher than the survival base rate $\left(\overline{d}_1 = 41.13\ \%\right)$. The marked tendency to endorse upper optimistic categories in 1-year survival forecasts is a major cause of the large positive bias of 39.15 %. For FA peers, a majority of 1-year survival estimates were assigned to the middle probability deciles that are slightly above the mean survival index $\left(\overline{d}_1 = 49.65\ \%\right)$, resulting in a relatively small bias of 8.8 %. Both downgrade models for FAs/FA peers assign 2-year survival forecasts to the probability categories that are lower than the survival base rate ($\overline{d}_2 = 77.27\ \%$ for FAs and $\overline{d}_2 = 72.22\ \%$ for FA peers), resulting in a negative bias.

On the index of slope, the downgrade models for FAs/FA peers perform poorly in discriminating survived states from downgraded states. Both models, however, achieve negligible scatters (variance).

72.6.1.3 Sources of Forecast Errors

For both FAs and their peers, the static upgrade models exhibit some predictive ability, whereas the static downgrade models perform poorly. The following discussion focuses on two factors that may contribute to the forecast performance of the static models.

The macroeconomic conditions in the estimation and holdout periods markedly differ (Table 72.1). Similarly, the migration propensity of FAs in the estimation period is not representative of the migration pattern in the holdout period. Relative to estimation FAs, holdout FAs are more vulnerable to downgrades and exhibit an unfavorable tendency towards upgrades (Tables 72.2 and 72.3). The poor economic development and the rapid deterioration in the credit quality of FAs in the holdout period present challenges to both downgrade and upgrade models.

The proportional models employ a static estimation framework and include time-fixed covariates which capture the economic conditions at the beginning of each rating state. The information embedded in the time-fixed macroeconomic covariates becomes increasingly stale and less relevant as the rating duration unfolds. As suggested by Amato and Furfine (2004), the date of a rating change is generally close to the time the actual credit review takes place. Any decision by credit rating agencies is influenced by the economic conditions prevailing at the time of the rating change. Downgrades departing from low ratings are particularly sensitive to the prevailing economic environment. This suggests the need to develop dynamic hazard models that include time-varying macroeconomic covariates updated at each event time.

72.6.2 The Predictive Accuracy of Dynamic Cox's Hazard Models

The following discussion focuses on the comparative performance of the dynamic models for FAs/FA peers in comparison to the respective static models examined above (Table 72.5).

72.6.2.1 Upgrade Models for FAs/FA Peers

The dynamic upgrade model for FA peers exhibits similar good forecast performance at the 1-year horizon and outperforms the respective static model at the 2-year horizon. Both the static and dynamic models for FA peers show strong calibration ability as evidenced by a minimal bias across forecast horizons. Of particular interest, the dynamic model achieves a zero (perfect) bias at the 1-year horizon. On the index of slope, the dynamic model performs slightly better in discriminating rating states that survived from states that were upgraded at the 2-year forecast horizon. This corresponds to a 5 % steeper slope and a smaller Brier score.

The dynamic upgrade model for FAs performs far better than the corresponding static model and achieves a competitive skill component score at both 1- and 2-year forecast horizons. The predictive accuracy is particularly good at the 1-year horizon. Relative to the corresponding static model, the dynamic upgrade model for FAs uses the optimistic probability category above 80 % to a much greater extent and employs middle and lower probability deciles to a lesser extent. Unlike the respective proportional model, the dynamic model seldom makes pessimistic survival forecasts. The marked propensity to place a majority of 1- and 2-year survival estimates into the optimistic categories that are close to the survival base rate ($\overline{d}_1 = 90.78\ \%$ and $\overline{d}_2 = 75\ \%$) translates into a small (superior) positive bias.

72.6.2.2 Dynamic Downgrade Models for FAs and FA Peers

The forecast performance of the dynamic downgrade models for FAs and FA peers at 1-year horizon is disappointing. This can be attributed to the deterioration in their calibration ability. Relative to the respective static models, the dynamic downgrade models for FAs/FA peers use optimistic probability categories to a greater extent and employ middle and lower probability deciles to a lesser extent. As a result, the dynamic downgrade models overestimate the 1-year survival forecasts by a larger margin and exhibit a larger (inferior) positive bias.

However, at the 2-year forecast horizon, both dynamic downgrade models have an edge over the respective static models. The dynamic models underestimate the 2-year survival forecasts by a smaller margin and obtain a smaller (more competitive) negative bias. It is noticeable that the dynamic model for FAs offers well-calibrated 2-year survival estimates.

72.6.2.3 Limitations

This study suffers from some limitations which contribute to the poor forecast performance of the downgrade models for FAs/FA peers at the 1-year horizon.

Speculative grade issuers, particularly FAs, are naturally under the close scrutiny of market participants and are potential targets for negative rating reviews and downward revisions following their fall dates. The common concept is that credit rating agencies focus more resources on quantifying deteriorations in the credit profile of speculative grade issuers than analyzing any improvement in their earnings (Holthausen and Leftwich 1986). Downgrades tend to be quick, whereas upgrades tend to lag behind credit quality improvement. Furthermore, ratings are assigned in a pro-cyclical manner.[22] Amato and Furfine (2004) suggested that credit rating agencies exhibit a propensity to overreact to the prevailing macroeconomic conditions when they revise ratings. They tend to exhibit excessive pessimism during economic downturns. This tendency could contribute to an acceleration of credit deterioration in the volatile holdout period 2001–2005. The survival forecasts based on the "average" migration experience in the estimation period 1984–2000 fail to capture the rapid deterioration in the credit quality of holdout FAs.

Furthermore, the employment of static macroeconomic data to form forecasts for holdout issuers in Eq. 72.12 is likely to dampen the predictive accuracy of the dynamic downgrade models. This is more pronounced given the strong effect of macroeconomic variables on downgrades (Panel A of Table 72.4) and the dramatically changing economic conditions in the holdout period (Table 72.1).

The absence of rating outlook and credit watch data in this study is also likely to diminish the predictive power of the downgrade models. It is suggested that rating outlook and CreditWatch data diminish the impact of rating history (Hamilton and Cantor 2004) and exhibit predictive accuracy in forecasting future rating changes (Vazza et al. 2005b; Hill et al. 2010; Guttler 2011).

72.7 Conclusion

Using Standard & Poor's CreditPro 2005 dataset and the survival analysis framework (Allison 1984), this study aimed to address two issues: first, to examine issuer-heterogeneity and time-heterogeneity in the rating migration dynamics of FAs and their speculative grade-rated peers over the period 1984–2000 and, second, to assess the predictive accuracy of the static and dynamic hazard models in

[22]The pro-cyclicality in rating actions may be attributed to the possibility that business cycle fluctuations coincide with permanent changes in credit quality (Loffler 2012).

forecasting future rating changes over the subsequent period, 2001–2005. The principal conclusions emerging from this study are as follows:

Past rating behaviors, industry sectors, and macroeconomic conditions are the key determinants of the rating migration process. The significant factors differ between FAs and their peers and vary between downgrades and upgrades. A key finding is that past rating behaviors persist and have a strong effect in the presence of time-varying macroeconomic covariates. As seen in the dynamic model, a prior rating withdrawal makes a FA 185 % more likely to return to the investment grade universe. A one-migration increase in the annual regrade volatility and a downgrade at lag-two rating state, respectively, makes a FA peer 725 % and 248 % more likely to continue the downward journey.

In both static and dynamic estimation frameworks, FAs and FA peers close to the speculative grade boundary (BB+, BB, BB−) are resilient and several times more likely to become rising stars. The current rating – the rating a FA received on the fall date – is not significant in determining its subsequent downgrade probability. For other rating migration processes, the current rating has a modest impact compared to past rating behaviors.

Volatile macroeconomic conditions and accelerated credit deterioration in the holdout period bring challenges to the models' forecast performance. In the aggregate, all models do well in removing irrelevant information but at some expense of failing to incorporate important information. The tendency to assign survival estimates in the probability categories that are either beyond or below the survival base rate results in some bias.

Switching from a static to a dynamic estimation framework improves the calibration power of the downgrade and upgrade models for FAs/FA peers, mostly at the 2-year forecast horizon. Of particular interest, the dynamic upgrade model for FAs, which includes only rating variables and features a strong effect of rating history, substantially outperforms the respective static model at both the 1- and 2-year forecast horizons. The results are in direct contrast to the evidence that the Markov property adequately holds within 1 or 2 years (Kiefer and Larson 2007; Frydman and Schuermann 2008).

The implication is that rating history provides important diagnostic information in making well-calibrated estimates for FAs. Financial institutions and regulators should condition models of rating migrations by reference to the path an issuer has followed to the current rating state. If, for example, an issuer is a FA, then it may need a FA dynamic hazard model to obtain appropriate time-varying survival probability estimates.

A natural extension of this study would be to employ rating outlook or credit watch data as a time-varying covariate to capture the changing credit quality of issuers over rating durations. The study could also be extended by using a moving window and continually updating/recalibrating the models to capture the acceleration in credit deterioration. Future work could also be directed to estimate separate dynamic hazard models over economic contraction and expansion periods and to examine credit quality changes under favorable and unfavorable macro environment.

Appendix 1: Maximum Partial Likelihood Estimation

The expression in Eq. 72.1 for the Cox's proportional hazard model and the expression in Eq. 72.3 for the full partial likelihood are repeated here as Eqs. 72.17 and 72.18 for convenience:

$$h_m[t, Z] = h(0, t)\exp[Z^m \beta] \tag{72.17}$$

$$PL = \prod_{m=1}^{N} L_{t_m}^m = \prod_{m=1}^{N} \left[\frac{\exp(\beta \ Z^m)}{\sum_{i \in R(t_m)} \exp\left(\beta \ Z^i\right)} \right] \tag{72.18}$$

The log partial likelihood function can be written as

$$PL = \sum_{m=1}^{N} \left[(\beta \ Z^m) - \ln \left[\sum_{i \in R(t_m)} \exp\left(\beta \ Z^i\right) \right] \right] \tag{72.19}$$

The derivative of Eq. (72.19) with respect to β is

$$\frac{\partial PL}{\partial \beta} = \sum_{m=1}^{N} \left[Z^m - \frac{\sum_{i \in R(t_m)} Z^i \exp\left(\beta \ Z^i\right)}{\sum_{i \in R(t_m)} \exp\left(\beta \ Z^i\right)} \right] \tag{72.20}$$

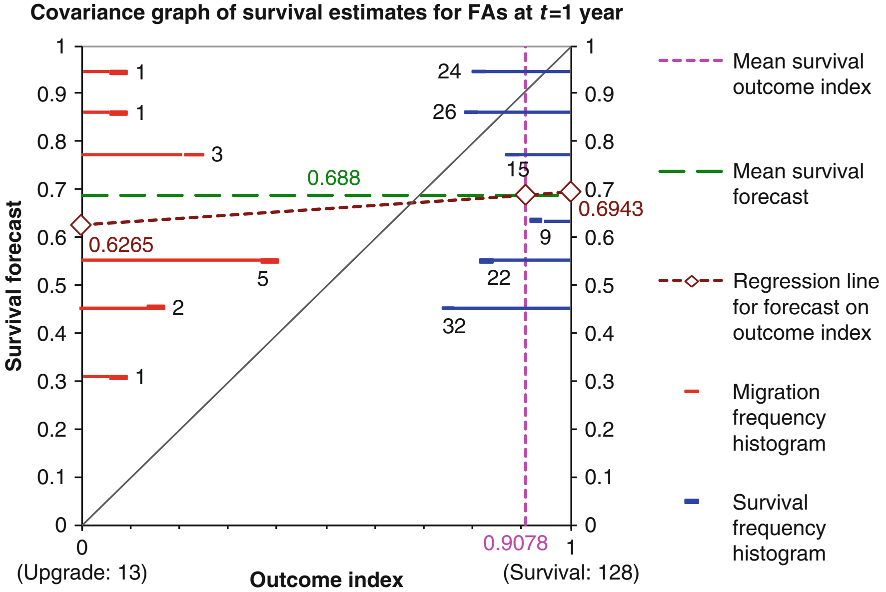
$$\frac{\partial PL}{\partial \beta} = \sum_{m=1}^{N} \left[Z^m - \sum_{i \in R(t_m)} w_{im}(\beta) Z^i \right]$$

$$\frac{\partial PL}{\partial \beta} = \sum_{m=1}^{N} \left[Z^m - \overline{Z}^{w_{im}} \right]$$

where $w_{im}(\beta) = \frac{\exp\left(\beta \ Z^i\right)}{\sum_{i \in R(t_m)} \exp\left(\beta \ Z^i\right)}$ and $\overline{Z}^{w_{im}} = \sum_{i \in R(t_m)} w_{im}(\beta) Z^i$

The estimated coefficient vector $\hat{\beta}$ can be obtained by setting the derivative in Eq. 72.20 equal to zero and solving for the unknown parameter (Hosmer et al. 2008, pp. 75–76).

Appendix 2: Covariance Graph



Brier score $B_t = 0.1578$; outcome index variance $\overline{d}_t(1-\overline{d}_t) = 0.0837$; bias $\overline{f}_t - \overline{d}_t = -0.2198$; slope $\overline{f}_1 - \overline{f}_0 = 0.0677$; forecast variance (Scatter) $S^2_{f_t} = 0.0372$

Yates (1982, pp. 143–148) and Arkes et al. (1995, pp. 121–123) provide detailed descriptions of a covariance graph. For illustration purpose, the above covariance graph depicts the characteristics of the 1-year survival forecasts generated by the proportional Cox's hazard upgrade model for FAs.

The abscissa shows the survival outcome index. The two possible outcomes for a FA in the upgrade model are "upgrade," which is denoted as 0 on the left, and "survival" (non-upgrade), which is denoted as 1 on the right. Of 141 holdout FAs available at 1-year lead time (forecast time $t = 1$ year), 128 FAs survived, and 13 FAs were upgraded. A vertical dotted line is located at the survival base rate, or the overall mean survival outcome index $\overline{d} = 0.9078$, on the abscissa. On the ordinate are the probability survival forecasts, categorized in deciles. A horizontal dotted line is located at the overall mean survival forecasts $\overline{f} = 0.688$ on the ordinate. The 45° solid line represents unbiased estimates. Bias can be measured as the vertical distance from the 45° line to the point where the vertical survival base rate line and the horizontal mean survival forecast line cross (marked as ◊). If a model produces unbiased forecasts, the vertical and horizontal dotted lines will cross on the 45° line, corresponding to a zero bias. If the two dotted lines meet below (above) the diagonal line, the model underestimates (overestimate) the survival outcome, corresponding to a negative (positive) bias.

In the covariance graph, the static upgrade model for FAs is 21.98 % too pessimistic in making 1-year survival forecasts.

On the vertical lines above the survival outcome (1) and the upgrade outcome (0) indices are the histograms for survival forecasts of 128 FAs that actually survived and 13 FAs that were upgraded, respectively. Survived and non-survived holdout FAs are stratified into distinct decile categories in the order of estimated survival probabilities. In this setting, FAs with survival forecasts varying from 0 % to 10 % are put together, those with forecasts ranging from 11 % to 20 % in another decile category and so on. The bars on the histograms illustrate the percentage of survival forecasts made at the individual probability deciles. The number of survival forecasts observed within each decile was attached to the corresponding bar for an easy reference. The further the histogram bars spread along the vertical lines, the greater the scatter (variance) of the survival forecasts.

The outcome index line extending vertically from 1 (on the right edge) includes the mean survival forecasts given to FAs that actually survived, $\bar{f}_1 = 0.6943$. The outcome index line drawn vertically from 0 (on the left edge) contains the average survival forecasts given to FAs that were actually upgraded, $\bar{f}_0 = 0.6265$. The dotted line linking $\bar{f}_1$ and $\bar{f}_0$ is the regression line for survival forecast on outcome index. The slope of the regression line is the difference between $\bar{f}_1$ and $\bar{f}_0$, or $\left(\bar{f}_1 - \bar{f}_0\right) = 6.77\ \%$. The further the regression line diverges from the horizontal line, the more discriminative the forecasts of the survived and upgraded groups.

Acknowledgments I am grateful to Capital Market Cooperative Research Center (CMCRC) for funding support. I wish to thank Paul Allison, Tony Hsiu His Chen, Sam Li Sheng Chen, and Amy Ming-Fang Yen for programming advice. Thanks are also due to Bob Kelly, Cheng-Few Lee, Graham Partington and Margaret Woods and for participants of the CEQURA Conference on Advances in Financial and Insurance Management for helpful comments.

References

Allison, P. D. (1984). *Event history analysis*. Beverly Hills: Sage.

Allison, P. D. (Ed.). (1995). *Survival analysis using SAS: A practical guide*. Cary: SAS Institute.

Altman, E. (1992). Revisiting the high yield bond market. *Financial Management, 21*(2), 78–92.

Altman, E. (1998). The importance and subtlety of credit rating migration. *Journal of Banking and Finance, 22*, 1231–1247.

Altman, E., & Kao, D. (1991). *Corporate bond rating drift: An examination of credit quality rating changes over time*. Charlottesville: The Research Foundation of the Institute of Chartered Financial Analysts.

Altman, E., & Kao, D. L. (1992a). Rating drift of high yield bonds. *Journal of Fixed Income, 2*, 15–20.

Altman, E., & Kao, D. (1992b). The implications of corporate bond rating drifts. *Financial Analysts Journal*, May–June, 64–75.

Amato, J., & Furfine, C. (2004). Are credit ratings procyclical? *Journal of Banking and Finance, 28*, 2641–2677.

Andersen, P. K. (1992). Repeated assessment of risk factors in survival analysis. *Statistical Methods in Medical Research, 1*, 297–315.

Arkes, H., Dawson, N., Speroff, T., Harrell, F., Alzola, C., Phillips, R., Desbiens, N., Oye, R., Knaus, W., Connors, A., & Support Investigators. (1995). The covariance decomposition of the probability score and its use in evaluating prognostic estimates. *Medical Decision Making, 15*(2), 120–131.

Bangia, A., Diebold, F., Kronimus, A., Schagen, C., & Schuermann, T. (2002). Rating migrations and the business cycle, with applications to credit portfolio stress testing. *Journal of Banking and Finance, 26*, 445–474.

Bessler, D., & Ruffley, R. (2004). Prequential analysis of stock market returns. *Applied Economics, 36*(5), 399–412.

Blossfeld, H. P., & Rohwer, G. (1995). *Techniques of event history modelling*. Mahwah: Lawrence Erlbaum.

Blossfeld, H.-P., Hamerle, A., & Mayer, K. U. (1989). *Event history analysis: Statistical theory and application in the social sciences*. Hillsdale: Lawrence Erlbaum.

Blume, M., Keim, D., & Patel, S. (1991). Returns and volatility of low-grade bonds 1977–1989. *Journal of Finance, 46*(1), 49–74.

Brier, G. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review, 78*(1), 1–3.

Cantor, R., & Packer, F. (1997). Differences of opinion and selection bias in the credit rating industry. *Journal of Banking and Finance, 21*, 1395–1417.

Carty, L. V. (1997). *Moody's rating migration and credit quality correlation, 1920–1996*. Special Comment (July 1997), Moody's Investor Service.

Carty, L. V., & Fons, J. S. (1994). Measuring changes in corporate credit quality. *Journal of Fixed Income, 4*(1), 27–41.

Caton, G., & Goh, J. (2003). Are all rivals affected equally by bond rating downgrades? *Review of Quantitative Finance and Accounting, 20*, 49–62.

Chen, L., Yen, M., Wu, H., Liao, C., Liou, D., Kuo, H., & Chen, T. H. (2005). Predictive survival model with time-dependent prognostic factors: Development of computer-aided SAS Macro program. *Journal of Evaluation in Clinical Practice, 11*(2), 181–193.

Cox, D. (1972). Regression models and life tables. *Journal of Royal Statistical Society Series B (Methodological), 34*, 187–220.

Dang, H. (2010). *Rating history, time and the dynamic estimation of rating migration hazard*. Ph.D. thesis, University of Sydney. http://hdl.handle.net/2123/6397. Accessed 30 June 2011.

Ederington, L. H., & Goh, J. C. (1999). Cross-sectional variation in the stock market reaction to bond rating changes. *The Quarterly Review of Economics and Finance, 39*(1), 101–112.

Figlewski, S., Frydman, H., & Liang, W. (2012). Modeling the effects of macroeconomic factors on corporate default and credit rating transitions. *International Review of Economics and Finance, 21*(1), 87–105.

Fledelius, P., Lando, D., & Nielsen, J. (2004). Nonparametric analysis of rating transition and default data. *Journal of Investment Management, 2*(2), 71–85.

Frydman, H., & Schuermann, T. (2008). Credit ratings dynamics and Markov mixture models. *Journal of Banking and Finance, 32*, 1062–1075.

Grunert, J., Norden, L., & Weber, M. (2005). The role of non-financial factors in internal credit ratings. *Journal of Banking and Finance, 29*, 509–531.

Guttler, A. (2011). Lead-lag relationships and rating convergence among credit rating agencies. *Journal of Credit Risk, 7*, 95–119.

Hamilton, D., & Cantor, R. (2004). *Rating transitions and default conditional on watchlist, outlook and rating history*. Special Comment, Moody's Investor Service.

Hill, P., Brooks, R., & Faff, R. (2010). Variations in sovereign credit quality assessments across rating agencies. *Journal of Banking and Finance, 34*(6), 1324–1343.

Hite, G., & Warga, A. (1997). The effects of bond rating changes on bond price performance. *Financial Analysts Journal*, (May/June), 35–51.

Holthausen, R. W., & Leftwich, R. W. (1986). The effect of bond rating changes on common stock prices. *Journal of Financial Economics, 17*, 57–89.

Hosmer, D., Lemeshow, S., & May, S. (2008). *Applied survival analysis: regression modeling of time-to-event data* (2nd ed.), Wiley Series in Probability and Statistics. Hoboken, N.J.: Wiley-Interscience, c2008.

Johnson, R. (2004). Rating agency actions around the investment grade boundary. *Journal of Fixed Income, 13*(4), 25–37.

Johnstone, D. (2002). More on the relative efficiency of probability scoring rules. *Journal of Statistical Computation and Simulation, 72*, 11–14.

Jorion, P., & Zhang, G. (2007). Information effects of bond rating changes: The role of the rating prior to the announcement. *Journal of Fixed Income, 16*, 45–59.

Kadam, A., & Lenk, P. (2008). Bayesian inference for issuer heterogeneity in credit rating migration. *Journal of Banking and Finance, 32*, 2267–2274.

Kalbfleisch, J. D., & Prentice, R. L. (1980). *The statistical analysis of failure time data*. New York: Wiley.

Kiefer, N., & Larson, C. (2007). A simulation estimator for testing the time homogeneity of credit rating transitions. *Journal of Empirical Finance, 14*, 818–835.

Lando, D. (2004). *Credit risk modeling: Theory and applications*. Princeton: Princeton University Press.

Lando, D., & Skodeberg, T. (2002). Analyzing ratings transitions and rating drift with continuous observations. *Journal of Banking and Finance, 26*, 423–444.

Lawless, J. F. (2003). *Statistical models and methods for lifetime data*. New York: Wiley.

Lee, E. T. (1980). *Statistical methods for survival data*. Belmont: Wadsworth.

Loffler, G. (2012). Can rating agencies look through the cycle? *Review of Quantitative Finance and Accounting*, 1–24. http://dx.doi.org/10.1007/s11156-012-0289-9.

Lucas, D., & Lonski, J. (1992). Changes in corporate credit quality 1970–1990. *Journal of Fixed Income, 1*, 7–14.

Mah, S., & Verde, M. (2004). *Rating path dependency: An analysis of corporate and structured finance rating momentum*. New York: Fitch Ratings Report.

Mann, C., Hamilton, D., Varma, P., & Cantor, R. (2003). *What happens to fallen angels? A statistical review 1982–2003*. Special comment. Moody's Investor Service.

Murphy, A. H. (1973). A new vector partition of the probability score. *Journal of Applied Meteorology, 12*, 595–600.

Murphy, A. H., & Winkler, R. (1977). Reliability of subjective probability forecasts of precipitation and temperature. *Applied Statistics, 26*, 41–47.

Namboodiri, K., & Suchindran, C. M. (1987). *Life tables and their applications*. Orlando: Academic.

Nickell, P., Perraudin, W., & Varotto, S. (2000). Stability of ratings transitions. *Journal of Banking and Finance, 24*, 203–222.

Samuelson, W., & Rosenthal, L. (1986). Price movements as indicators of tender offer success. *Journal of Finance, 41*, 481–499.

Sander, F. (1963). On subjective probability forecasting. *Journal of Meteorology, 2*, 191–201.

Standard & Poor's. (2001). *Playing out the credit cliff dynamics*. New York: Standard & Poor's [Re-printed Oct. 7, 2004].

Tsaig, Y., Levy, A., & Wang, Y. (2011). Analysing the impact of credit migration in a portfolio setting. *Journal of Banking and Finance, 35*, 3145–3157.

Vazza, D., Aurora, D., & Schneck, R. (2005a). *Crossover credit: A 24-year study of fallen angel rating behavior*. Global Fixed Income Research. New York: Standard & Poor's.

Vazza, D., Leung, E., Alsati, M., & Katz, M. (2005b). *Credit watch and ratings outlooks: Valuable predictors of rating behaviour*. Global Fixed Income Research (May). New York: Standard & Poor's.

Wei, L. J., Lin, D. Y., & Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *Journal of the American Statistical Association, 84*, 1065–1073.

Winkler, R. L. (1996). Scoring rules and the evaluation of probabilities. *Test, 5*(1), 1–60.

Yamaguchi, K. (1991). *Event history analysis*. Beverly Hills: Sage.

Yao, Y., Partington, G., & Stevenson, M. (2005). Run length and the predictability of stock price reversals. *Journal of Accounting and Finance, 45*, 653–671.

Yates, J. (1982). External correspondence: Decompositions of the mean probability score. *Organizational Behaviour and Human Performance, 30*, 132–156.

Creation and Control of Bubbles: Managers Compensation Schemes, Risk Aversion, and Wealth and Short Sale Constraints

73

James S. Ang, Dean Diavatopoulos, and Thomas V. Schwarz

Contents

J.S. Ang (✉)
Department of Finance, College of Business, Florida State University, Tallahassee, FL, USA
e-mail: jang@cob.fsu.edu; jang@garnet.acns.fsu.edu

D. Diavatopoulos
Finance, Villanova University, Villanova, PA, USA
e-mail: dean.diavatopoulos@villanova.edu

T.V. Schwarz
Stetson University, DeLand, FL, USA
e-mail: tschwarz@stetson.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_73,

Abstract

Persistent divergence of an asset price from its fundamental value has been a subject of much theoretical and empirical discussion. This chapter takes an alternative approach of inquiry – that of using laboratory experiments – to study the creation and control of speculative bubbles. The following three factors are chosen for analysis: the compensation scheme of portfolio managers, wealth and supply constraints, and the relative risk aversion of traders. Under a short investment horizon induced by a tournament compensation scheme, speculative bubbles are observed in markets of speculative traders and in mixed markets of conservative and speculative traders. These results maintain with super-experienced traders who are aware of the presence of a bubble. A binding wealth constraint dampens the bubbles as does an increased supply of securities. These results are unchanged when traders risk their own money in lieu of initial endowments provided by the experimenter.

The primary method of analysis is to use live subjects in a laboratory setting to generate original trading data, which are compared to their fundamental values. Standard statistical techniques are used to supplement analysis in explaining the divergence of asset prices from their fundamental values.

Keywords

Speculative bubbles • Laboratory experimental asset markets • Fundamental asset values • Tournament • Market efficiency • Behavioral finance • Ordinary least squares regression • Correlation

73.1 Introduction

The purpose of this study is to investigate the formation of speculative bubbles in asset prices under a laboratory setting.[1] Specifically, we investigate how to create, control, and dismantle bubbles as well as the conditions in which bubbles may or may not arise.

Speculative bubbles are induced in this study under a laboratory setting, where a New York Stock Exchange type of double-oral auction market (without a specialist) involving many traders is modeled. Speculative bubbles occur when buyers are willing to bid higher and higher prices for an asset which, in retrospect, are far in excess of its worth based on fundamentals. The bubbles ultimately burst and prices drop to a much lower level.[2] The stock market crash in

[1]The paper was previously published as Ang et al. (2009). The creation and control of speculative bubbles in a laboratory setting. In Lee, A., and Lee, C.F. (Eds.), *Handbook of Quantitative Finance and Risk Management* (pp. 137–164). Springer, New York.

[2]Stiglitz (1990), in his overview of a symposium on bubbles, defines the existence of bubbles to be: "if the reason that the price is high today is only because investors believe that the selling price will be high tomorrow – when 'fundamental' factors do not seem to justify such a price." Similarly, he defines the breaking of a bubble as marked price declines which occur without any apparent new information.

the USA in 1987 and in Japan in 1991–1992, the dot-com bubble in 2,000, and the recent housing bubble in the USA are examples.[3] In recent years, academicians and practitioners are slowly but grudgingly coming to the realization that the extant theories of stock market behavior, e.g., efficient market hypothesis and capital asset pricing theory, fail to explain the magnitude of fluctuations in the stock market. Not only have stock prices been found to fluctuate too much relative to fundamentals, but also there have been occurrences of speculative bubbles that could not be explained by arrival of new information. Several plausible explanations for bubbles are offered such as rational bubbles (Shiller 1988; West 1988), irrational bubbles (Ackert et al. 2002; Lei et al. 2001), judgement error (Ackert et al. 2006), and herding behavior (Froot et al. 1992).[4]

While some work has been done to show that bubbles can be abated with experience (Dufwenberg et al. 2005), an understanding of the formation of speculative bubbles is still important to researchers for several reasons. First, bubbles could cause significant disruptions in the asset market, not only by creating a large redistribution of wealth among investors but also by adversely affecting the supply of funds to the market as well as resource allocation among and within firms. Second, the identification of factors affecting the formation of bubbles is crucial in aiding regulators in designing policies to reduce the occurrence or magnitude of bubbles. In particular, if bubbles can be replicated in a laboratory setting, then various proposals to dampen bubbles could also be tested and compared for their effectiveness. Roll (1989) summarizes the difficulty with examining recent empirical results of the 1987 crash in this regard. Third, an understanding of the dynamic process of bubble formation would contribute to our knowledge of how to model the behavior of asset prices DeLong et al. (1989), Cutler et al. (1989).

In spite of some interesting recent theoretical developments, empirical research on the existence of bubbles tends to be inconclusive and with low power, e.g., Gurkaynak (2005), West (1988), Flood and Hodrick (1990). A major problem is the difficulty of specifying the fundamental value of an asset, since bubbles are defined as the price in excess of the fundamental values Bierman (1995), Robin and Ruffieux (2001). Without being able to calculate the time series of the asset's fundamental value, price movement could simply be caused by factors affecting the fundamental valuation of the asset, e.g., change in risk aversion and arrival of

---

[3]Other notable examples of bubbles include the Dutch tulip mania in the seventeenth century, the South Sea Islands Company bubbles Voth and Temin (2003), John Law's Mississippi Company scheme bubbles of the eighteenth century, the US stock market boom of the late 1920s, the Florida land price bubbles of the 1920s, the great bull market of the 1950s and 1960s, the high-tech stock boom of the early 1980s, and the boom and bust of the California and Massachusetts housing markets in recent years. However, due to the difficulties in specifying the fundamentals, there are still disagreements as to whether these cases could be explained by the fundamental, e.g., Garber (1990) versus White (1990).

[4]Outstanding surveys of this literature are provided by Porter and Smith (2003), Camerer (1989), Sunder (1992).

new information. And if the fundamental value could only be measured imperfectly using proxies such as past dividends, the imprecise estimates would, of course, reduce the power of any test.

The experimental approach reduces this problem (Cason and Noussair 2001). By design, the value of the fundamentals can be specified in advance; hence, there is no measurement problem. Any gross and persistent divergence of the asset price from the prespecified fundamental value can now be attributed to bubbles (Siegel 2003). In addition to reducing the identification/measurement problem, performing laboratory experiments to study asset bubbles has two other advantages. First, it allows different characteristics of the market institutions and participants to be introduced in a controlled manner. That is, relevant factors may be manipulated to create or discourage the formation of bubbles. This is an important feature because some of these factors may not be isolated in the real world for detailed study, while other factors are simply proposals in the design of market institutions of the future. Second, by controlling the information available to market participants, we can control the role played by unrelated or exogenous events, e.g., sunspots. Thus, the laboratory experiment approach to study asset market behavior complements the theory/model building process. The three types of variables chosen for analysis in this study are:

1. The compensation scheme of a portfolio manager. Allen and Gorton (1988) have argued that compensation schemes for portfolio managers may induce bubbles even in a finite horizon. Also, recent literature on tournaments (see James and Issac (2000), Ehrenberg and Bognanno (1990), and others) has shown that the level and structure of relative compensation influence participant behavior, while Hirota and Sunder (2005) have found that short horizons are important factors in the emergence of bubbles. Three types of compensation structure are used in these experiments: a linear compensation scheme based on portfolio performance and two versions of compensation based on relative performance in a short-term horizon.
2. Wealth constraint (tight/loose), supply of securities. An infinite number of trades (e.g., overlapping generations and the availability of credit) are often cited as a prerequisite for bubbles. Ricke (2004) discusses how credit made available from margin could generate bubbles. High liquidity leads to bubbles in the work of Caginalp et al. (2001). Scheinkman and Xiong (2003) and Hong et al. (2004) analyze the effect of a short sales constraint on the formation of bubbles. Therefore, experimenting with wealth constraints may provide valuable insights into the effectiveness of certain policies (such as margin rule change, credit availability) to control bubbles.
3. The type of investors in the market (speculative/conservative). This variable tests the Keynes-Hicks theory of speculation where differences in traders' willingness to take risks are the foundation of speculative markets. Traders in the experiments are pretested for attitudes toward risk taking.

   Bubbles are observed under the following conditions:

1. A market of speculators with short-term investment horizon.
2. A market of mixed conservative and speculative traders with short-term investment horizon.

3. A market of mixed trader types with short-term investment horizon using their own money.

On the other hand, bubbles are dampened under the following investment environments:

1. A market of conservative traders with a short-term horizon;
2. A market of mixed trader types with a long-term investment horizon;
3. A market of mixed trader types when the wealth of the traders, especially the bulls, is constrained;
4. A single-period trading environment.

The remaining part of this chapter is organized into four sections. Section 73.1 presents the hypotheses to be tested by incorporating them into the experimental design, which is discussed in greater detail in Sect. 73.2. The results are reported in Sect. 73.3 with Sect. 73.4 summarizing and concluding the chapter.

73.2 Bubbles in the Asset Markets

The possibility of asset bubbles has long been recognized; however, more formal theoretical development is of relatively recent vintage. Harrison and Kreps (1978), for instance, suggest that in general the right to resell the asset makes traders willing to pay more for it than they would if obliged to hold it forever. Thus, market price could exceed fundamental value. Literature on rational bubbles emphasizes that once a bubble is started, it would be rational to price the bubble component even if it is expected to burst with positive probability. Brunnermeier and Nagel (2003) examine stockholdings of hedge funds during the recent Nasdaq tech bubble and find that the portfolios of these sophisticated investors were heavily tilted toward (overpriced) technology stocks. However, this does not seem to be the result of unawareness of the bubble on the part of hedge funds.[5] At an individual stock level, hedge funds reduced their exposure before the prices collapsed, suggesting awareness and implicit pricing of the bubble component. On the other hand, whether bubbles can even get started has been questioned (Diba and Grossman 1987). Essentially, if there are a finite number of periods, starting from the next to the last period, the expectation that the bubble might end may be sufficient to keep it from ever starting. By the process of backward induction or an unraveling argument, bubbles will not exist. Moreover, if the number of trades is finite, withdrawal of early trades at a profit means the remaining traders would be at a negative sum game, i.e., with finite trades, will there be a "greater fool" who gets stuck when the bubble bursts.

[5]Griffin et al. (2003) examine the extant theoretical literature about bubbles which includes models where naive individuals cause excessive price movements and smart money trades against (and potentially eliminates) a bubble versus models where sophisticated investors follow market prices and help drive a bubble. In considering these competing views over the tech bubble period on Nasdaq, they find evidence which supports the view that institutions contributed more than individuals to the spectacular Nasdaq rise and fall.

Still, perturbing the model by adding uncertainties on the length of the horizon among traders or market size may preserve the possibility of asset bubbles.

Smith et al. (1988) are among the first to investigate the incidence of bubbles. Their design was to give traders common beliefs (according to one of Tirole's requirements) and long horizons of up to 15 trading periods. Bubbles are observed in several of their experimental markets. It is unclear, however, what institutional setting, other than long trading periods, induces bubbles in their study.

In a speculative market where bubbles could be present, speculative traders are more likely to purchase shares (and even more so, if bubbles are rationally priced) than risk-averse traders. Not only are they more willing to put a higher value on risky assets, they are also more likely to take the chance that they might not be able to sell out their inventory before the bubble bursts. Therefore, our first hypothesis is that bubbles are more likely to be formed in a market of risk-taking traders (speculators).

The compensation scheme could also affect the behavior of traders. For instance, Allen and Gorton (1988) and Allen and Gale (2002) show that an option-type compensation scheme for portfolio managers could induce speculative bubbles in asset prices. Portfolio managers are encouraged via incentive rewards to generate short-term trading gains even in a finite horizon world. The current practice of publishing and ranking the short-term investment performance of portfolio managers and the very substantial incentives to hedge fund managers' performance that may be based on unrealized gains on illiquid assets could give rise to adverse incentives. Portfolio managers who are concerned about these rankings will either take on a riskier strategy for the possibility of outshining their peers or they will simply play it safe and follow the crowd. Both portfolio strategies could lead to the formation of bubbles. The play safe by "following the herd" strategy will cause asset prices to have a strong positive correlation in the short term. Portfolio managers would be buying when others are buying, thus creating an upward price trend, and selling when others were selling, thus bursting the bubble it created. On the other hand, the pursuit of a risky strategy may be sufficient to create price leadership that is followed by others in the market. This would be more likely in an uncertain valuation environment. Temin and Voth (2003) suggest that riding the bubble may actually be a profitable strategy.

Portfolio managers are subject to an occupational hazard: unless they produce winning results, they stand a good chance of being fired. On the other hand, star performers receive seven or even eight figure incomes as new cash flows into the funds they manage. This compensation system is similar to tournament models where participants are paid according to their relative performance among a group of peers rather than on their absolute performance. Tournament systems are likely to produce increased performance when (1) there is difficulty in monitoring the activities of the agent (Rosen 1981), (2) when the agent possesses valuable information (Baker 1992), and (3) when good performance measures are available (Baker 1992). All three of these conditions exist in the realm of professional money management, and therefore it is probable that a relative performance compensation system will be effective in increasing manager performance.[6] Thus, it is hypothesized that a tournament incentive scheme that encourages a short-term horizon for portfolio managers is more likely to create bubbles.[7]

Finally, an important policy question has been whether restricting the availability of credit or the supply of securities in a market (by raising the margin requirement or allowing short sales) could reduce or even eliminate the formation of speculative bubbles. Ackert et al. (2006) find that price run-ups and crashes are moderated when traders are allowed to short sell). Most countries, including the USA and Japan, have adjusted these conditions in the recent past through adjustments in the use of stock index futures and by easing credit conditions. These changes have tended to occur subsequent to large declines in the country's equity markets. With limited or asymmetric ability to go short versus long, speculators on the long side have an advantage in acquiring funds for investment. Additionally, if the life of a bubble is uncertain and relatively long lasting, costly short sell will not be profitable even if the bubbles eventually burst. The usual experimental design often endows traders with a relatively large initial wealth such that the budget constraint is not binding. This experiment will test the effect of a tighter budget constraint by both reducing the initial endowment and increasing the supply of securities. It is hypothesized that a wealth constraint and/or relative increase in the supply of securities will reduce the incidence of bubbles.

To summarize, the effect of three factors, attitude toward risk, investment horizon, and wealth constraint, is examined as to their contribution to the creation and control of asset bubbles. They are tested by incorporating them into the experimental design of a laboratory setting described below; the importance of these factors in the formation and control of asset price bubble, singly and jointly, can now be formally examined.

73.3 Experimental Design

The evolutionary nature of laboratory experimental research is such that the results of any study act as a catalyst for new questions and therefore new experiments. As with Smith et al. (SSW) (1988), we note that many of our latter experiments were directly motivated by the results obtained from our earlier ones. This progression of thought and analysis will be apparent in the later section on results. Herein, however, we present the method of our investigation in comprehensive form.

[6]Becker and Huselid (1992), Ehrenberg and Bognanno (1990) have documented in field studies that such tournament compensation systems are effective in raising performance in professional golf and auto racing competitions.

[7]It is possible that if there is sufficient number of short horizon portfolio managers herding in the manner described by Froot et al. (1992), a bubble can start on basis of any information. Shleifer and Vishny (1990) also propose that the portfolio managers have short horizon; however, it is the risk of uncertain return from investing in the longer horizon that prevented disequilibrium to be arbitraged away.

Table 73.1 Experimental design

Participant groups[a]	Design	Initial endowment[b]	Investment horizon[c]	Risk aversion[d]	Experiments[e]
Las Vegas	1	2 securities 10,000 francs	Two period	Mixed	1, 2, 3, 5$
Las Vegas	2	2 securities 10,000 francs	ShortenedMixed		4m, 6m$
Las Vegas	3	5 securities 3,000 francs	ShortenedMixed		7x, 8x, 9t, 10t
FSU1	4	2 securities 10,000 francs	Two period	Single type	11s, 13c
FSU1	5	2 securities 10,000 francs	ShortenedSingle	Type	12sm, 14cm
FSU2	2	2 securities 10,000 francs	ShortenedMixed		15, 16
FSU2	6	10 securities 1,000 francs	ShortenedMixed		17, 18
Albania	1	2 securities 10,000 francs	Two period	Mixed	23
Albania	2	2 securities 10,000 francs	ShortenedMixed		24, 25
Albania	6	10 securities 1,000 francs	ShortenedMixed		26
Albania	7	2 securities 5,000 francs	Single Period	Mixed	19
Albania	8	2 securities 5,000 francs	Single Period/ Tournament	Mixed	20, 21
Albania	9	20 securities 500 francs	Single Period/ Tournament	Mixed	22

This table categories five designs of 14 experiments used to examine the impact of risk aversion, investment horizon, and credit/supply constraints (initial endowment) upon the formation and control of asset bubbles.

[a]The participant groups consist of the following:

Las Vegas represents students from the University of Las Vegas at Nevada.

FSU1 and FSU2 represent students from the Florida State University at two different time periods.

Albania represents students from the University of Tirana in Albania.

[b]The initial endowment refers to traders wealth position at the beginning of each trading year of an experiment. This endowment allows traders to sell (using provided securities) or buy (using francs, the currency used in these experiments). The additional securities and reduced currency endowments provided in Design 3 serves to better equate relative purchase and selling abilities.

[c]Investment horizon refers to the horizon within which traders effectively operate. A two-period horizon refers to a market where period A securities are based on the dividends paid in both periods (A and B) of a trading year. In a shortened investment horizon, the trader is induced (via the tournament compensation schedule of Table 73.3) to operate with a horizon which is shorter than the two-period environment in which the securities will pay dividends.

[d]Mixed risk aversion means that traders with various risk preferences were participants within the same market. Single type means that only speculative (s) or conservative (c) traders made up that market. The designations (s) and (c) appear next to experiments 11–14 in the last column.

[e]Notation is as follows:

$ represents a market where traders provided $20 of their own money to trade, the sum of which became the pool of money dispersed according to relative profit performance.

m,x,t represents the number of traders receiving the tournament prize as outlined in Table 73.3. This tournament compensation was used to induce a shortened horizon market and was differentially paid to the top two (t) or the top six (x) traders. In experiments marked (m), the first three trading years paid a bonus to the top six traders followed by years where only the top two traders received bonuses.

The creation of "bubbles" within asset markets is examined under the control of three primary factors: (1) the degree of trader risk aversion, (2) trader investment horizon, and (3) available investment capital/supply of securities. Table 73.1 summarizes the design of 14 experiments used to investigate these factors upon the presence of asset bubbles. Each of these experiments uses a common market mechanism that builds upon the earlier work of Forsythe et al. (1982), Plott and Sunder (1982), Ang and Schwarz (1985). These common features are summarized below.

73.3.1 General Market Design

1. A double-oral auction, similar to that used on the floor of major US exchanges, is replicated. The recruited traders are physically present within a single room during the course of trading. These traders are independent and trade solely for their own account. There are no specialists or other privileged traders.
2. Only those shares of a single generic security are traded. The sole attribute of these shares is the payment of dividends at the end of each period.
3. Each market (experiment) has ten trading periods. These periods are further categorized into five trading years, each of which consists of two contiguous trading periods (A and B). Endowments (discussed below) are reinitialized at the end of the second period of each year. Thus, the initial market represents a two-period model with each security entitled to two payoffs (dividends), one at the end of period A and the other at the end of period B.[8]
4. Each trading period lasts for 6 min, with opening, warning (at 5 and 5 1/2 min), and closing bells. Consequently, each experiment has a total of 60 (6 min $\times$ 10 periods) trading minutes. During the 6-min periods, traders can observe the continually updated bid-ask and past transacted prices.

73.3.2 Dividend Design

1. At the beginning of each year, each trader is endowed with trading capital and shares of the generic security. Each share pays dividends at the end of the first (A) and second (B) periods. The second period dividend is a liquidating dividend. Reinitializing of position at the beginning of each year allows for replication of decision making in experimental markets.[9]

---

[8]In the experiment, a trader has at least the following choices available:

(a) Maintain the endowed position by not trading and receiving the stochastic payoffs at the end of each period.
(b) Hold the securities through period A and sell in period B, in which case the investor will receive the first period dividend and the selling price.
(c) Sell the initial holdings in period A to receive the sale price.
(d) Buy additional shares in period A, receive dividends at the end of the period, and then sell the securities in period B.
(e) Sell the securities in period A and then buy back securities in period B in order to receive the dividends.
(f) Purchase a net amount of shares in both periods.
(g) Purchase and sell shares within each period.

[9]See Smith et al. (1988) for an example of when reinitialization is not used.

2. The dividends to be paid at the end of each period are stochastic. Two equally likely dividend outcomes are possible, the good (G) state and the bad (B) state. The realized state is announced at the end of each 6-min trading period as determined by the flip of a coin by the experimenter.
3. The dollar amount of the dividends paid at the end of each period depends upon the trader's type and the realized state. The 12 traders who make up each market are classified into three types to allow for differences in induced values. The dividend payouts for these three trader types are summarized in Table 73.2. As an example, at the beginning of year 1, the four traders of type I are privately informed that for period A they will receive either 350 or 110 for the good (G) and bad (B) states, respectively, and for period B either 250 or 150 for the good and bad states, respectively.
4. The trader type with the highest expected dividend (.5 × good dividend +.5 × bad dividend) is rotated each period so as to enhance trader uncertainty about equilibrium prices. Virtually all previous experimental studies have documented that given sufficient learning (through repeated trading) in a stationary dividend payout environment, prices will rather quickly approach the rational equilibrium level. This learning has two sources: (a) observation that one's own payouts are not changing and (b) observation that market-generated bids, offers, and transacted prices are not changing. Our expectation is that the greater the trader's reliance upon market-generated (as opposed to prior dividend) information, the more likely bubbles are to occur due to bandwagon and other crowd psychologies. If, instead of bubbles, we should observe that prices converge to rational equilibrium prices (as in the constant dividend studies), then this would strengthen our knowledge concerning efficiency in these laboratory markets. This result would also suggest that trading methods based upon historical prices alone would not have value.

73.3.3 Investment Horizon

1. Three types of investment horizon are provided within these experiments: a single-period, a two-period, and a shortened horizon. Initially, at the beginning of each trading year, a trader is entitled to two stochastic dividends for each security held, one each at the end of periods A and B. Therefore, at the beginning of period A, a rational trader will value the security for both its period A and period B stochastic dividends. Hence, all A period pricing should reflect a two-period investment horizon. Subsequent to the termination of period A trading and the announcement and payment of the period A dividend, period B trading proceeds. As the security is now only entitled to the B period dividend, a single-period investment horizon results for all B periods. Our hypothesis is that a shortened investment horizon increases the possibility of an asset pricing bubble. We test for this by creating

Table 73.2 Dividend design

Years[a]	Trader type[b]	Period A dividend state[c] G	B	Expected dividend[d]	Period B dividend state[c] G	B	Expected dividend[d]	Yearly expected dividend[e]
1	I	350	110	230*	250	150	200	430*
	II	250	150	200	200	140	170	370
	III	200	140	170	350	110	230*	400
2	I	200	140	170	350	110	230*	400
	II	350	110	230*	250	150	200	430*
	III	250	150	200	200	140	170	370
3	I	250	150	200	200	140	170	370
	II	200	140	170	350	100	230*	400
	III	350	110	230*	250	150	200	430*
4	I	350	110	230*	250	150	200	430*
	II	250	150	200	200	140	170	370
	III	200	140	170	350	110	230*	400
5	I	200	140	170	350	110	230*	400
	II	350	110	230*	250	150	200	430*
	III	250	150	200	200	140	170	370

This table presents the cash flow payoffs which a single asset will provide to its owner. This payoff is different for Trader Types I, II, and III and therefore provides for different fundamental valuations. Rational Expectations Equilibrium are determined by the trader type with the highest valuation for that period.

[a]Each experiment is composed of five trading years, each of which contains two trading periods A and B. Ownership of an asset in period A entitles the bidder of both period A and Period B dividends (two-period valuation) whereas period B ownership merits only that period's dividend (single-period valuation).

[b]There are three trader types in each trading year with four traders in each category. These trader types only differ by the amount of dividend cashflows that the single traded asset will provide its holder. The four traders within each category are rotated within the other categories so as to maintain an uncertain valuation environment.

[c]Dividend States refer to the stochastic payoff that will be provided to specific trader types given the occurrence of the G (Good) or B (Bad) state. The realization of the state of nature is determined at the end of each trading period by flipping a fair coin.

[d]Given equal fifty percent probability of occurrence of G or B, the expected dividend is the simple average of period G and B payoffs.

[e]The yearly expected dividend represents the summation of expected dividend for both periods A and B.

*Signifies trader type with the highest expected value

a tournament compensation package in period A. The incentive for traders is to concentrate upon their single (A)-period performance over the concerns of a rational two-period price. This incentive results in a shortened investment horizon.[10]

[10]It is important to note that there is a difference between a one- and two-period horizon and a shortened horizon. In a one-period model, only a single dividend is valued. In a two-period model, two dividends are valued. In our shortened investment horizon, the trader is induced to operate within a horizon that is different from that of his operating environment. That is, within

2. In this study, initially, a trader's dollar compensation is defined by the following in which we alter the compensation structure to induce a change in the length of a trader's investment horizon profit function:

$$P_i = f\left[d_{i,j,A} \cdot X_{i,A} + d_{i,j,B} \cdot X_{i,B} + (R_i - C_i)\right] \tag{73.1}$$

where:

P_i = dollar profit per trading year for trader i. It consists of dividend income and trading gains (losses) from both periods.

f = the conversion rate of francs into dollars.[11]

$d_{i,j,t}$ = the dividend paid in francs to trader i, given state j occurs in period t; j = G or B; t = A or B.

$X_{i,t}$ = the number of shares held by trader i at the end of period t; t = A or B.

R_i = revenues in francs for trader i for all shares sold during periods A and B.

C_i = costs in francs for trader i for all shares purchased during periods A and B.

In order to induce pressure for a shortened investment horizon, an additional compensation package is introduced in period A of some experiments as identified in Table 73.1. This tournament compensation system is based on the traders' relative performance as measured by the Tournament Performance Index (TPI) below:

$$TPI_i = R_i - C_i + MX_{i,A} \tag{73.2}$$

where R_i, C_i, and $X_{i,A}$ are as previously defined and M represents the closing market value of the shares. This closing market value is taken to be the price of the second to last transacted price for that period. This procedure is introduced in order to reduce the possibility of manipulating market value by collaboration on a final transaction. It represents a simplified version of the price-averaging process that takes place on most organized exchanges for the setting of opening and closing prices.

3. The tournament compensation system provides traders with an incentive to outperform each other in period A only. This incentive system increases the importance of single-period performance (in A) over two-period concerns; that is, it induces a shorter investment horizon in period A. A trader's compensation is dependent upon his relative rank as summarized in Table 73.3. In Schedule Six, the top six (of 12) traders are rewarded with francs ranging from 1,500 to 200.

---

a two-period operating environment, the trader is given an incentive to operate with a shorter (possibly single)-period horizon. This is quite different from a single-period model. This shortened horizon is a stronger test of market efficiency, in that the pressures are away from rather toward rational equilibrium prices, (as defined in Eq. 73.4, subsequently). The methodology is meant to emulate modern portfolio managers operating in an environment of perpetual horizon stock securities yet receiving tournament incentives to outperform colleagues on a short-term basis.

[11] Francs are the currency used within this study. They have been used successfully by Plott and Sunder (1982), Ang and Schwarz (1985), as well as others. Their primary benefit is to avoid the technical problem of dealing with small dollar amounts.

Table 73.3 Tournament compensation schedule

Schedule	Rank	TPI	Compensation
	1	Highest	1,500 francs
Six	2		1,000
(s)	3		700
	4		400
	5		200
	6		200
	7–12	Lowest	0
	1	Highest	3,000 francs
Two	2		1,000
(t)	3–12	Lowest	0

This table presents the additional tournament compensation schedule provided to traders based on their relative profitability in period A of certain experiments (see shortened investment horizon listed in Table 1). Relative profitability is measured by:

$$TPI_i = R_i - C_i + MX_{i,A}$$

where TPI_i is the tournament performance index for trader i, R_i is the revenues received from the sale of assets in period, C_i is the cost of assets purchased, M is the closing market value for the period, and $X_{i;A}$ represents the end-of-period asset holdings. Together the index measures the total of realized and unrealized capital gains. The addition of the tournament compensation to period A provides an incentive for traders to prefer period A capital gains over equivalent period B dividends and thereby induces a shortened (from the two period model) investment horizon.

[a]Two compensation schedules are introduced. The first provides for those traders who do better than the average (i.e., the top six) receive the additional compensation list. In the second schedule, only the top two "superstars" are richly rewarded. The tournament literature (e.g., Baker 1991) suggests that tournament systems, and especially schedule two, provide effective incentive systems to increase performance. This design is meant to emulate the short-term performance pressures faced by professional money managers.

Schedule Two is an alternative schedule which is hypothesized to induce even greater competitive pressure as only the top two traders are compensated greatly.[12] Ehrenberg and Bognanno (1990), Becker and Huselid (1992) find that the reward spread does cause increased performance incentives. Therefore, we expect Schedule Two to increase incentives for short-term pricing behavior. Table 73.1 summarizes the experimental use of the performance reward schedule.

[12]The compensation schemes depict the different ways portfolio managers are being rewarded: those who are above the average or beaten the market (Schedule Six) and those who are the superstars (Schedule Two).

73.3.4 Risk Aversion

1. Prior to selection, each potential trader was given a lengthy questionnaire. Intermingled within this material were two psychological tests on risk taking: the Jackson Personality Inventory (1976) and the Jackson et al. (1972) tests.[13] These two tests have been applied in laboratory Ang and Schwarz (1985) and field studies (Durand et al. 2006) and are more practical to administer than the theoretical risk measures found in the economics literature.[14] Those persons who score in the top 12, signifying the least risk averse, and the bottom 12, or the most risk averse, are invited to participate in the second stage of the experiment.
2. Traders for experiments 1–10 were students from the University of Las Vegas at Nevada and were recruited from a senior-level options class. These students had all taken two statistics, a corporate finance, a valuation, a portfolio analysis, and an options course. They were well trained in arbitrage, present value, and expected value. From this pool of students, 12 were chosen to participate based upon their attribute ranking in risk aversion. Participants were chosen so that a mix of risk aversion types was represented in the same market. Included were those who ranked at all levels of the scale, from high- to low-risk aversion. This was done so that differences in individual risk behavior could be tracked within an identical market environment.
3. Experiments 11–14 were conducted at Florida State University (FSU1), and as summarized in Table 73.1, these experiments were designed so that an experimental market consisted entirely of traders who were either relatively more risk averse (conservatives) or less risk averse (speculators). This experimental form allowed for evaluation of whether risk aversion is uniquely a necessary or sufficient condition for the presence of bubbles.
4. Experiments 15–26 were conducted at a later date at Florida State University (FSU2) and the University of Tirana in Albania. This was done to confirm the robustness of our results. We intentionally chose students from two

[13]The authors are aware of the work of Holt and Laury (2002) which was not available at the time of this study. According to Holt and Laury, their experiment shows that increases in the payoff level increase RRA. However, when estimating RRA, Holt and Laury assume that subject's utilities depend only on payments in the experiment. They fail to account for the wealth subjects have from other sources (see Heinemann 2003).

[14]The Jackson Personality Inventory is scientifically designed questionnaire for the purpose of measuring a variety of traits of interest in the study of personality. It was developed for use on populations of average or above average ability. Jackson states (1976, p. 9), "It is particularly appropriate for use in schools, colleges, and universities as an aid to counseling, for personality research in a variety of settings, and in business and industry." Of the 16 measurement scales of personality presented, one scale directly measures monetary risk taking using a set of 20 true and false questions. Mean and standard deviation measures for 2,000 male and 2,000 female college students are provided. Jackson et al. (1972) demonstrate four facets of risk taking: physical, monetary, social, and ethical. The authors' questionnaires are situational in that the respondent is asked to choose the probability that would be necessary to induce the respondent to choose a risky over a certain outcome. Jackson (1977) presents high internal consistency correlation between the risk measurement techniques.

different universities with different backgrounds to represent the two extremes in our test. Experiments 19–26 were administrated to subjects in Albania, who have a low degree of familiarity with capital markets, while experiments 15–18 were administered to Florida State University students (FSU2) who had taken a financial engineering course and completed another more involved laboratory asset market experiment. Hence, we consider these FSU2 students to be super-experienced relative to the students from Albania.

73.3.5 Validation Procedures

The following procedures are incorporated into the experimental design to ensure the reliability and external validity of the results:

1. To guard against the possibility that subjects' experience with trading could change their attitudes toward risk taking, they were retested. Subsequent to the first four experiments, additional risk questionnaires were given to the participants. A Spearman rank correlation (with initial risk rankings) was .902 with a t-statistic of 6.61 indicating that there had been no significant change in the relative risk attributes of the traders.
2. Videos were used to verify recorded information, to identify possible irregularities, and to train new subjects.
3. Subjects were given extensive training on the operation of the game; the main experiments were conducted on groups of experienced, if not super-experienced, traders.
4. Lengthy post-experiment questionnaires were also given to the subjects. Among other things, these were used to verify that the traders considered their trading strategies taken at the time of trade to be rational.

73.4 Results and Analysis

73.4.1 Control Experiments

For experiments 1–14, 5 experimental designs were used to test for the effects of risk aversion, investment horizon, and capital endowment upon the presence of asset bubbles. These designs are summarized in Table 73.1. The first design consisting of experiments 1, 2, 3, and 5 was control market where the two-period model was tested without extraneous influence from the three treatment variables mentioned above. Figures 73.1–73.5 plot the series of resulting prices. Bid and ask prices are represented by a "+" symbol and are connected by a vertical solid line. Transacted prices are identified by a solid horizontal line connecting each trade. From earlier laboratory studies, we would expect prices to converge to rational expectation equilibrium levels after an initial period of learning.

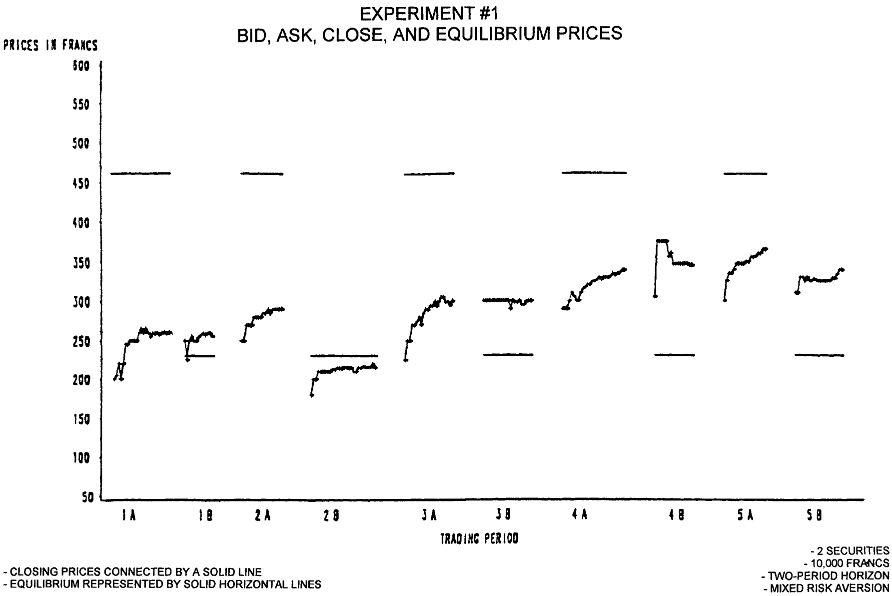


Fig. 73.1 Experiment 1: bid ask, close and equilibrium prices, two dividends with five trading periods each, and mixed risk aversion. Subjects are from Las Vegas

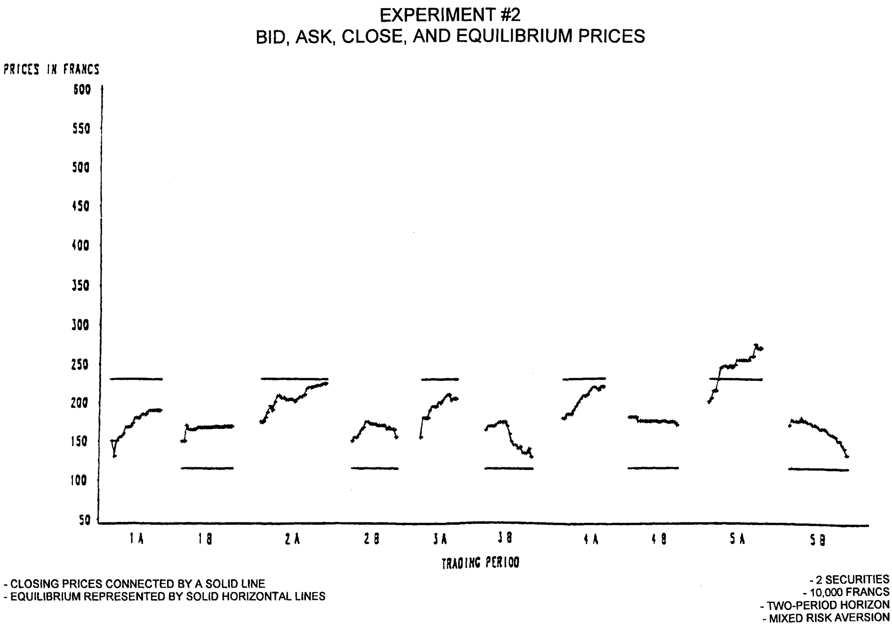


Fig. 73.2 Experiment 2: bid ask, close and equilibrium prices, two dividends with five trading periods each, and mixed risk aversion. Subjects are from Las Vegas

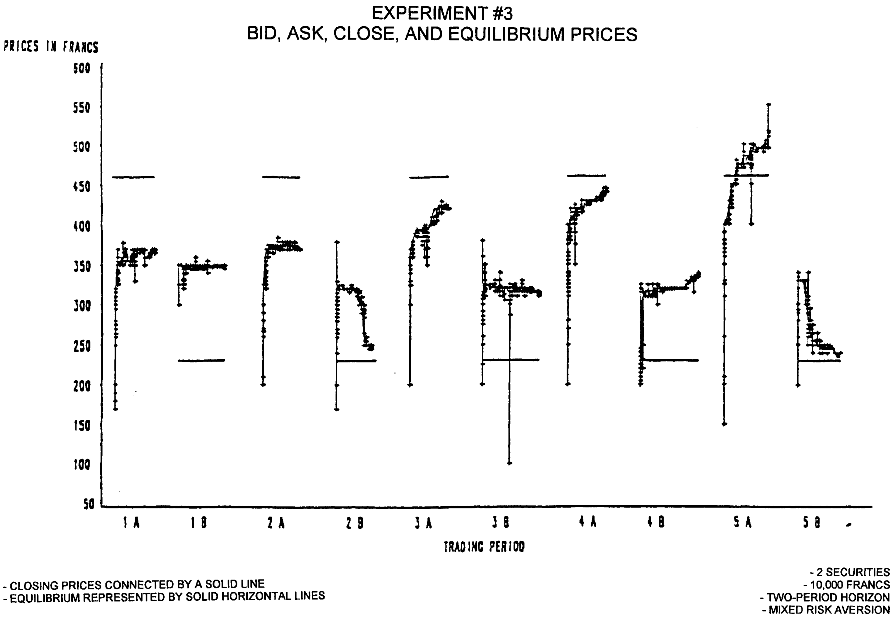


Fig. 73.3 Experiment 3: bid ask, close and equilibrium prices, two dividends with five trading periods each, and mixed risk aversion. Subjects are from Las Vegas

Two relevant concepts of equilibrium prices in these markets have been proposed (see Forsythe et al. 1982). The first is the naive equilibrium (NE) price. The NE is the highest price any trader in the market is willing to pay based upon his individual valuation of the expected dividends for the two periods or:

$$\mathrm{NE} = \mathrm{Max}_{\mathrm{k}}\left[\mathrm{E}\left(\mathrm{D}_{\mathrm{A,\,k}}\right) + \mathrm{E}\left(\mathrm{D}_{\mathrm{B,\,k}}\right)\right] \tag{73.3}$$

where:

k classifies the trader type (1, 2, or 3) based upon prior expected dividend valuations (see Table 73.2).

$E(D_{A,k})$ and $E(D_{B,k})$ are the values of expected dividends in periods A and B to the kth trader type.

The NE price is the market price that will prevail if the traders use only their private information to determine value. It is naïve in the sense that traders do not learn about the valuations of other traders from the market trading information. These traders also ignore the option value to trade, e.g., hold a security for a period and then sell it to another trader who would value it most in the remaining period.

The second is the perfect foresight equilibrium (PFE) price. It is equal to the highest total value that successive owners of the same share will pay or, in the experiment, the sum of the highest expected payoffs for periods A and B for all traders or:

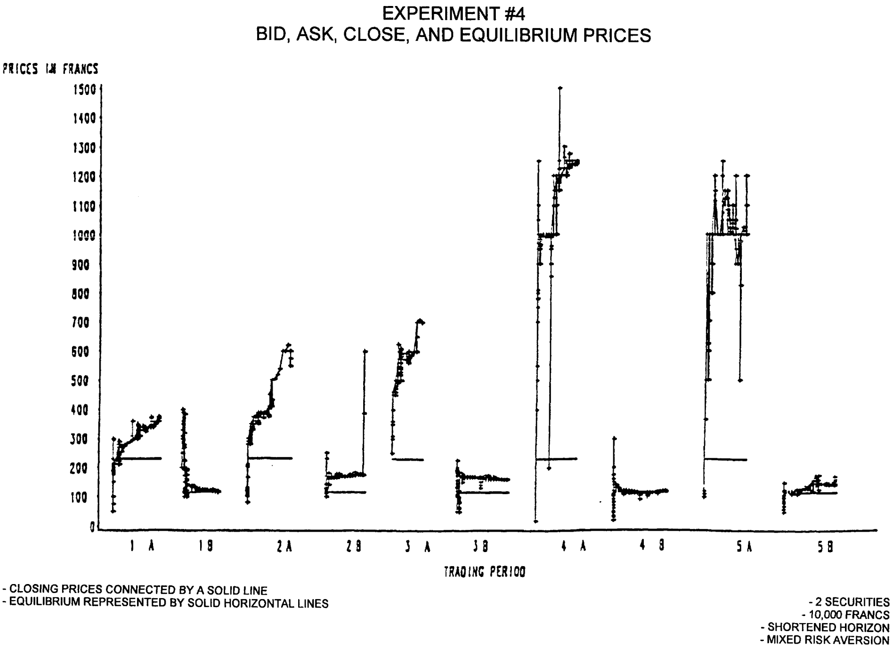


Fig. 73.4 Experiment 4: bid ask, close and equilibrium prices, one dividend with 5 trading periods, and mixed risk aversion. Subjects are from Las Vegas. Mixed bonus: to top 6 traders in first half and top 2 traders only in the second half

$$\mathrm{PFE} = \mathrm{Max}_k\,\mathrm{E}\big(\mathrm{D}_{A,k}\big) + \mathrm{Max}\,\mathrm{E}\big(\mathrm{D}_{B,k}\big) \tag{73.4}$$

These prices represent two extreme benchmarks in the continuum of the value of capital market in discovering information through trading. The NE price gives no role to capital market in price discovery, while the PFE price assumes full discovery, i.e., trading in capital market can correctly identify the share's highest value in each future holding period. They define, respectively, the lower and upper bounds of the share's fundamental value. Thus, with payoffs to traders and across holding periods under the control of the experimenters, we can now identify with certainty whether a stock is undervalued (when price is below NE) or overvalued (when price is above PFE) or is in a bubble (when price is grossly below NE or above PFE, as in a negative or positive bubble). If the experimental market captures a well-functioning capital market, learning and repeated trials would cause prices to converge toward PFE.

There are two properties in Eqs. 73.3 and 73.4 that are worth noting. First, NE and PFE prices are identical in a one-period world when price determination is closer to a simple auction of a single-period payoff. Second, when the payoff in the equations is dollars, as in cash dividends and capital gains or losses, NE and PFE give the risk neutral prices. In the absence of risk neutrality, a negative

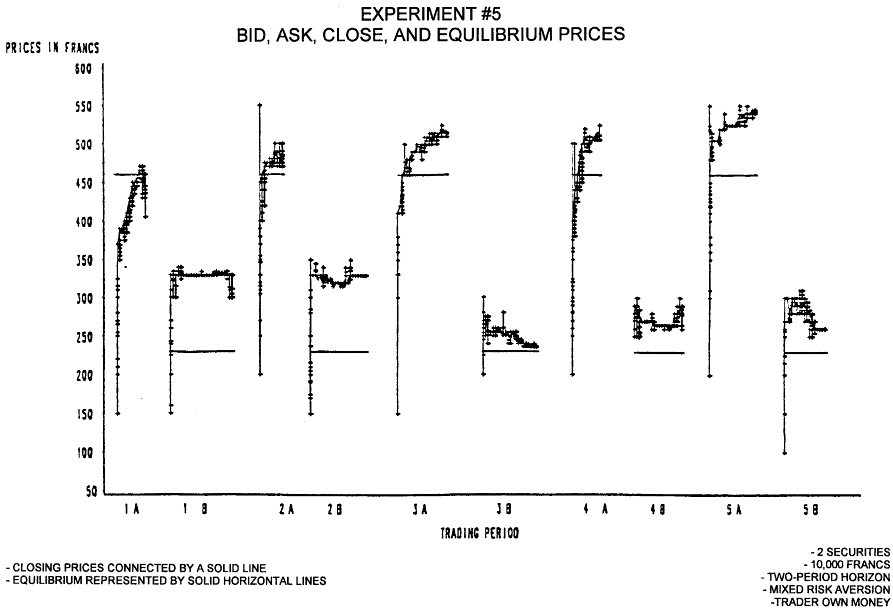


Fig. 73.5 Experiment 5: bid ask, close and equilibrium prices, two dividends periods with five trading periods each, and mixed risk aversion. Subjects are from Las Vegas and *use own funds*

difference between observed prices and these prices may be interpreted as a risk premium.

The results illustrated in Figs. 73.1–73.3 establish the validity of our experimental design as we are able to produce results similar to those obtained in previous experimental studies. In particular, we are able to reproduce the result that prices converge to PFE with learning and repeated trials. These prices are plotted as a solid horizontal line and are greater than the NE prices.[15] The inexperience of traders in experiment 1 is greatly reduced in experiments 2 and 3 as traders learn to cope with the large uncertainty in valuations (introduced by design). This pricing uncertainty

[15]Note that all odd-numbered experiments used the dividend design in Table 73.2. In order to differentiate between (1) learning about a stationary environment and (2) learning efficient valuation within laboratory markets, we created nonstationarity in equilibrium prices across experiments. In particular, for all even-numbered experiments, the dividend payoffs of Table 73.2 were simply cut in half so that rational equilibrium prices were also one-half that of the odd-numbered experiments. When this equilibrium dividend rotation is viewed in conjunction with the previously mentioned rotation of trader types, it becomes apparent that each individual trader was likely to view the environment (at least initially) as nonstationary. Consequently, any results that we show regarding equilibrium pricing and convergence would suggest that learning about valuation methods rather than a stationary environment creates rational valuation. That is, we are concerned about learning which takes place within the trader (how he values) not about the environment (stationary value). We are able to pursue this expanded question due to our debt to earlier authors who have already well established the presence of the latter.

increases the traders' reliance upon "market-generated information" in order to determine valuation. A microanalysis of traders' accounts in experiment 2 shows that some traders became actively involved in arbitrating between the A and B periods of a trading year. As a consequence, these prices tended toward their PFE equilibrium levels.

Traders' learning contributed to further pricing efficiencies in experiment 3. Some earlier "irrational" trades by selected individuals had resulted in substantial losses creating a "once-bitten" effect, and more rational decisions were followed subsequently. By the end of this experiment, prices in both the A and B periods were close to the PFE price.[16] A final examination of the validity of the experimental design was performed by requiring each trader from Las Vegas experiments 5 and 6 to "invest" his own money ($20) into the markets. As a result, it was possible for traders to lose as well as to win. The results, illustrated in Fig. 73.5, show continued price convergence toward equilibrium levels.[17] Of interest is the pattern of the bid-ask spread within a period. The data suggests that the primary resolution of uncertainty is obtained during the first transaction of a period. Subsequent trading tends to vary little from earlier levels with subsequently smaller bid-ask spread levels.

We conclude the control section noting that the experimental design creates price-revealing trades that foster PFE equilibrium pricing. While consistent with earlier research, these results extend our knowledge into a much more uncertain (nonstationary) valuation environment more typical of real-world asset markets. In addition, the validity of these results is not affected by whether or not a dollar investment is required from traders; trading behavior is similar under both environments.

73.4.2 The Formation of Bubbles

With a well-functioning experimental design established, we now sequentially introduce our hypothesized treatment variables. In experiment 4, we introduce the shortened trading horizon with a tournament prize as described in the experimental design. At this point, we have an advantage over previous studies in that we were able to recruit the identical 12 traders back. This level of experience will lead to converging equilibrium prices as opposed to bubble formation.[18]

[16]While period A prices exceeded the calculated PFE price of 460, this price is somewhat unknown to traders at this point. Prior trading results had created a history of B period prices averaging 320. Consequently, it was rational for a PFE trader to pay up to 550 (230 for A period plus 320 for B period sales price). The last trade in period 5A of 505 was well below that level. A more detailed presentation of the experimental results further reveals the rationality of these prices and is available from the authors upon request.

[17]Again, period A prices seem to drift upward due to initial excess pricing in period B.

[18]Our design is to eliminate the bubble effect of miscalculation caused by inexperienced traders as suggested by White (1990) and King et al. (1990). It is more useful and realistic to study the formation and control of bubbles in markets of experienced traders.

The effect of the tournament compensation is to shorten the traders' investment horizon in period A from a PFE two-period model. By providing tournament payment based on period A relative ranking, there is an increased incentive to generate period A capital gains over equivalent period B dividends. The tournament compensation, while increasing the incentive to win, does not necessarily equate to higher equilibrium prices. The prize is paid to the largest (realized and unrealized) relative capital gains which can be achieved in either a bull or bear market.

Extraordinary results are shown in Fig. 6.4 where five massive price bubbles are observed in each of the A periods. At this point, a new learning phase was initiated as traders competed strategically for the tournament prize. The dominant initial strategy centered on buying all available assets at increasing price levels, thereby creating artificial price support for capital gains. While this often resulted in achieving the prize, it also meant dealing with an inventory of overvalued assets in period B. Some traders actually lost money for the year even though they obtained the prize. It is important to note that the bubbles did not discourage the traders from participating, and at least for awhile the number willing to participate actually increased. Examination of asset holdings reveals that there were four to five active prize seekers in later bubbles versus one to two initially. In addition, seven to nine traders continued to hold securities at the periods' end rather than to sell out at extremely high bubble levels.

The much higher increased tournament reward structure for "superstar" performers of periods 4 and 5 (see Schedule Two of Table 73.3) resulted in the largest bubbles (consistent with our predictions) and with the greatest variability in prices and bid-ask spreads. Again, the buying frenzy in period 5 was lead by different traders than those in period 4. This continued rotation in trading leadership highlights that the results are not driven by a few misinformed traders. In fact, period B prices are very stable and efficiently priced. Furthermore, a trader questionnaire survey at the end of experiment 4 revealed that traders were fully cognizant of expected dividend value, yet they looked to both dividends and market-generated information to determine value. Traders stated that they were influenced by the behavior of their peers and were motivated to earn as much as possible. Several traders noted that the introduction of the tournament compensation stimulated them to take on more risk. The net result of these effects was to create a herd or bandwagon effect centered on market-generated information.

Despite the earlier findings of experiment 5, we tested the validity of these bubbles in an environment where traders used their own money rather than the experimenters.[19] Would such wild speculation occur when a trader's own money was at risk? Figure 73.6 clearly shows this answer to be yes. In all five A periods, average prices are over twice the equilibrium value. As before, period B pricing is very efficient and stable. That is, even though our traders engaged in bubble pricing, they arrived at it through rational means (Figures 73.7–73.10).

[19]To the author's knowledge, this is the first time traders in an experimental market of this type have used their own money to trade and still produced bubbles.

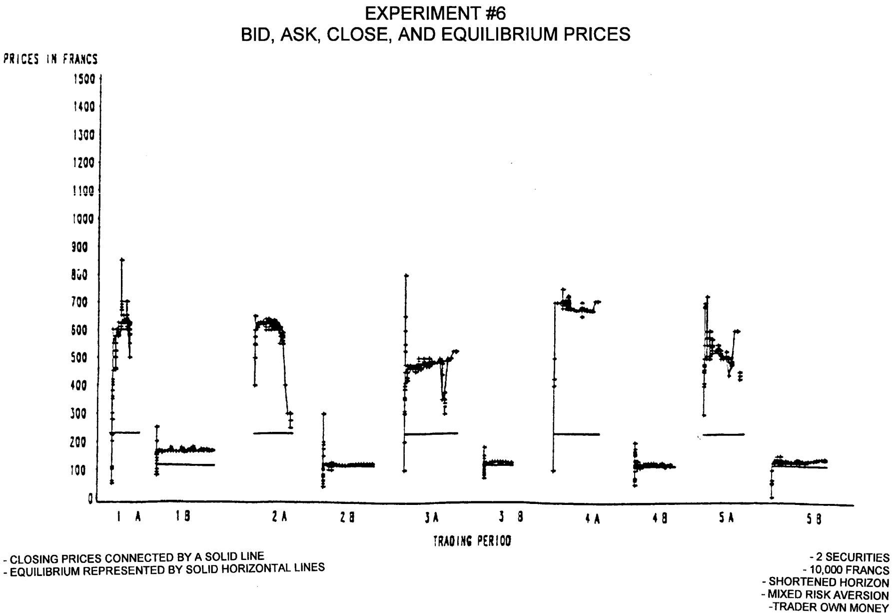


Fig. 73.6 Experiment 6: bid ask, close and equilibrium prices, one dividend with 5 trading periods, mixed risk aversion. Mixed bonus: to top 6 traders in first half and top 2 traders only in the second half. Subjects are from Las Vegas and *use own funds*

The traders in these markets had now participated in six experiments, the most of any research to date. Yet, even in the presence of super-experienced traders, we continue to find bubble formation. In addition, these traders were aware of the situation and made every opportunity to profit from the bubble.[20]

73.4.3 The Control of Bubbles

It became readily apparent from the earlier experiments that restrictions on the supply side of the market were having an influence on market prices. Many traders found themselves bound in their actions by the institutional makeup of the

[20]For instance, new strategies were employed at various stages (which perpetuated continuing uncertainty in the markets). At one time, the market actually stood still for an extended period. Then traders began to liquidate at any price rather than to replicate their earlier strategy of waiting until late in the period to sell out at bubble prices. Other traders began to try and scalp the market by driving prices both up and down, thereby generating capital gains in both price directions. Even others began to try and force losses on traders with large inventories and thereby improve their relative ranking. This was accomplished successfully in period 2A by selling at a loss (at a price below market prices) in order to create a low settle price, M (the second to last trade). Other attempts at this strategy followed in all remaining A periods. Nevertheless, bubbles persisted and many traders were frustrated in their inability to arbitrage them away.

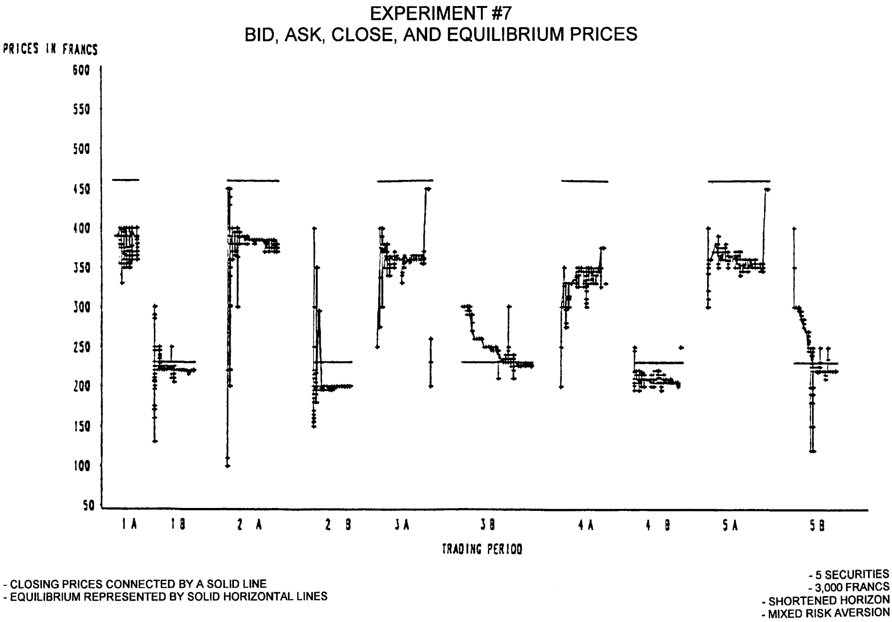


Fig. 73.7 Experiment 7: bid ask, close and equilibrium prices, one dividend with 5 trading periods, mixed risk aversion, and increase supply of shares. Bonus to top six traders. Subjects are from Las Vegas

experimental markets. Many of the traders suggested that they be allowed to short sell in future experiments so as to implement sell strategies in overvalued markets.

As previously mentioned, the tournament compensation system does not alter PFE prices since the prize can be achieved in any type of market environment and with any type of price pattern. Given the results of our previous experiments as well as traders' comments,[21] it appeared that buyers (longs) had an advantage over sellers (shorts). Is it possible that the bubbles we observe were due to differential market position in addition to a shortened investment horizon? In order to answer this question, we conducted four more experiments that provided traders with initial endowments and better equated the position of buyers and sellers. Rather than being endowed with two securities and 10,000 francs of trading capital as before, each trader is initially endowed with five securities and 3,000 francs (see Table 73.3).[22]

[21]Traders completed survey questionnaire at the completion of experiments 4, 6, and 10.

[22]Given that in experiment 6, period A prices averaged around 600, initial trading capital of 3,000 francs would provide buying power of roughly five securities. Consequently, the new buying power and selling power were a priori relatively equal. Even though period A prices turned out to be quite a bit lower in experiments 7–10, this did not create a great advantage to buyers since the supply of securities (5 traders $\times$ 12 traders $=$ 60) was relatively large for a 6-min trading period. As such, there was an ample supply of securities relative to buying power in order to drive prices down should traders turn bearish.

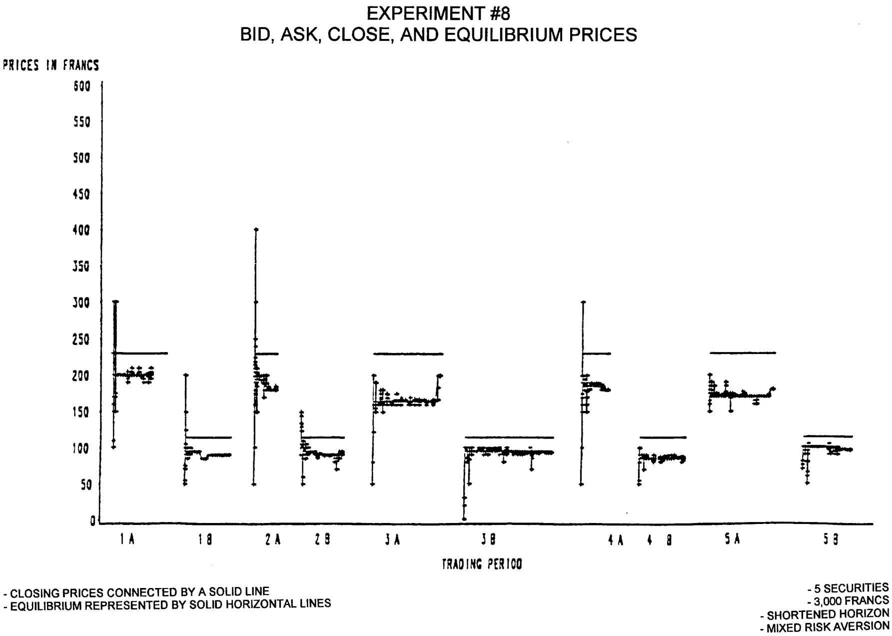


Fig. 73.8 Experiment 8: bid ask, close and equilibrium prices, one dividend with 5 trading periods, mixed risk aversion and increase supply of shares. Bonus to top six traders. Subjects are from Las Vegas

The price patterns of experiments 7–10 are as startling as the dramatic bubbles earlier. We find that the market is immediately priced at a discount to PFE.[23] This had never happened in any of the tournament periods before. If this had been simply the result of learning, we would have expected a gradual decline from the lofty levels of experiments 4 and 6. Rather, we see an immediate discount price which generally remains at a discount throughout all four experiments.[24] Overall, we consider this to be strong evidence that a necessary condition for the creation of

[23]We are unable to recruit all 12 traders back for experiments 7–10 due to graduation, taking of jobs, etc. We were, however, able to retain 7 of the original 12 traders. These traders had now participated in six previous experiments. The five replacements were drawn from the original pool of subjects that had completed the risk attribute questionnaires. These new traders were chosen to replace the risk types that had vacated so that in general, we maintained a wide dispersion of risk types within the market. In addition, some of these new traders had sat in as observers to previous experiments. Others viewed videos of the earlier experiments. All were instructed in the past experimental results, and the various strategies previously used were explained. As such, we do not believe that this change is a critical factor in the continuation of our investigation.

[24]An analysis of many of the last trades of period A for experiments 7–10 often shows either a sharp spike up or down. This illustrates that the traders had become very efficient (through learning) in their manipulation of closing prices. Given the large supply of securities available to squelch a price bubble, speculators were no longer singularly (due to large initial endowments of

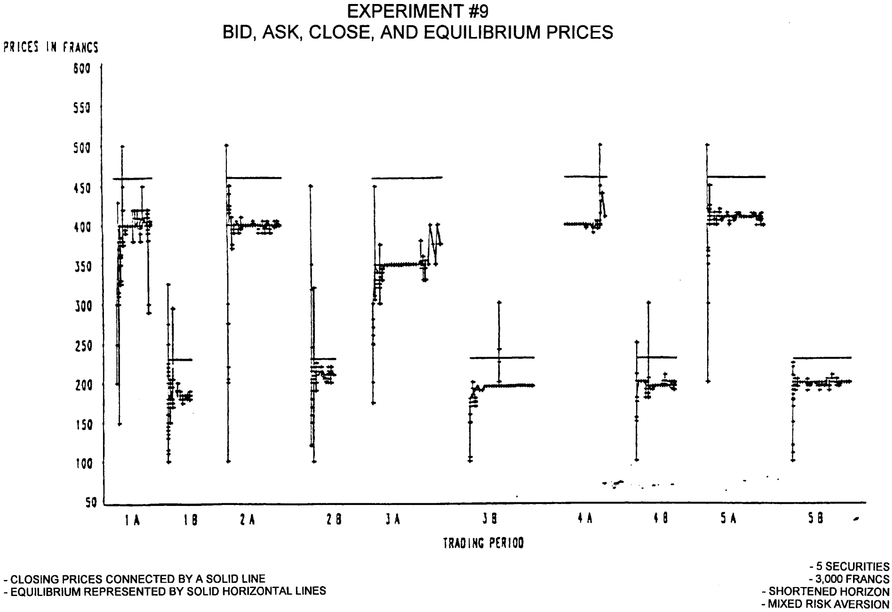


Fig. 73.9 Experiment 9: bid ask, close and equilibrium prices, one dividend with 5 trading periods, mixed risk aversion and increase supply of shares. Bonus to top two traders. Subjects are from Las Vegas

the large bubbles of these markets is that the institutional environment be biased toward more purchasing ability relative to that of selling.

In summary, experiments 7–10 highlight the importance of the supply of securities and the supply of investable funds that may be augmented by short selling. Bubbles observed in experiments 4 and 6 are immediately eliminated when the relative purchasing advantage of long traders is removed. Rational pricing

---

trading capital) able to create capital gains by driving market prices up. With this constraint, they quickly learned that all they needed to accomplish was to purchase the most securities at present prices and then drive the market up on the final few trades. This was often easily accomplished in that 1) only the second to last trade needed to higher in line with the calculation rules of the TPI and 2) as no surprise, there were always many traders who were willing to sell their securities at a price above the current level. The art to this strategy became a matter of timing; do not try to buy the market too early lest you run out of capital, and do not be too late lest you be unable to make the second to the last trade. There did not appear to be too much of a problem for buyers in accomplishing this in experiments 7 and 8; however, starting in experiment 9, some short traders, having become annoyed at bullish traders getting the tournament prize, began jockeying in these last seconds with the long traders in order to drive prices down. The results of such feuds appear in periods 3A, 4A, and 5A of experiment 9 and each A period of experiment 10. The winner of these duels increasingly became the trader who was best able to execute his trade. Eventually, trading activity become so enraged in the last 15 s of trading that the open outcry systems of double-oral auction began to break down.

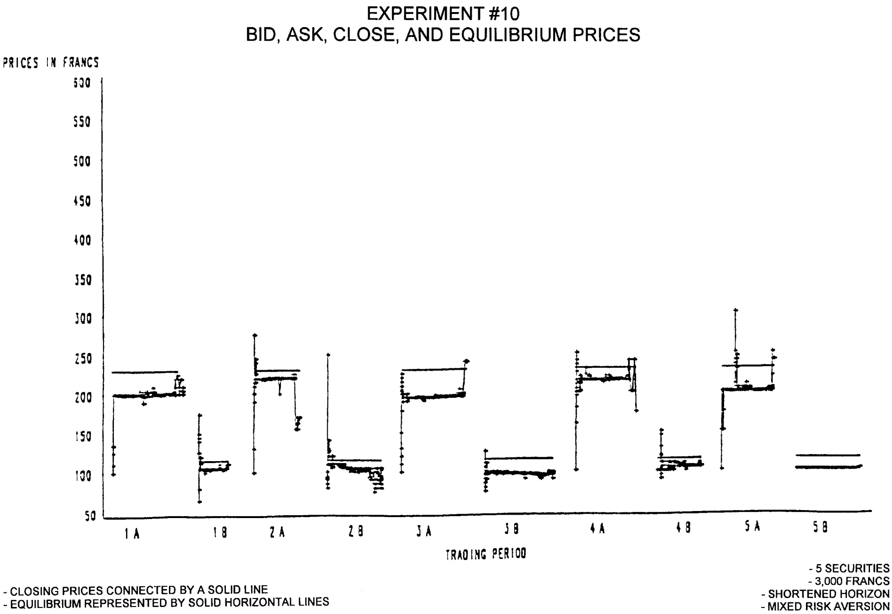


Fig. 73.10 Experiment 10: bid ask, close and equilibrium prices, one dividend with 5 trading periods, mixed risk aversion and increase supply of shares. Bonus to top two traders. Subjects are from Las Vegas

reflecting a modest risk premium results even when traders are faced with a shortened investment horizon.

73.4.4 The Impact of Risk Aversion

The results of previous experiments, especially 6, showed that trader risk aversion was an important factor in determining trader strategy and therefore price patterns. In general, it was found that speculative traders were more likely to seize upon the opportunity created by the introduction of uncertainty (via the tournament period) in search of capital gains. In contrast, the more conservative traders were likely to allow the speculators to act first by creating a positive price trend and would simply sell at inflated prices, or they would allow speculators to first initiate the "burst" of the bubble and then follow in their footsteps. Consequently, the conservative traders were often those responsible for the perpetuation of a direction initially set by speculators. The purpose of experiments 11–14 was to further test these relationships.[25]

[25]Experiments 11–14 were conducted at a second university, and therefore, the results provide information about the external validity of our experiments outside the setting of a single university.

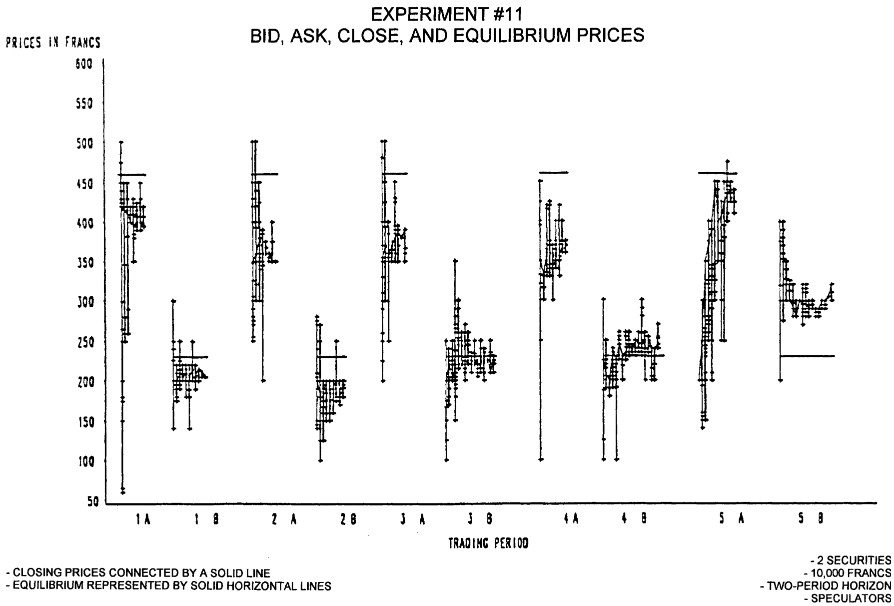


Fig. 73.11 Experiment 11: bid ask, close and equilibrium prices, two dividends and five trading periods each, with speculative traders. Subjects are from FSU, experiment 1

We chose at this time to create two separate trading groups according to risk aversion, each composed of 12 traders. These 24 traders were chosen from a pool of 70 students that completed the risk ranking questionnaire described earlier. The 70 respondents were rank ordered from highest to lowest in risk aversion. The top 12 and bottom 12 students were chosen to participate in the experiments. This method allows us to obtain good separation according to risk aversion. Contrary to our previous experiments, these markets would be made up entirely of one risk aversion class. We label these two risk classes as speculators and conservatives. This is a relative nomenclature as all of these traders are considered to be risk averse, and we only presume to provide an ordinal measure of risk aversion.

The design of these experiments follows that of experiments 1–6, as we wish to test for the presence of bubbles, and the initial endowments of experiments 7–10 have already been shown to eliminate bubbles. All of these traders had previously participated in two experimental markets and therefore can be considered experienced. Nevertheless, we test for rationality of pricing in experiments 11 and 13 before introducing the shortened horizons in experiments 12 and 14.

Figures 73.11–73.14 reveal that both markets are quite rational in that they charge a discount from PFE as a risk premium. As expected, the conservative traders of experiment 13 charge a larger risk premium than the speculative traders of experiment 11. This result provides strong evidence in support of our

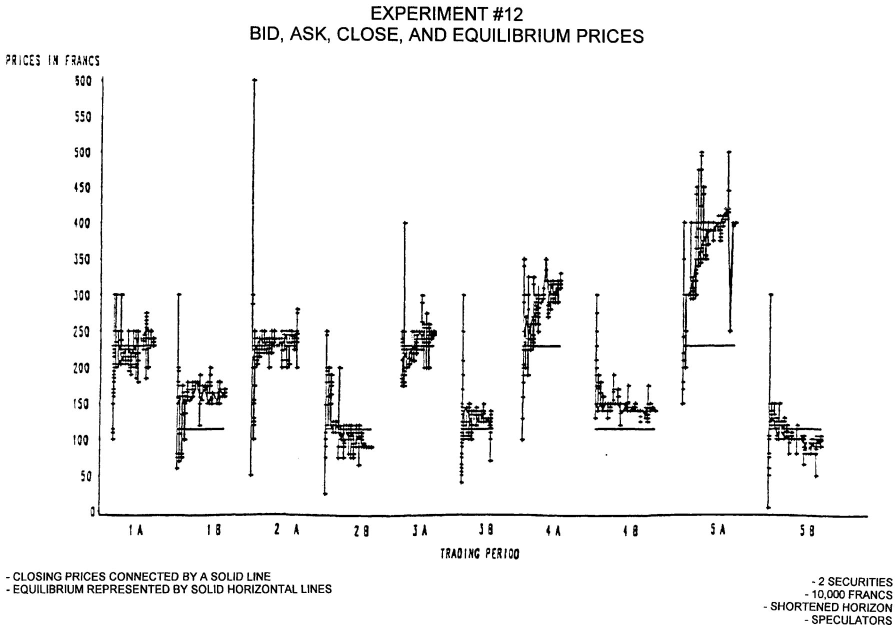


Fig. 73.12 Experiment 12 bid ask, close and equilibrium prices, one dividend with 5 trading periods, with speculative traders. Subjects are from FSU, experiment 1. Mixed bonus: to top 6 traders in first half and top 2 traders only in the second half. Subjects are from FSU, experiment 1

measure/separation of risk aversion. We also note that the speculative group exhibits prices above the PFE levels of period B. This is consistent with our earlier results where this was found in the single-period case.

Experiments 12 and 14 introduce the tournament compensation schedule to induce a shorter investment horizon. As expected, the speculative group seizes upon the opportunity and price bubbles are generated in the latter periods. Also to no surprise, the conservative group does not create the pressure necessary to cause bubbles to form. As a result, we conclude that a necessary condition for asset bubbles is the presence of speculators.[26]

[26]A detailed examination of individual trades reveals the speculative group of traders who are found to be more innovative in designing new trading strategies both in the creating and bursting of bubbles. The finding is consistent with the observation made by Benjamin Friedman (1992) in his review of a dozen NBER working papers on asset pricing. He finds these recent research results demonstrate that rational speculative behaviors such as an attempt by investors to learn from other investors, to affect another's opinion, or to simply engage in protective trading could in some context, such as imperfect information, magnify price fluctuations.

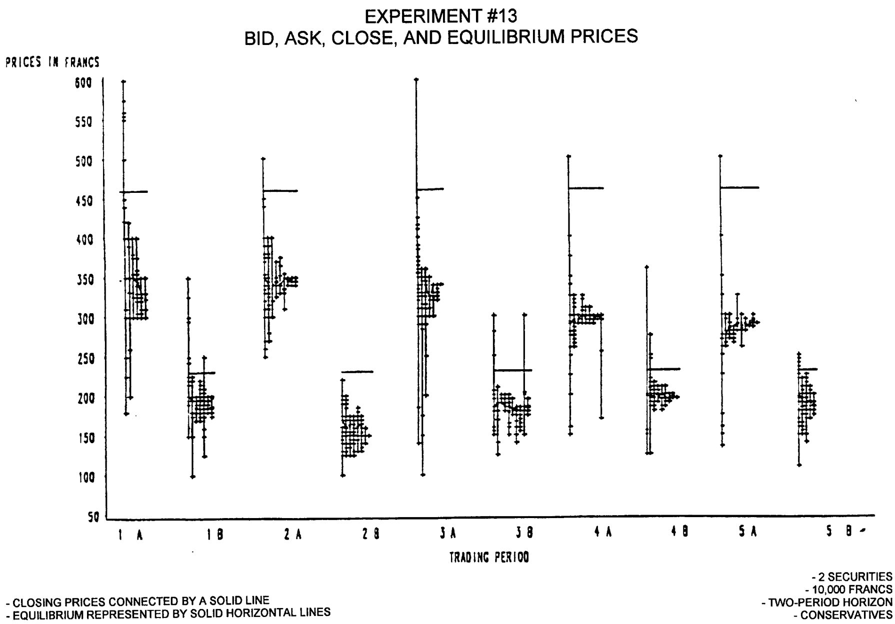


Fig. 73.13 Experiment 13 bid ask, close and equilibrium prices, two dividends and five trading periods each, with speculative traders. Subjects are from FSU, experiment 1. Mixed bonus: to top 6 traders in first half and top 2 traders only in the second half. Subjects are from FSU, experiment 1

73.4.5 The Formation of Negative Bubbles

We have just learned that the effect of the reduced investment horizon is to increase the incentive for short-term speculative gains and that speculative traders are those most eager to earn these profits. We now extend the research design to investigate the question of whether negative bubbles are also possible. We test this proposition by conducting four new experiments (labeled as experiments 15–18 in Table 73.1). We conduct experiments 15 and 16 as "controls" to replicate the positive bubble environment found in experiments 4, 6, and 12. Experiments 15 and 16 validate our previous results with a new set of experimental subjects, while Figs. 73.15 and 73.16 plot the pattern of close prices relative to the equilibrium level (horizontal line). In both experiments, large positive bubbles emerge in most trading years.

We now pose the following question, "Would an environment opposite to that of Design 2 lead to negative bubbles?" We keep the structure of Design 2, but since it was the unequal endowment effect (more purchasing power versus selling pressure, under 2 securities, 10,000 francs) that created the ability to pursue profits in a positive bubble environment, we reverse the endowment effect in experiments 17 and 18 by providing 10 securities and 1,000 francs to each trader. This one change provides traders in experiments 17 and 18 with a much greater ability to buy relative to sell.

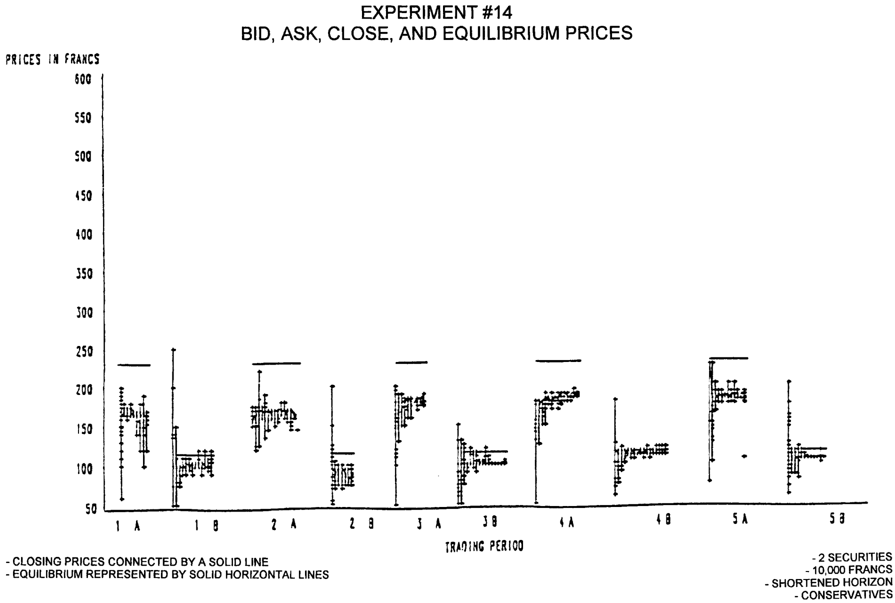


Fig. 73.14 Experiment 14 bid ask, close and equilibrium prices, one dividend with 5 trading periods, with conservative traders. Mixed bonus: to top 6 traders in first half and top 2 traders only in the second half. Subjects are from FSU, experiment 1

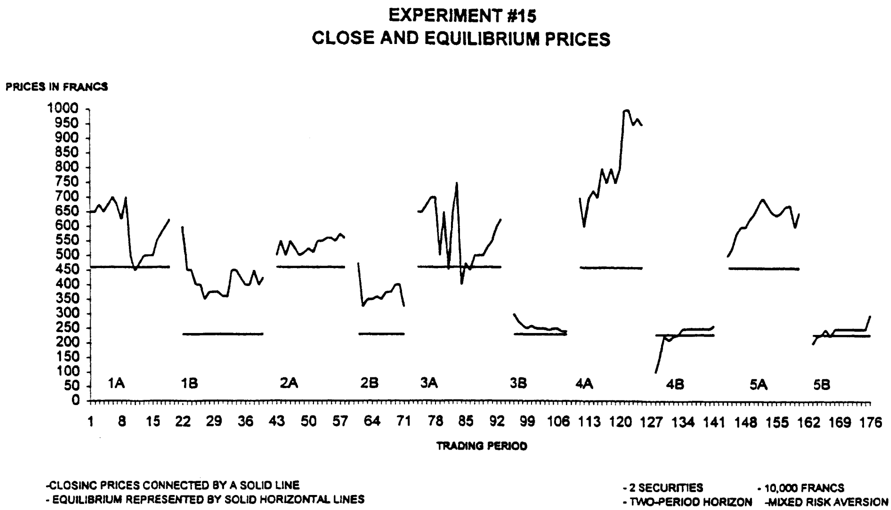


Fig. 73.15 Experiment 15 bid ask, close and equilibrium prices, two periods, mixed risk aversion. Subjects are from FSU, experiment 2. Considered most sophisticated

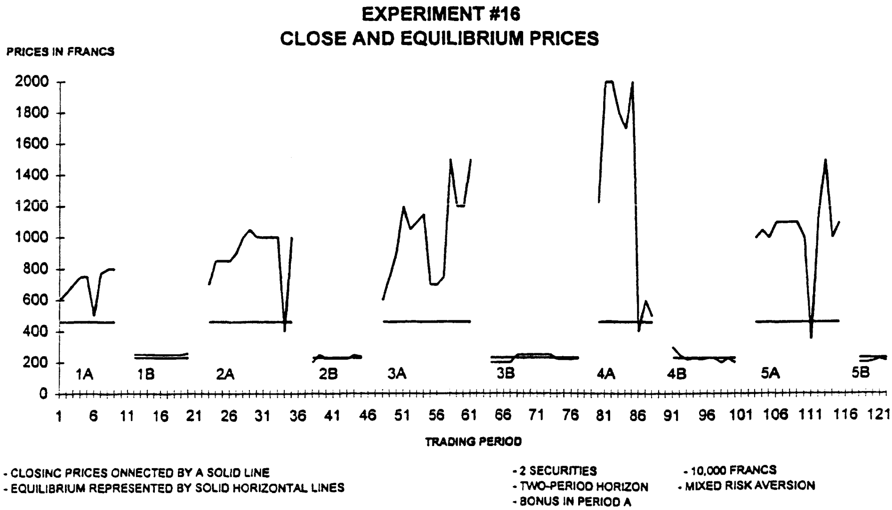


Fig. 73.16 Experiment 16 bid ask, close and equilibrium prices, one dividend with 5 trading periods, and mixed risk aversion. Subjects are from FSU, experiment 2. Considered most sophisticated

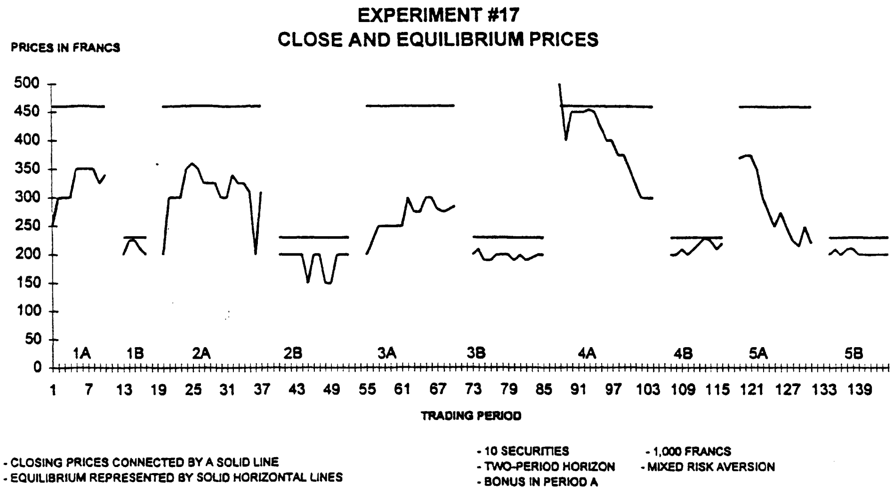


Fig. 73.17 Experiment 17 bid ask, close and equilibrium prices, two periods, and mixed risk aversion. Increase supply of shares. Subjects are from FSU, experiment 2. Considered most sophisticated

The results plotted in Figs. 73.17 and 73.18 show a preponderance for negative bubbles. While the initial 4 years of experiment 17 show some learning adjustment to this new and difficult trading scheme, large price discounts emerge to the extent that period 5A's closing price is insignificantly different than period 5B's which is

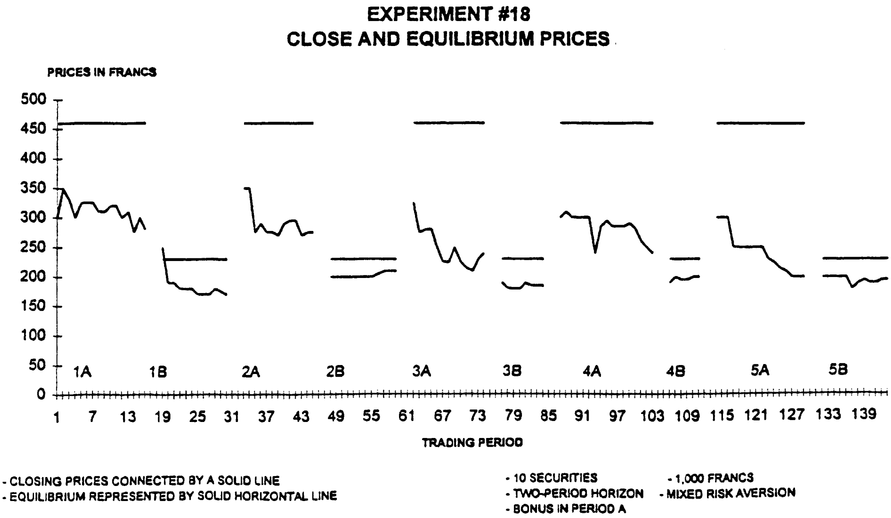


Fig. 73.18 Experiment 18 bid ask, close and equilibrium prices, two period, and mixed risk aversion. Increase supply of shares. Subjects are from FSU, experiment 2. Considered most sophisticated

a single period receiving only a single dividend. By experiment 18, each year shows downward trending markets in each A period. The reader may notice that the positive bubbles seem to burst, while the negative bubbles don't. However, since our design did not allow more cash to be made available through borrowing or infusion, correction may not be observed in the short trading period.

73.5 Conclusions

The results of this study have a number of implications for real-world markets. Experiments 1–6 seem to imply that within an environment that restricts selling pressures, a shortened investment horizon is sufficient to create asset bubbles. In application to the real world, short-term performance of traders, portfolio managers, etc., could create pressures leading to price bubbles. Experiments 7–10 provide restrictions to the previous conclusion in that a shortened investment horizon creates bubble pressure only when the market environment favors buyers over sellers. Unfortunately, most of our real-world securities markets do have such a bias via restricted short sales, asymmetric leverage for longs versus shorts, restricted options and futures, and the like. Experiments 11–14 add to the puzzle by demonstrating the role of speculators within bubble formation. As a whole, the study suggests that necessary and sufficient conditions for the formation of asset bubbles are a shortened investment horizon, restricted selling activity relative to buyers, and the presence of speculators. We have also shown that repeated replication of these experiments under different settings still produces robust results.

The first and third variables are a matter of fact within US securities markets, while restricted selling activity relative to buyers can take many forms. Either enhancing the buyer's position or restricting the seller's position is sufficient. Examples include increasing purchasing (speculative) ability through reduced stock margin levels, introduction of high leverage stock index futures, and, in macroeconomic terms, a growing money supply or savings level. This latter variable may help explain the previous high levels of the Japanese equity market. The high level of Japanese savings creates very large endowments available for investment purchase. Given a limited supply of securities, our experimental markets show that these conditions will lead to a bubble. They also suggest that the bubble will burst when there is greater equating between the supply and demand. Recent changes in the Japanese institutional framework may, as predicted by this study, have led to the bursting of that bubble.

The primary prescription put forth for regulatory authorities in eliminating unnecessary market volatility resulting from asset bubbles is to create an institutional environment that does not restrict the transfer of information to the market. Structure the variables so that both bulls and bears have equal costs in executing their trades.

Appendix 1: Statistical Analysis

Table 73.4 summarizes the ordinary least squares regression analyses of the impact upon the divergence of asset prices from their PFE levels in period A.[27] In particular, we test the following relation:

$$\mathrm{P^L - PFE} = f(\mathrm{I, E, I * E, T, I * E * T, S, I * S, A, I * A, \ \$}) \tag{73.5}$$

where:

$\mathrm{P^L}$ – PFE = the deviation from equilibrium for period A of each trading year where $\mathrm{P^L}$ is the last trade of the period and PFE is the perfect foresight equilibrium price,

f = a linear additive model,

I = a dummy variable representing the shortened investment horizon according to Table 73.1 I = 1 for shortened horizon and 0 otherwise (i.e., experiments 4, 610, 12, 14),

E = a dummy variable representing the endowment effect according to Table 73.1. E = 1 when 2 securities are issued and 0 otherwise (i.e., experiments 1–6, 11–14),

I*E = an interaction dummy variable representing both a shortened investment horizon and two-security endowment (i.e., experiments 4, 6, 12, 14),

[27] See the classic textbooks by Greene (2012), Wooldridge (2010), or Hayashi (2000) for details on the implementation and interpretation of OLS.

Table 73.4 The impact of investment horizon, credit/supply constraints, risk aversion, and other variables

Model	Intercept	I	E	I*E	T	I^*E^*T	S	I^*S	A	I^*A	$\$R^2$
Panel A											
Experiments 1–10 (n = 50)[b]											
1	NOINT[c]	−45.1 (–1.48)	−38.5 (−1.26)	512.3^{***} (8.40)							.67
2	−48.1 (−1.43)		9.6 (.23)	311.8^{***} (6.28)	6.1 (0.13)	382.3^{***} (4.57)					.79
3	-178.1^{***} (−2.73)		145.1^{*} (1.80)	195.9^{***} (2.23)	17.3 (0.38)	282.5 (3.08)		0.8 (0.01)	292.4^{*} (1.77)	−23.6 (−0.54)	.80
Panel B											
Experiments 11–14 (n = 20)											
4	-150.5^{***} (−11.17)	109.1^{***} (6.98)					89.9^{***} (4.92)				.75
5	-139.0^{***} (−9.30)	61.8^{**} (2.71)				61.1^{***} (2.83)	67.0^{***} (3.17) 45.8 (1.53)				.83

This table shows the extent to which certain variables cause a deviation from perfect foresight equilibrium values.[a] The following regression is estimated separately for experiments 1–10 and 11–14 according to the experimental design of Table 1.

P^L-PFE = f(I, E, I^*E, T, I^*E^*T, S, I^*S, A, I^*A, \$)

[a] t-values in parentheses. Variables defined as follows:

PL-PFE represents the deviation from equilibrium for Period A of each trading year where PL is the last trade of the period and PFE is the Perfect Foresight Equilibrium price

I is a dummy variable representing the shortened Investment horizon according to Table 73.1. I = 1 for shortened horizon, and 0 otherwise (i.e., experiments 4,6–10,12,14)

E is a dummy variable representing the Endowment effect according to Table 73.1 E = 1 when 2 securities are issued, and 0 otherwise. (i.e., experiments 1–6, 11–14)

I*E is an interaction dummy variable representing both a shortened investment horizon and two security endowment (i.e., experiments 4,6,12,14)

T is a dummy variable representing the Tournament effect according to Table 73.1. T = 1 when there is a tournament prize for two traders only, and 0 otherwise (i.e., experiments 4,6,12,14 (years 4&5) and 9,10)

I*E*T is an interaction dummy variable representing a shortened investment horizon, a two security endowment, and a tournament effect (i.e., experiments 4,6 (years 4&5), and 9, 10)

S is a dummy variable representing experiments composed entirely of Speculators according to Table 73.1. S = 1 for experiments 11 and 12 and 0 otherwise

I*S is an interaction dummy variable representing the shortened investment horizon and a pure speculative trader market (i.e., experiment 12)

A represents the ratio of end-of-period Asset inventory for speculative traders to total asset holdings. Speculative traders are those who scored in the top one-half of the risk measurement questionnaires

I*A is an interaction variable for shortened investment horizon and ratio asset holdings for speculators (experiments 4,6–10,12,14)

$ is a dummy variable representing experiments where traders risked their own money according to Table 73.1. $ = 1 when own money is used, and 0 otherwise (i.e., experiments 5 and 6)

[b]As reported in Table 73.1, experiments 1–10 consisted of traders with a wide range (mixed) of risk aversion. Experiments 11 and 12 were composed only of speculative traders with experiments 13 and 14 composed of conservatives.

[c]NOINT means the regression was run by suppressing the intercept.

*, **, ***Signify statistical significance levels at .10, .05, and .01, respectively.

T = a dummy variable representing the tournament effect according to Table 73.1. T = 1 when there is a tournament prize for two traders only and 0 otherwise (i.e., experiments 4, 6, 12, 14 (years 4 and 5), 9, and 10),

I*E*T = an interaction dummy variable representing a shortened investment horizon, a two-security endowment, and a tournament effect (i.e., experiments 4, 6, 9, 10 (years 4 and 5)),

S = a dummy variable representing the extent to which speculators participated in the experiments according to Table 73.1. S = 1 for experiments 11 and 12 and 0 otherwise,

I*S = an interaction dummy variable representing the shortened investment horizon and a pure speculative trader market (i.e., experiment 12),

A = the ratio of end-of-period asset inventory for speculative traders to total asset holdings. Speculative traders are those who scored in the top one-half of the risk measurement questionnaires,

I*A = an interaction variable for shortened investment horizon and ratio asset holdings for speculators (experiments 4, 6–10, 12, 14),

$ = a dummy variable representing experiments where traders risked their own money according to Table 73.1. $ = 1 when their own money is used and 0 otherwise (i.e., experiments 5 and 6).

Due to their differential design, the results for experiments 1–10 appear separately in Panel A and those for experiments 11–14 in Panel B.

Model 1 of Panel A tests the impact of (1) I = 1, a shortened horizon; (2) E = 1, a restricted endowment effect (wealth and supply effects); and (3) I = 1, E = 1, an interaction of a shortened horizon with restricted initial endowment. Given that the regression was run with no intercept, the coefficients represent estimates of each variable's independent impact. The results suggest that neither a shortened investment horizon nor a biased endowment effect (advantage to "bulls" versus "bears") is sufficient to induce bubble behavior. However, the interaction of these two variables is highly significant in explaining the bubble results of these experiments. That is, an environment that provides both the incentive and the ability to profit from a bubble will likely result in positive price divergence.

As hypothesized earlier, we test for the heightened effect of tournament incentives (i.e., T = 1) by examining the effect of "superstar" prizes paid to only the top two traders (as outlined in Tables 73.1 and 73.3). We also test for an interaction effect with a shortened horizon (I = 1) and restricted endowment (E = 1). Model 2 results are consistent with Model 1 in that a tournament effect is not sufficient in itself (t = 0.13 on T variable); however, in conjunction with a reduced horizon and restricted endowment, the tournament interacts to explain a significant part (t = 4.57 on I*E*T) of the bubbles in these experiments.

In Model 3, we observe the impact of speculative traders vis-á-vis conservatives by introducing a measure of asset purchase activity. The end-of-period asset holdings for the speculative group (the top one-half of traders in risk ratings) are compared to the total asset endowment for all traders. In the absence of any effect, assets should be evenly divided, and this ratio, A, should be equal to.5. The results of Model 3 indicate that speculators independently do not impact the presence of a bubble (t = 0.01 for A); however, when speculators operate within a shortened horizon (I*A), they do significantly differentiate themselves from conservatives by buying more and contributing to positive price bubbles. Finally, the impact of the use of the trader's own money is shown not to significantly alter the effects of the price bubbles (t = −0.54 for $). The R2 of .80 suggests that the vast majority of price deviation from PFE levels can be explained by investment horizon, endowment effects, and risk aversion.

Panel B reports the results for experiments 11–14 where markets were composed of either all speculators (11, 12) or all conservatives (13, 14). Due to this makeup, variables A and I*A are not defined in these regressions although S and I*S are substituted in their place and represent the speculative markets (11 and 12) and the interaction of shortened horizon with a speculative market (12). In addition, a restricted endowment effect (E = 1) is imposed for experiments 11–14 since experiments 7–10 clearly established their necessity in creating bubbles. Model 4 results highlight the significant positive effect of the combined shortened horizon/restricted endowment effect (t = 6.98 for I). More importantly, the speculative group statistically differs from conservatives with an additional mean price difference of 89.9 (t = 4.92). Model 5 supports the results of experiments 1–10 in that 1) a shortened investment horizon with restricted endowments leads to price bubbles (t = 2.71) for I, 2) a heightened tournament incentive will heighten short-term horizons and lead to positive price effects (t = 2.83 for I*T), and 3) speculators contribute to positive price bubbles in restricted endowment environments (t = 3.17 for S).[28]

The visual analysis of experiments 15–18 (negative bubble experiments) is confirmed by the regression results reported in Table 73.5. The variables are as defined earlier under Eq. 73.5 albeit the EN representing a dummy variable for the negative endowment effect. EN = 1 when the initial endowment equals 10 securities and 1,000 firms and 0 otherwise. In addition, since the shortened horizon variable I occurs for all years except 1A of each experiment, I and E are highly correlated. The design is therefore set to only measure the interaction effects of a shortened horizon and endowment. The four periods (1A of each experiment) are the control periods where a shortened horizon is not present (dummy NI = 1 for not I).

The parameter estimates of Model 6 show significant positive results for both positive and negative bubbles. The joint presence of a shortened horizon induced by a tournament payoff along with a buy side endowment (2 securities, 10,000 firms), that is, I*E = 1, leads to an average increase of 400.6 francs in price levels. The single alteration of the endowment to sell side (10 securities, 1,000 francs) in the presence of a tournament leads to an average decrease in price of 201.3 francs. The estimate for NI reflects the insignificant impact of the control periods where the endowment effect is present but without the tournament payoff inducing a shortened horizon. So as in the earlier results, the combined effect of the incentive

[28]Furthermore, although insignificant, the p-value for I*S is equal to.14, suggesting that the speculative difference may be even greater under a shortened investment horizon.

Table 73.5 Negative bubble and single period results

Model	Intercept	NI	$I^{*}E$	$I^{*}E^{N}$	R^2
Panel A					
Experiments 15–18 (n = 20)					
6	NOINT[b]	45.0 (0.37)	400.6*** (4.67)	−201.3** (−2.34)	.55
Panel B					
Experiments 19–22 (n = 40)					
7	NOINT[b]	−22.8 (−1.60)	22.5** (2.36)	−92.9***(−6.89)	.57
Panel C					
Experiments 23–26 (n = 20)					
8	−130.0** (−3.61)		174.5*** (3.95)	−208.0*** (−4.08)	.80

This table shows the extent to which endowment in conjunction with other variables causes a deviation from perfect foresight equilibrium values[a]. The following regression is estimated separately for experiments 15–18, 19–22, and 23–26 according to the experimental design of Table 73.1.

P^{L}-PFE = *f*(.NI; E; I*E; EN; I*E^{N})

[a]t-values in parentheses. Variables defined as follows:

PL-PFE represents the deviation from equilibrium for Period A of each trading year where P^{L} is the last trade of the period and PFE is the Perfect Foresight Equilibrium price

I is a dummy variable representing the shortened Investment horizon according to Table 73.1. I = 1 for shortened horizon, and 0 otherwise

E is a dummy variable representing the Endowment effect according to Table 73.1. E = 1 when 2 securities are issued, and 0 otherwise

I*E is an interaction dummy variable representing both a shortened investment horizon and two security endowment

E^{N} is a dummy variable representing the sell side of the Endowment effect hypothesized to lead Negative bubbles. E = 1 when 10 securities are issued and 0 otherwise

I*E^{N} is an interaction dummy variable representing both a shortened investment horizon and a ten security environment

[b]NOINT means the regression was run by suppressing the intercept.

*, **, ***Signify statistical significance levels at .10, .05, and .01, respectively.

(i.e., the tournament) and the ability (i.e., the endowment) works to create both positive and negative price bubbles.

Appendix 2: Additional Tests

To check the robustness of our results, we conduct eight final experiments in a unique and different setting, the former Communist country of Albania.[29] Of its many unique characteristics, one of the most important is its history of being the most isolated (politically and economically) country in Europe since World War II. Since democratic reforms opened in 1991, a new business school was opened in the second largest city of Albania, Shkodra, where the third year students served as traders. Would the students whose country didn't have a securities market

[29]In addition, we perform further OLS regressions and report the results in Table 73.5.

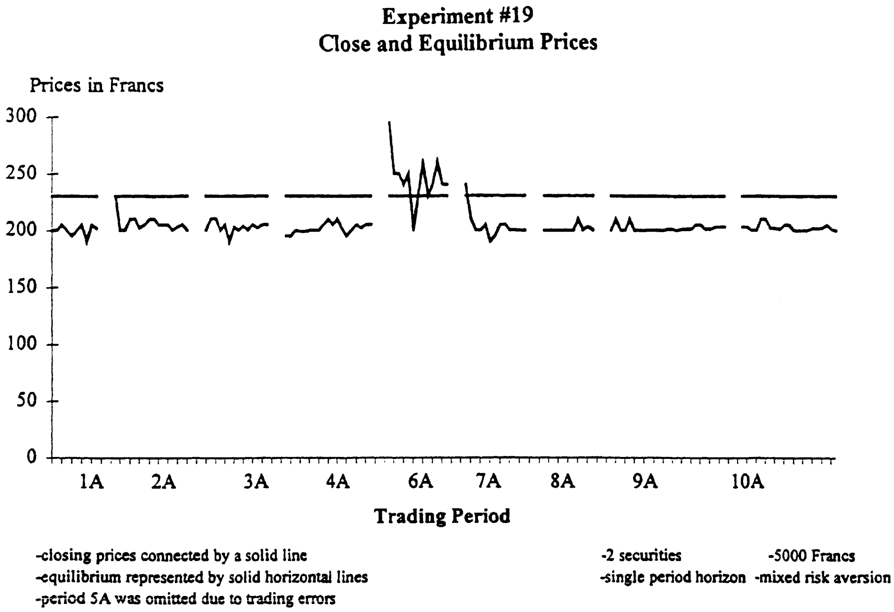


Fig. 73.19 Experiment 19 bid ask, close and equilibrium prices, one dividend with 5 trading periods, mixed risk aversion. Subjects are from Albania

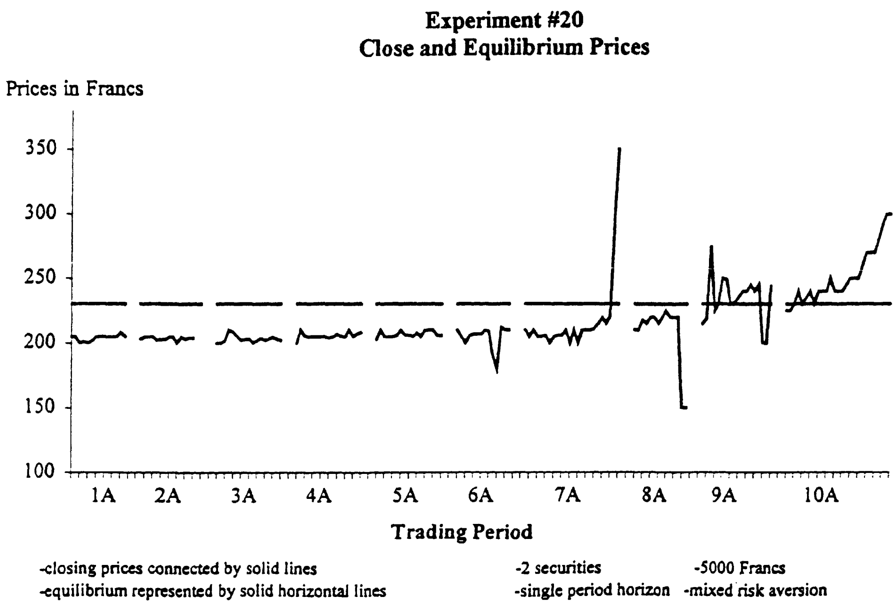


Fig. 73.20 Experiment 20 bid ask, close and equilibrium prices, one dividend with 5 trading periods, mixed risk aversion, and bonus to top traders. Subjects are from Albania

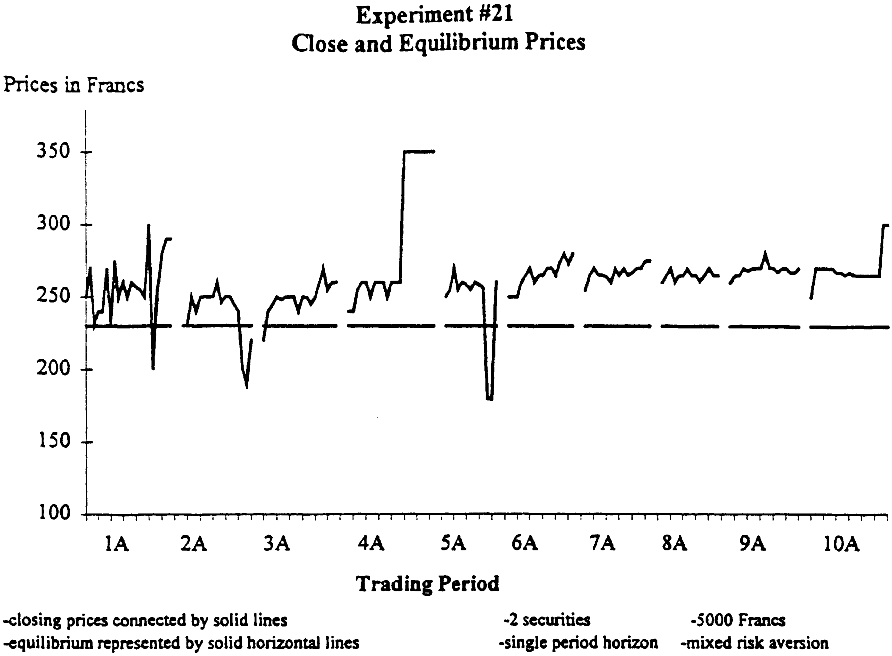


Fig. 73.21 Experiment 21 bid ask, close and equilibrium prices, one dividend with 5 trading periods, mixed risk aversion and bonus to top traders. Subjects are from Albania

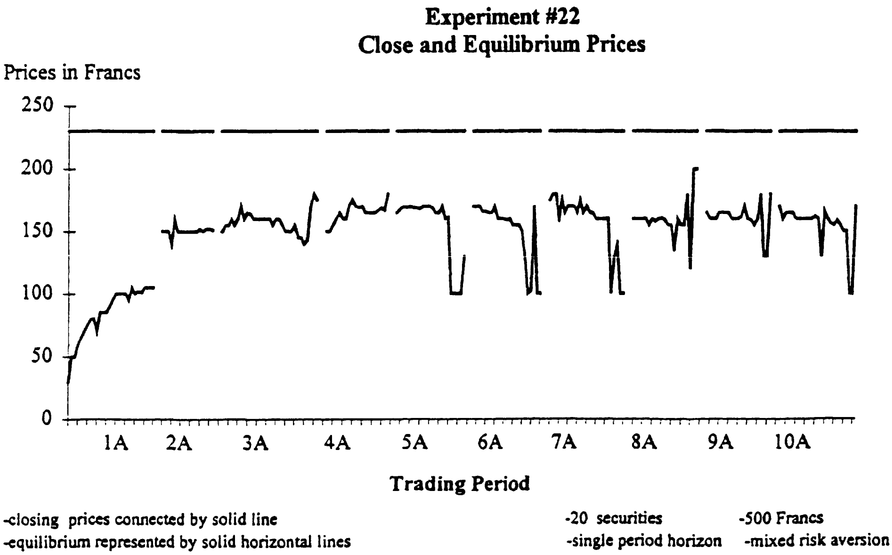


Fig. 73.22 Experiment 22 bid ask, close and equilibrium prices, two periods, mixed risk aversion, and bonus to top traders. Increase supply of shares. Subjects are from Albania

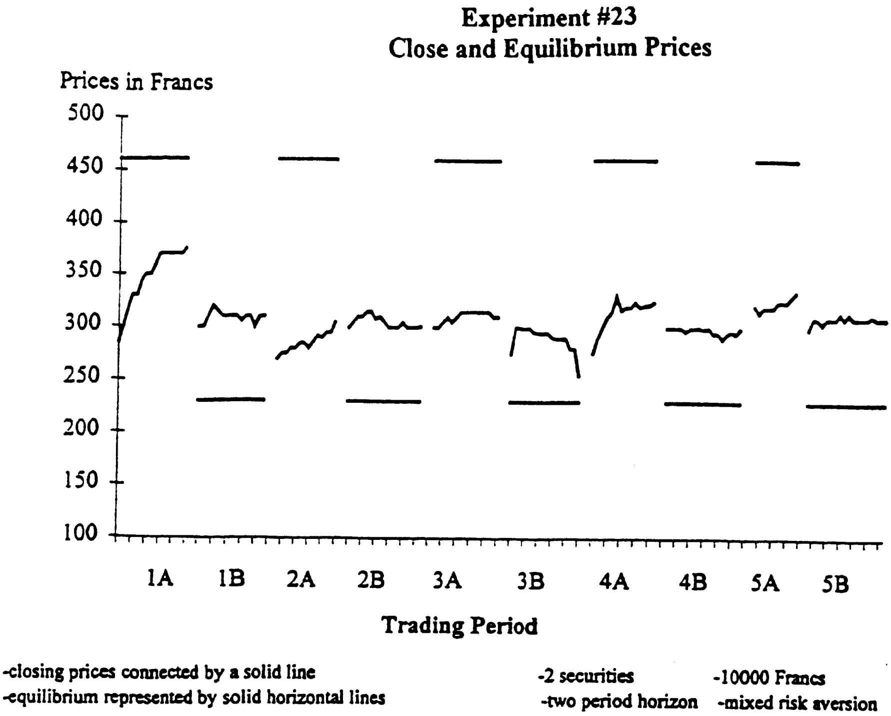


Fig. 73.23 Experiment 23 bid ask, close and equilibrium prices, two dividends with five trading periods each, and mixed risk aversion. Subjects are from Albania

or a history of free market trade show the same results as we had found at US universities? While our previous experiments had the most experienced traders ever used in a study, these may indeed represent the least experienced traders examined to date which may be regarded as an extreme test of the validity of our results.

Because of the newness of the trading experience for these students, a single-period design was used in the first four experiments. For each experiment's ten trading years (no period B), asset payoffs were for a single dividend payoff. The amounts used were the same as those of Table 73.2 so that equilibrium levels remained at 230 for each year. As shown in Table 73.1, Design 7 (experiment 19) consists of a single-period security without a tournament effect. Design 8 (experiments 20 and 21) introduces the tournament payoff of Table 73.3 (Schedule Two) within the single-period environment. This allows us to test for the presence of bubbles in the simpler pricing environment while also easing the learning experience of the Albanian students toward two-period tournament pricing.

The pricing results for these three experiments can be seen in Figs. 73.19–73.21. Without the tournament in experiment 19, pricing is rational and typical showing a discount (risk premium) of about 30 francs from the equilibrium level of 230. Near the end of experiment 20, the tournament effect appears to have created some

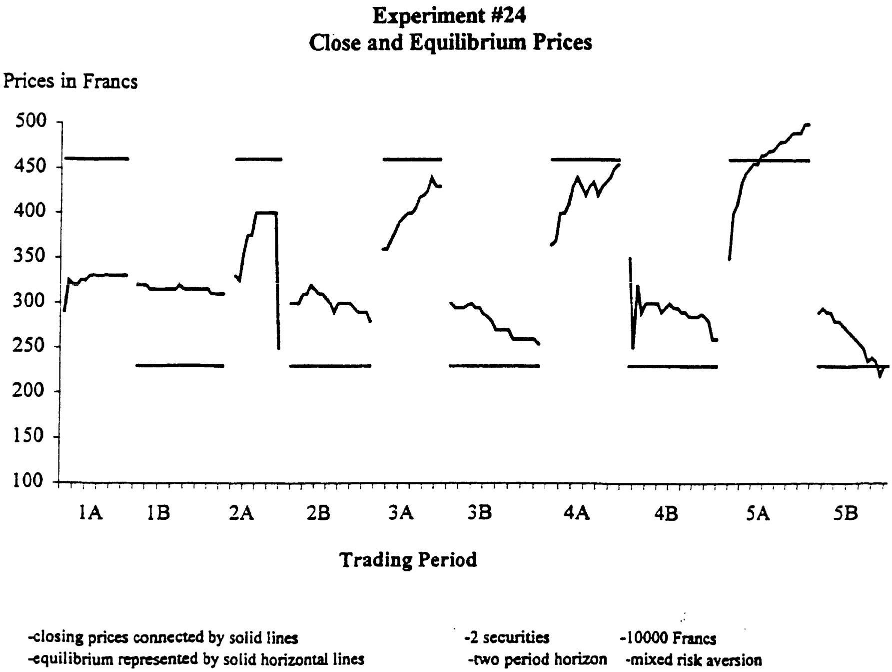


Fig. 73.24 Experiment 24 bid ask, close and equilibrium prices, one dividend with 5 trading periods, and mixed risk aversion. Subjects are from Albania

price movement above equilibrium. This pressure continues into experiment 21 where prices trade at an average premium of 30 francs. While these premiums do not constitute a bubble, it is clear they had a significant positive impact on pricing levels. Would this effect be eliminated (reversed) by changing the buy/sell pressure as was done earlier under Design 6 where negative bubbles were induced? Design 9 tests this proposition by changing the endowment from 2 securities and 5,000 francs to 20 securities and 500 francs. The results, reported in Fig. 73.22, show that even in these simple markets, the endowment effect combined with tournament payoff leads to pricing away from equilibrium. These observations are confirmed by the regression results of Model 7 in Table 73.5 where buy side preference (I*E = 1) leads to significant increase in prices, while sell side preference (I*EN = 1) leads to lower prices. The absence of a tournament payoff (NI = 1) leads to insignificant price effects as investment horizon cannot be altered in a single-period market.

The Albanian students had now participated in four single-period experiments and were ready to attempt two-period pricing. Experiment 23 was a simple two-period pricing environment without any tournament payoff as in Design 1 (control). The plot of prices in Fig. 73.23 shows that the students initially struggled with two-period pricing since period A prices (two payoffs) differed little

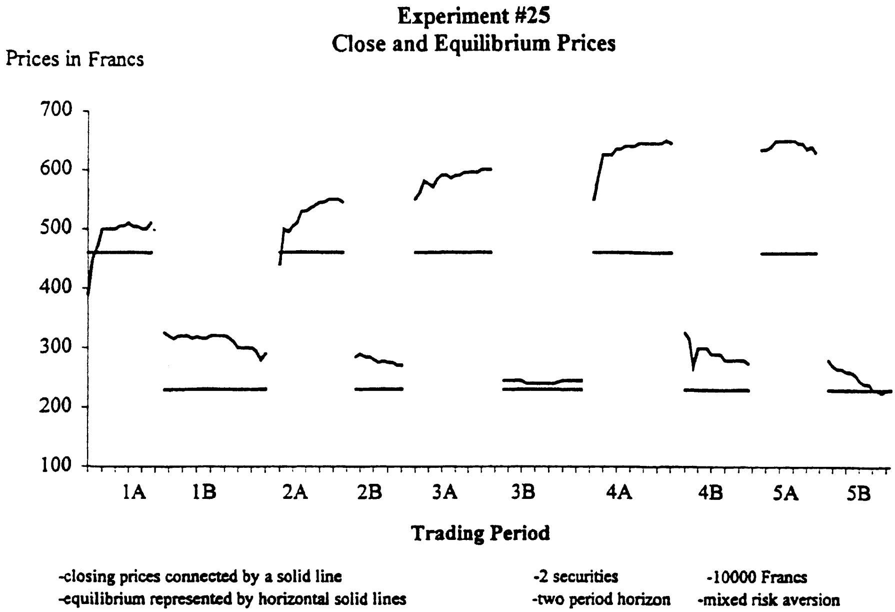


Fig. 73.25 Experiment 25 bid ask, close and equilibrium prices, one dividend with 5 trading periods, mixed risk aversion. Subjects are from Albania

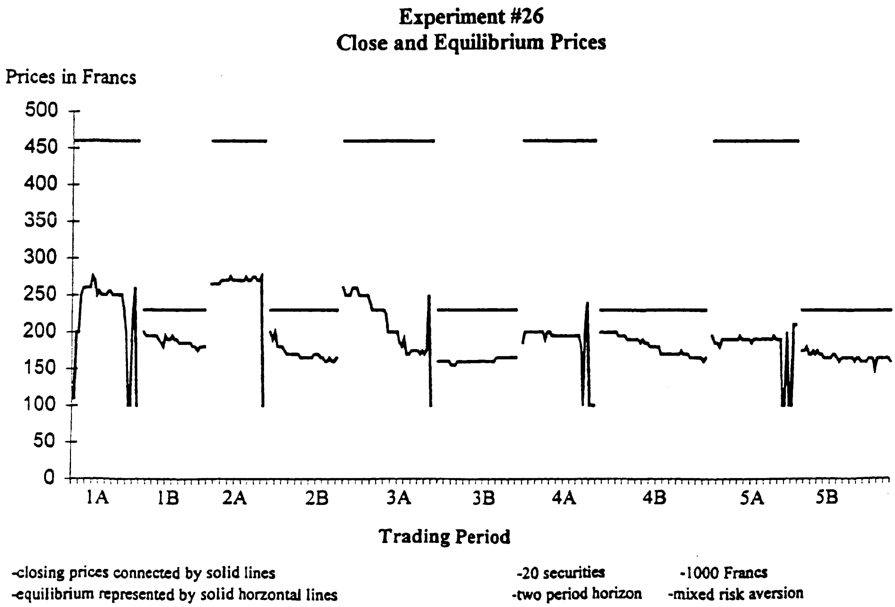


Fig. 73.26 Experiment 26 bid ask, close and equilibrium prices, one dividend with 5 trading periods, and mixed risk aversion. Increase supply of shares. Subjects are from Albania

from period B prices (single payoff), though by the end of the experiment enough learning had developed.

Experiments 24 and 25 introduce the shortened investment horizon (tournament effect) within the two-period framework as in Design 2 earlier. The price patterns in Figs. 73.24 and 73.25 show the creation of positive price bubbles to levels approaching 650 francs. Despite the historical background of this country and these students, they responded to market pressures in the same bubble-like manner. The last experiment, 26, alters the endowment to the sell side as before to see if negative bubbles can also be obtained. Price paths in Fig. 73.26 show a general downward trend of prices. The prices in period A show significant and growing discounts from the equilibrium levels of 460. These observations are confirmed by the regression results reported in Panel C of Table 73.5 with buy side endowment contributing 174.5 francs and sell side endowment reducing levels by 208.0 francs.

References

Ackert, L. N., Charupat, B. C., & Deaves, R. (2002). Bubbles in experimental asset markets: Irrational exuberance no more. Working paper 2002-24.

Ackert, L. N., Charupat, R. D., & Kluger, B. (2006). The origins of bubbles in laboratory asset markets. Federal Reserve Bank of Atlanta. Working paper 2006-6.

Allen, F., & Gale, D. (2000). Bubbles and crises. *Economic Journal, 110*, 236–255.

Allen, F., & Gorton, G. (1988). Rational finite bubbles. In L. Rodney (Ed)., *White center for financial research*. Working paper no. 41–88.

Ang, J., & Schwarz, T. (1985). Risk aversion and information structure: An experimental study of price variability in the securities market. *Journal of Finance, 40*, 924–844.

Ang, J. S., Davitopolous, D., & Schwarz, T. (2009). The creation and control of speculative bubbles in a laboratory setting. In A. Lee & C. F. Lee (Eds.), *Handbook of quantitative finance and risk management* (pp. 137–164). New York: Springer.

Baker, G. P. (1992). Incentive contracts and performance measurement. *Journal of Political Economy, 100*, 598–614.

Becker, B. E., & Huselid, M. A. (1992). The incentive effects of tournament compensation systems. *Administrative Science Quarterly, 37*, 336–350.

Bierman, H., Jr. (1995). Bubbles, theory and market timing. *Journal of Portfolio Management, 22*, 54–56.

Caginalp, G., Porter, D., & Smith, V. (2001). Financial bubbles: Excess cash, momentum, and incomplete information. *The Journal of Psychology and Financial Markets, 2*(2), 80–99.

Camerer, C. (1989). Bubbles and fads in asset price: A review of theory and evidence. *Journal of Economic Surveys, 3*, 3–41.

Cason, T., & Noussair, C. (2001). The experimental study of market behavior. *Advances in experimental markets* (pp.1–14). Springer.

Cutler, D. M., Poterba, J. M., & Summers, L. H. (1991). Speculative dynamics. *Review of Economic Studies, 58*, 529–546.

DeLong, J., Bradford, A., Schleifer, L., Summers, H., & Waldmann, R. T. (1989). The size and incidence of the losses from noise trading. *Journal of Finance, 44*, 681–696.

Diba, B. T., & Grossman, H. (1985). Rational bubbles in stock prices? NBER Working paper no. 1779.

Diba, B. T., & Grossman, H. (1987). On the inception of rational bubbles. *Quarterly Journal of Economics, 102*, 197–700.

Dufwenberg, M., Lindqvist, T., & Moore, E. (2005). Bubbles and experience: An experiment. *The American Economic Review, 95*, 1731–1737.
Duncan, J., & Isaac, R. M. (2000). Asset markets: How they are affected by tournament incentives for individuals. *American Economic Review, 90*(4), 995–1004. American Economic Association.
Durand, R., Newby, R., & Sanghani, J. (2006). An intimate portrait of the individual investor. Working paper.
Ehrenberg, R. G., & Bognanno, M. L. (1990). Do tournaments have incentive effects? *Journal of Political Economy, 98*, 1307–1324.
Flood, R. P., & Hodrick, R. J. (1990). On testing for speculative bubbles. *Journal of Economic Perspectives, 4*, 85–102.
Forsythe, R., Palfrey, T., & Plott, C. (1982). Asset valuation in an experimental model. *Econometrica, 50*, 537–563.
Friedman, B. M. (1992). Monetary economics and financial markets: Retrospect and prospect. NBER Reporter. (pp. 1–5).
Froot, K. A., Scharfstein, D. S., & Stein, J. C. (1992). Herd on the street: Informational inefficiencies in a market with short-term speculation. *Journal of Finance, 47*, 1461–1484.
Garber, P. (1990). Famous first bubbles. *Journal of Economic Perspectives, 4*, 35–54.
Greene, W. H. (2012). Econometric analysis (7th ed.). Prentice Hall.
Gurkaynak, R. S. (2005). Econometric tests of asset price bubbles: Taking stock. REDS Working paper no. 2005-04.
Harrison, M., & Kreps, D. (1978). Speculative investor behavior in a stock market with heterogeneous expectations. *Quarterly Journal of Economics, 92*, 323–336.
Hart, O. D., & Kreps, D. M. (1986). Price destabilizing speculation. *Journal of Political Economy, 94*, 927–953.
Hayashi, F. (2000). Econometrics. Princeton University Press.
Heinemann, F. (2003). Risk aversion and incentive effects: Comment. Working paper (Ludwig-Maximilians-Universitat Munchen).
Hirota, S., & Sunder, S. (2005). Price bubbles sans dividend anchors: Evidence from laboratory stock markets. Working paper.
Holt, C. A., & Laury, S. K. (2002). Risk aversion and incentive effects. *American Economic Review, 92*, 1644–1655.
Hong, H. G., Scheinkman, J. A., & Xiong, W. (2006). Asset float and speculative bubbles. *Journal of Finance, 61*, 1073–1117.
Jackson, D. (1976). *Jackson personality inventory manual*. Goshen: Research Psychologists Press.
Jackson, D. (1977). Reliability of the Jackson personality inventory. *Psychological Reports, 40*, 613–614.
Jackson, D., Hourany, D., & Vidmar, N. (1972). A four dimensional interpretation of risk taking. *Journal of Personality, 40*, 433–501.
King, R. R., Smith, V. L., Williams, A., & Van Boening, M. (1990). *The robustness of bubbles and crashes*. Seattle: University of Washington. Working paper.
Lei, V., Noussair, C., & Plott, C. (2001). Nonspeculative bubbles in experimental asset markets: Lack of common knowledge of rationality vs actual irrationality. *Econometrica, 69*(4), 831–859.
Noussair, C. N., & Tucker, S. (2006). Futures markets and bubble formation in experimental asset markets. *Pacific Economic Review, 11*, 167–184.
Plott, C., & Sunder, S. (1982). Efficiency of experimental markets with insider information: An application of rational expectations models. *Journal of Political Economy, 90*, 663–698.
Porter, D. P., & Smith, V. L. (2003). Stock market bubbles in the laboratory. *Journal of Behavioral Finance, 4*(1), 7–20.
Ricke, M. (2004). What is the link between margin loans and stock market bubbles? University of Muenster, Department of Banking No. 03-01.
Robin, S., & Ruffieux, B. (2001). Price bubbles in laboratory asset markets with constant fundamental values. *Experimental Economics, 4*, 87–105.

Roll, R. (1989). Price volatility, international market links, and their implications for regulatory policy. *Journal of Financial Services Research, 3*, 211–246.

Scheinkman, J. A., & Xiong, W. (2003). Overconfidence and speculative bubbles. *Journal of Political Economy, 111*, 1183–1219.

Shiller, R. (1984). Stock prices and social dynamics. *Brookings Papers on Economic Activity, 15*, 457–498.

Shiller, R. J. (1988). Fashions, fads, and bubbles in financial markets. In J. Coffee Jr., L. Lowenstein, & S. Ross-Ackerman (Eds.), *Knights, raiders, and targets*. New York: Oxford Press.

Shleifer, A., & Vishny, R. W. (1990). Equilibrium short horizons of investors and firms. *American Economic Review, 80*, 148–153.

Siegel, J. J. (2003). What is an asset price bubble? An operational definition. *European Financial Management, 9*, 11–24.

Smith, V., Suchanek, G., & Williams, A. (1988). Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica, 56*, 1119–1151.

Stiglitz, J. (1990). Symposium on bubbles. *Journal of Economic Perspective, 4*, 13–18.

Sunder, S. (1992). Experimental asset markets: A survey. In J. H. Kagel & A. E. Roth (Eds.), *Handbook of experimental economics*. Princeton: Princeton University Press.

Tirole, J. (1982). On the possibility of speculation under rational expectations. *Econometrica, 50*, 1163–1187.

Voth, H. -J., & Temin, P. (2003). Riding the South Sea Bubble. MIT Economics Working paper no. 04-02.

West, K. (1987). A specification test for speculative bubbles. *Quarterly Journal of Economics, 102*, 553–580.

West, K. (1988). Bubbles, fads, and stock price volatility tests: A partial evaluation. *Journal of Finance, 43*, 639–655.

White, E. (1990). The stock market boom and crash of 1929 revisited. *Journal of Economic Perspectives, 4*, 67–84.

Wooldridge, J. M. (2010). Econometric analysis of cross section and panel data. Second edition. The MIT Press.

Range Volatility: A Review of Models and Empirical Studies

74

Ray Yeutien Chou, Hengchih Chou, and Nathan Liu

Contents

R.Y. Chou (✉)
Institute of Economics, Academia Sinica and National Chiao Tung University, Taipei, Taiwan
e-mail: rchou@econ.sinica.edu.tw

H. Chou
Department of Shipping and Transportation Management, National Taiwan Ocean University, Keelung, Taiwan
e-mail: hcchou@mail.ntou.edu.tw

N. Liu
Department of Finance, Feng Chia University, Taichung, Taiwan
e-mail: nathanliu@mail.fcu.edu.tw

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_74,

Abstract

The literature on range volatility modeling has been rapidly expanding due to its importance and applications. This chapter provides alternative price range estimators and discusses their empirical properties and limitations. Besides, we review some relevant financial applications for range volatility, such as value-at-risk estimation, hedge, spillover effect, portfolio management, and microstructure issues.

In this chapter, we survey the significant development of range-based volatility models, beginning with the simple random walk model up to the conditional autoregressive range (CARR) model. For the extension to range-based multivariate volatilities, some approaches developed recently are adopted, such as the dynamic conditional correlation (DCC) model, the double smooth transition conditional correlation (DSTCC) GARCH model, and the copula method. At last, we introduce different approaches to build bias-adjusted realized range to obtain a more efficient estimator.

Keywords

Range • Volatility forecasting • Dynamic conditional correlation • Smooth transition • Copula • Realized volatility • Risk management

74.1 Introduction

Financial volatility is a key input in derivative pricing, asset allocation, investment decisions, hedging, and risk analysis; volatility modeling thus has became an important task in financial markets, and it has held the attention of academics and practitioners over the last three decades. Nevertheless, following Barndorff-Nielsen and Shephard (2005) or Andersen et al. (2003), financial volatility is a latent factor and hence it cannot be observed directly. Financial volatility thus can only be estimated using its signature on certain known market price processes; when the underlying process is more sophisticated or when observed market prices suffer from market microstructure noise effects, the results are less clear.

It is well known that the time series of asset prices usually exhibit volatility clustering or autocorrelation. In incorporating the characteristics into the dynamic process, the generalized autoregressive conditional heteroskedasticity (GARCH) family of models proposed by Engle (1982) and Bollerslev (1986) and the stochastic volatility (SV) models advocated by Taylor (1986) are two popular and useful alternatives for estimating and modeling time-varying conditional financial volatility.

However, as pointed by Alizadeh et al. (2002), Brandt and Diebold (2006), Chou (2005), and others, both GARCH and SV models are inaccurate and inefficient, because they are based on the closing prices of the reference period, failing to use the information content inside the reference. In other words, the path of the price inside the reference period is totally ignored when volatility is estimated by these models. Especially in turbulent days with drops and recoveries in the markets, the traditional close-to-close volatility indicates a low level, while the daily price range shows correctly that the volatility is high.

The price range, also known as high/low range or range volatility, is basically defined as the difference between the highest and lowest market prices over a fixed sampling interval. The price range has been known for a long time and has recently experienced renewed interest as a proxy of the latent volatility. The information contained in the opening, highest, lowest, and closing prices of an asset is widely used in Japanese candlestick charting techniques and other technical analysis indicators, such as the directional movement indicator (DMI). Early applications of range in the field of finance can be traced to Mandelbrot (1971) and the academic work on the range-based volatility estimator which began in the early 1980s. Several authors, back to Parkinson (1980), developed several volatility measures which were far more efficient than the classical return-based volatility estimators.

Building on the earlier results of Parkinson (1980), many studies[1] showed that one can use the price range information to improve volatility estimation. In addition to being significantly more efficient than the squared daily return, Alizadeh et al. (2002) also demonstrated that the conditional distribution of the *log range* is approximately Gaussian, thus greatly facilitating maximum likelihood estimation of stochastic volatility models. Moreover, as pointed out by Alizadeh et al. (2002) and Brandt and Diebold (2006), the range-based volatility estimator appears robust to microstructure noise such as bid-ask bounce. By adding microstructure noise to the Monte Carlo simulation, Shu and Zhang (2006) also supported the finding of Alizadeh et al. (2002) that range estimators are fairly robust toward microstructure effects.

Cox and Rubinstein (1985) explained the problem that despite the elegant theory and the support of simulation results, the range-based volatility estimator has performed poorly in empirical studies. Chou (2005) argued that the failure of all the range-based models in the literature is caused by their ignorance of the temporal movements of price range. Using a proper dynamic structure for the conditional expectation of range, the conditional autoregressive range (CARR) model, proposed by Chou (2005), successfully resolves this puzzle and retains its superiority in empirical forecasting abilities. The in-sample and out-of-sample volatility forecasting using S&P 500 index data shows that the CARR model does provide more accurate volatility estimator compared with the GARCH model. Similarly,

[1]See Garman and Klass (1980), Beckers(1983), Ball and Torous (1984), Wiggins (1991), Rogers and Satchell (1991), Kunitomo (1992), Yang and Zhang (2000), Alizadeh et al. (2002), Brandt and Diebold (2006), Brandt and Jones (2006), Chou (2005, 2006), Cheung (2007), Martens and van Dijk (2007), Chou and Wang (2007), Floros (2009), Chou et al. (2009), and Chou and Liu (2010, 2011).

Brandt and Jones (2006) formulated a model that is analogous to Nelson's (1991) EGARCH model but uses the square root of the intraday price range in place of the absolute return. Both studies find that the range-based volatility estimators offer a significant improvement over their return-based counterparts. Moreover, Chou et al. (2009) extended CARR to a multivariate context using the dynamic conditional correlation (DCC) model proposed by Engle (2002a). They found that this range-based DCC model performs better than other return-based volatility models in forecasting covariances. In this chapter, we also review alternative range-based multivariate volatility models in Sect. 74.3.

Recently, many studies have used high-frequency data to get an unbiased and highly efficient estimator for measuring volatility; see Andersen et al. (2003) and McAleer and Medeiros (2008) for a review. The volatility built by nonparametric methods is called *realized volatility*, which is calculated by the sum of nonoverlapping squared returns within a fixed time interval. Martens and van Dijk (2007) replaced the squared return with the price range to get a more efficient estimator, namely, the realized range. In their empirical study, the realized range was a significant improvement over realized return volatility. In addition, Christensen and Podolskij (2007) independently develop the *realized range* and showed that this estimator is consistent and relatively efficient under some specific assumptions.

The remainder of the chapter is laid out as follows. Section 74.2 introduces the price range estimators. Section 74.3 describes the range-based volatility models, including univariate and multivariate ones. Section 74.4 presents the realized range. The financial applications of range volatility are provided in Sect. 74.5. Finally, the conclusion is showed in Sect. 74.6.

74.2 The Price Range Estimators

A few price range estimators and their estimation efficiency are briefly introduced and discussed in this section. The price ranges which can be calculated by the daily opening, highest, lowest and closing prices are readily available for many assets. Most data suppliers provide daily highest/lowest prices as summaries of intraday activity. For example, Datastream records the intraday price range for most securities, including equities, currencies, and commodities, going back to 1955. Thus, range-based volatility proxies are easily calculated. When using this record, the additional information yields a great improvement when used in financial applications. Roughly speaking, knowing these records allows us to get closer to the real underlying process, even if we do not know the whole path of asset prices. For an asset, let's define the following variables:

$O_t =$ the opening price of the tth trading day.
$C_t =$ the closing price of the tth trading day.
$H_t =$ the highest price of the tth trading day.
$L_t =$ the lowest price of the tth trading day.

The efficiency for the Parkinson (1980) estimator intuitively comes from the fact that the price range of intraday trading gives more information regarding the future volatility than two arbitrary points in this series (the closing prices). Assuming that the asset price follows a simple diffusion model without a drift term, his estimator $\hat{\sigma}_P^2$ can be written as follows:

$$\hat{\sigma}_P^2 = \frac{1}{4\ln 2}(\ln H_t - \ln L_t)^2. \tag{74.1}$$

But instead of using two data points, the highest and lowest prices, four data points, the opening, closing, highest, and lowest prices, might also give extra information. Garman and Klass (1980) proposed several volatility estimators based on the knowledge of the opening, closing, highest, and lowest prices. Like Parkinson (1980), they assumed the same diffusion process and proposed their estimator $\hat{\sigma}_{GS}^2$ as

$$\begin{aligned}\hat{\sigma}_{GK}^2 = {} & 0.511[\ln(H_t/L_t)]^2 - 0.019\{\ln(C_t/O_t)[\ln(H_t) + \ln(L_t) - 2\ln(O_t)] \\ & -2[\ln(H_t/O_t)\ln(L_t/O_t)]\} - 0.383[\ln(C_t/O_t)]^2.\end{aligned} \tag{74.2}$$

As mentioned in Garman and Klass (1980), their estimator can be presented practically as $\hat{\sigma}_{GK'}^2 = 0.5[\ln(H_t/L_t)]^2 - [2\ln 2 - 1][\ln(C_t/O_t)]^2$. Molnár (2012) showed that in the absence of high-frequency data, returns normalized by their estimator are, approximately, distributed normally.

The price path cannot be monitored when markets are closed; however, Wiggins (1991) found that both the Parkinson estimator and Garman-Klass estimator were still biased downward compared to the traditional estimator, because the observed highs and lows were smaller than the actual highs and lows. Garman and Klass (1980) and Grammatikos and Saunders (1986), nevertheless, estimated the potential bias using simulation analysis and showed that the bias decreases with an increasing number of transactions. Therefore, it is relatively easy to adjust the estimates of daily variances to eliminate the source of bias.

Because the Parkinson (1980) and Garman and Klass (1980) estimators implicitly assumed that log-price follows a geometric Brownian motion with *no* drift term, further refinements were made by Rogers and Satchell (1991) and Kunitomo (1992). Rogers and Satchell (1991) added a drift term in the stochastic process that could be incorporated into a volatility estimator using only daily opening, highest, lowest, and closing prices. Their estimator $\hat{\sigma}_{RS}^2$ can be written as follows:

$$\begin{aligned}\hat{\sigma}_{RS}^2 = \frac{1}{N}\sum_{n=t-N}^{t} & \ln(H_n/O_n)[\ln(H_n/O_n) - \ln(C_n/O_n)] \\ & + \ln(L_n/O_n)[\ln(L_n/O_n) - \ln(C_n/O_n)].\end{aligned} \tag{74.3}$$

Rogers et al. (1994) reported that the Rogers-Satchell estimator yields theoretical efficiency gains compared to the Garman-Klass estimator. They also reported that the Rogers-Satchell estimator appears to perform well when changing drift with as few as 30 daily observations.

Different from Rogers and Satchell (1991), Kunitomo (1992) used the opening and closing prices to estimate a *modified* range corresponding to a hypothesis of a Brownian bridge of the transformed log-price. This basically tries to correct the highest and lowest prices for the drift term:

$$\hat{\sigma}_K^2 = \frac{1}{\beta_N} \sum_{p=t-N}^{t} \left[\ln\left(\hat{H}_n / \hat{L}_n \right) \right], \tag{74.4}$$

where. two estimators $\hat{H}_n = \underset{t_i}{\text{Arg}} \left[\underset{P_{t_i}}{\text{Max}} \{ P_{t_i} - [O_n + (C_n - O_n)/t_i] + (C_n - O_n) \mid t_i \in [n-1, n] \} \right]$ and $\hat{L}_n = \underset{t_i}{\text{Arg}} \left[\underset{P_{t_i}}{\text{Min}} \{ P_{t_i} - [O_n + (C_n - O_n)/t_i] + (C_n - O_n) \mid t_i \in [n-1, n] \} \right]$ are denoted as the end-of-the-day drift correction highest and lowest prices. $\beta_N = 6/(N\pi^2)$ is a correction parameter.

Finally, Yang and Zhang (2000) made further refinements by deriving a price range estimator that is unbiased, independent of any drift, and consistent in the presence of opening price jumps. Their estimator $\hat{\sigma}_{YZ}^2$ thus can be written as follows:

$$\begin{aligned} \hat{\sigma}_{YZ}^2 = {} & \frac{1}{(N-1)} \sum_{n=t-N}^{t} \left[\ln(O_n/C_{n-1}) - \overline{\ln(O_n/C_{n-1})} \right] \\ & + \frac{k}{(N-1)} \sum_{n=t-N}^{t} \left[\ln(O_n/C_{n-1}) - \overline{\ln(O_n/C_{n-1})} \right] + (1-k)\, \hat{\sigma}_{RS}^2, \end{aligned} \tag{74.5}$$

where $k = \frac{0.34}{1.34 + (N+1)/(N-1)}$. The symbol $\overline{X}$ is the unconditional mean of X, and σ_{RS}^2 is the Rogers-Satchell estimator. The Yang-Zhang estimator is simply the sum of the estimated overnight variance, the estimated opening market variance, and the Rogers and Satchell (1991) drift-independent estimator. The resulting estimator therefore explicitly incorporates a term for the closed market variance.

Shu and Zhang (2006) investigated the relative performance of the four range-based volatility estimators including Parkinson, Garman-Klass, Rogers-Satchell, and Yang-Zhang estimators for S&P 500 index data and found that the price range estimators all perform very well when an asset price follows a continuous geometric Brownian motion. However, significant differences among the various range estimators are detected if the asset return distribution involves an opening jump or a large drift.

In terms of efficiency, all previous estimators exhibit substantial improvements. Defining the efficiency measure of a volatility estimator $\hat{\sigma}_i^2$ as the variance estimation compared with the close-close estimator, $\hat{\sigma}^2$, that is,

$$Eff\left(\hat{\sigma}_i^2\right) = \frac{Var\left(\hat{\sigma}^2\right)}{Var\left(\hat{\sigma}_i^2\right).} \tag{74.6}$$

Parkinson (1980) reported a theoretical relative efficiency gain ranging from 2.5 to 5, which means that the estimation variance is 2.5–5 times lower. Garman and Klass (1980) reported that their estimator has an efficiency of 7.4; while the Yang and Zhang (2000) and Kunitomo (1992) variance estimators resulted in a theoretical efficiency gain of 7.3 and 10, respectively.

In addition to the variance estimation, Rogers and Zhou (2008) proposed a new estimator for the correlation based on the opening, closing, high, and low prices of two asset prices. However, they concluded that the range-based estimator of correlation does not perform better than the simpler estimator based only on the opening and closing prices. Nevertheless, it still points to new possibilities for future research.

74.3 The Range-Based Volatility Models

This section provides a brief overview of the models used to forecast range-based volatility. In what follows, the models are presented in increasing order of complexity. For an asset, the range of the log-prices is defined as the difference between the daily highest and lowest prices in a logarithm type. It can be denoted by

$$R_t = \ln(H_t) - \ln(L_t). \tag{74.7}$$

According to the Christoffersen's (2002) result applied to the S&P 500 data, the range-based volatility R_t showed more persistence than the squared return based on estimated autocorrelations. Thus, the range-based volatility estimator of course could be used instead of the squared return for evaluating the forecasts from volatility models, and with the time series of R_t, one can easily construct a volatility model under the traditional autoregressive framework.

Instead of using the data of range, nevertheless, Alizadeh et al. (2002) focused on the variable of the log range, $\ln(R_t)$, since they found that in many applied situations, the log range follows an approximately normal distribution. Therefore, all the models introduced in the section except for Chou's CARR model are estimated and forecasted using the log range.

The following range-based volatility models were first introduced with some simple specifications, including random walk, moving average (MA), exponentially weighting moving average (EWMA), and autoregressive (AR) models.

Hanke and Wichern (2005) thought that these models were fairly basic techniques in the applied forecasting literature. Additionally, we also provide some models with a much higher degree of complexity, such as the stochastic volatility (SV), CARR, and range-based multivariate volatility models.

74.3.1 The Random Walk Model

The log range $\ln(R_t)$ can be viewed as a random walk. It means that the best forecast of the next period's log range is this period's estimate of log range. As in most papers, the random walk model is used as the benchmark for the purpose of comparison.

$$E[\ln(R_{t+1})|I_t] = \ln(R_t), \tag{74.8}$$

where I_t is the information set at time t. The estimator $E[\ln(R_{t+1})|I_t]$ is obtained conditional on I_t.

74.3.2 The MA Model

MA methods are widely used in time series forecasting. In most cases, a moving average of length N where $N = 20$, 60, 120 days is used to generate log range forecasts. Choosing these lengths is fairly standard because these values of N correspond to 1 month, 3 months, and 6 months of trading days, respectively. The expression for the N day moving average is shown below:

$$E[\ln(R_{t+1})|I_t] = \frac{1}{N}\sum_{j=0}^{N-1} \ln\left(\mathrm{R}_{t-j}\right). \tag{74.9}$$

74.3.3 The EWMA Model

EWMA models are also very widely used in applied forecasting. In EWMA models, the current forecast of log range is calculated as the weighted average of the one period past value of log range and the one period past forecast of log range. This specification appropriately provides the underlying log range series with no trend.

$$E[\ln(\mathrm{R}_{t+1})|I_t] = \lambda E[\ln(R_t)|I_{t-1}] + (1-\lambda)\ln(\mathrm{R}_t). \tag{74.10}$$

The smoothing parameter, λ, lies between zero and unity. If λ is zero then the EWMA model is the same as a random walk. If λ is one, then the EWMA model places all of the weight on the past forecast. In the estimation process the optimal value of λ was chosen based on the root mean squared error criteria. The optimal λ is the one that records the lowest MSE.

Like the above-mentioned EWMA model, Harris and Yilmaz (2010) combined the Parkinson range estimator and the open-to-close return to propose a hybrid EWMA variance model.

$$\hat{\sigma}_{t+1}^{\text{Hybrid}} = \lambda \hat{\sigma}_{t}^{\text{Hybrid}} + (1-\lambda)\hat{\sigma}'_{P,t}, \tag{74.11}$$

where $\hat{\sigma}'_P = \frac{1}{4\ln 2}(\ln H_t - \ln L_t)^2 + (O_t - C_{t-1})^2$.

74.3.4 The AR Model

This model uses an autoregressive process to model log range. It combines the dynamic volatility with the range information. There are n lagged values of past log range to be used as drivers to forecast one period ahead.

$$E[\ln(\mathrm{R}_{t+1})] = \beta_0 + \beta_i \sum_{i=1}^{n} \ln(\mathrm{R}_{t+1-i}). \tag{74.12}$$

Li and Hong (2011) introduced the range-based autoregressive volatility (AV) model which was first proposed by Hsieh (1991, 1993). Their empirical study showed that the range-based AV model performs better than the GARCH model in the in-sample and out-of-sample comparisons.

74.3.5 The Discrete-Time Range-Based SV Model

Alizadeh et al. (2002) presented a formal derivation of the discrete-time SV model from the continuous-time SV model. The conditional distribution of log range is approximately Gaussian:

$$\ln R_{t+1} \big| \ln R_t \sim N\left[\ln \overline{R} + \rho\left(\ln R_{t-1} - \ln \overline{R}\right), \beta^2 \Delta t\right], \tag{74.13}$$

where $\Delta t = T/N$, T is the sample period, and N is the number of intervals. The parameter β models the volatility of the latent volatility. Following Harvey et al. (1994), a linear state space system including the state equation and the signal equation can be written as

$$\ln R_{(i+1)\Delta t} = \ln \overline{R} + \rho_{\Delta t}\left(\ln R_{i\Delta t} - \ln \overline{R}\right) + \beta\sqrt{\Delta t}\upsilon_{(i+1)\Delta t}. \tag{74.14}$$

$$\ln\left|f\left(s_{i\Delta t,\,(i+1)\Delta t}\right)\right| = \gamma \ln R_{i\Delta t} + E\left[\ln\left|f\left(s^{*}_{i\Delta t,\,(i+1)\Delta t}\right)\right|\right] + \varepsilon_{(i+1)\Delta t}. \tag{74.15}$$

Equation 74.14 is the state equation and Eq. 74.15 is the signal equation. In Eq. 74.15, E is the mathematical expectation operator. The state equation errors are i.i.d. $N(0,1)$ and the signal equation errors have a mean of zero.

A two-factor model can be represented by the following state equation:

$$\begin{aligned} \ln R_{(i+1)\Delta t} &= \ln \overline{R} + \ln \overline{R}_{1,(i+1)\Delta t} + \ln \overline{R}_{2,(i+1)\Delta t}. \\ \ln R_{1,(i+1)\Delta t} &= \rho_{1,\Delta t} \ln R_{1,i\Delta t} + \beta_1 \sqrt{\Delta t}\upsilon_{1,(i+1)\Delta t}. \\ \ln R_{2,(i+1)\Delta t} &= \rho_{2,\Delta t} \ln R_{2,i\Delta t} + \beta_2 \sqrt{\Delta t}\upsilon_{2,(i+1)\Delta t}. \end{aligned} \tag{74.16}$$

The error terms υ_1 and υ_2 are contemporaneously and serially independent $N(0, 1)$ random variables. Compared with one-factor volatility model for currency future prices, the two-factor model shows more desirable regression diagnostics. Asai and Unite (2010) extended this model to capture the leverage and size effects, but their empirical result did not support Alizadeh et al. (2002) theory; on the contrary, they showed that the conditional distributions of the selected returns are non-normal.

74.3.6 The Range-Based EGARCH Model

Brandt and Jones (2006) incorporated the range information into the EGARCH model, named by the range-based EGARCH model. The model significantly improves both in-sample and out-of-sample volatility forecasts. The daily log range and log returns are defined as the followings:

$$\ln(R_t)\big|I_{t-1} \sim N\left(0.43 + \ln h_t, 0.29^2\right), \quad r_t\big|I_{t-1} \sim N\left(0, h_t^2\right), \tag{74.17}$$

where h_t is the conditional volatility of the daily log return r_t. Then, the range-based EGARCH for the daily volatility can be expressed by

$$\ln h_t - \ln h_{t-1} = \kappa(\theta - \ln h_{t-1}) + \phi X_{t-1}^R + \delta r_{t-1}/h_{t-1}, \tag{74.18}$$

where θ is denoted as the long-run mean of the volatility process and κ is denoted as the speed of mean reversion. The coefficient δ decides the asymmetric effect of lagged returns. The innovation

$$X_{t-1}^R = \frac{\ln(R_{t-1}) - 0.43 - \ln h_{t-1}}{0.29} \tag{74.19}$$

is defined as the standardized deviation of the log range from its expected value. It means ϕ is used to measure the sensitivity to the lagged log ranges. In short, the range-based EGARCH model replaces the innovation term with the standardized log range.

74.3.7 The CARR Model

This section provides a brief overview of the CARR model used to forecast range-based volatility. The CARR model is also a special case of the multiplicative error model (MEM) of Engle (2002b) extended from the GARCH approach. The MEM model is used to model a nonnegative valued process, such as trading volume, duration, realized volatility, and range.[2] The MEM model provides conditional expectations of the variables like the GARCH approach and avoids the effect of zeros as resorting to logs. It can be extended to a multivariate case through the use of copula functions (Cipollini et al. 2009).

Instead of modeling the log range in the previous parts of this section, Chou (2005) focused the process of the price range directly. With the time series data of price range R_t, Chou (2005) presented the CARR model of order (p, q) or CARR (p, q) as

$$\begin{aligned} R_t &= \lambda_t \varepsilon_t, \varepsilon_t \sim f(.), \\ \lambda_t &= \omega + \sum_{i=1}^{p} \alpha_i R_{t-i} + \sum_{j=1}^{q} \beta_j \lambda_{t-j}, \end{aligned} \tag{74.20}$$

where λ_t is the conditional mean of the range based on all information up to time t and the distribution of the disturbance term ε_t, or the normalized range, is assumed to have a density function $f(.)$ with a unit mean. Since ε_t is positively valued given that both the price range R_t and its expected value λ_t are positively valued, a natural choice for the distribution is the exponential distribution.

The equation of the conditional expectation of range can easily be extended to incorporate other explanatory variables, such as trading volume, time to maturity, and lagged return:

$$\lambda_t = \omega + \sum_{i=1}^{p} \alpha_i R_{t-i} + \sum_{j=1}^{q} \beta_j \lambda_{t-j} + \sum_{k=1}^{L} l_k X_k. \tag{74.21}$$

This model is called the CARR model with exogenous variables, or the CARRX model. The CARR model essentially belongs to a symmetric model. In order to describe the leverage effect of financial time series, Chou (2006) divided the whole price range into two single-side price ranges, upward range and downward range. Further, he defined UPR_t, the upward range, and DNR_t, the downward range, as the differences between the daily highs, daily lows, and the opening price, respectively, at time t. This can be expressed as follows:

$$UPR_t = \ln(H_t) - \ln(O_t), \tag{74.22}$$

$$DNR_t = \ln(O_t) - \ln(L_t). \tag{74.23}$$

[2]Please refer to Engle and Russell (1998), Engle and Gallo (2006), and Manganelli (2005).

Similarly, with the time series of single-side price range, UPR_t or DNR_t, Chou (2006) extended the CARR model to the asymmetric CARR (ACARR) model. In volatility forecasting, the asymmetric model also performed better than the symmetric model. Chen et al. (2008) proposed a range-based threshold conditional autoregressive (TARR) model which has superior ability in volatility forecasting. In addition, Lin et al. (2012) proposed a nonlinear smooth transition CARR model to capture smooth volatility asymmetries in international financial markets.

74.3.8 The Range-Based Multivariate Volatility Model

The multivariate volatility models have been extensively researched in recent studies. They provide relevant financial applications in various areas, such as asset allocation, hedging, and risk management. Bauwens et al. (2006) offered a review of the multivariate volatility models. As to the extension of the univariate range models, Fernandes et al. (2005) proposed one kind of multivariate CARR (MCARR) model using the formula $\mathrm{Cov}(X, Y) = [V(X + Y) - V(X) - V(Y)]/2$. Moreover, Lee and Shin (2008) drove conditions for stationarity, geometric ergodicity, and β-mixing with exponential decay. Analogous to Fernandes et al. (2005)) work, Brandt and Diebold (2006) used no-arbitrage conditions to build the covariances in terms of variances. However, this kind of method could substantially apply to a bivariate case.

Chou et al. (2009) combined the CARR model with the DCC model of Engle (2002a) to propose a range-based volatility model, which uses the ranges to replace the GARCH volatilities in the first step of DCC. They concluded that the range-based DCC model performs better than other return-based models (MA100, EWMA, CCC, return-based DCC, and diagonal BEKK) through the statistical measures, RMSE and MAE, based on four benchmarks of implied and realized covariance.

The DCC model is a two-step forecasting model which estimates univariate GARCH models for each asset and then calculates its time-varying correlation by using the transformed standardized residuals from the first step. The related discussions about the DCC model can be found in Engle and Sheppard (2001), Engle (2002a), and Cappiello et al. (2006). It can be viewed as a generalization of the constant conditional correlation (CCC) model proposed by Bollerslev (1990). The conditional covariance matrix $\mathbf{H}_t$ of a $k \times 1$ return vector $\boldsymbol{r}_t$ in CCC ($\mathbf{r}_t|\Omega_{t-1} \sim N(0, \mathbf{H}_t)$) can be expressed as

$$\mathbf{H}_t = \mathbf{D}_t\mathbf{R}\mathbf{D}_t, \tag{74.24}$$

Where $\mathbf{D}_t$ a $k \times k$ diagonal matrix with time-varying standard deviations $\sqrt{h_{i,t}}$ of the ith return series from GARCH on the ith diagonal. $\mathbf{R}$ is a sample correlation matrix of $\mathbf{r}_t$.

The DCC is formulated in the following specification:

$$\begin{aligned}
&\mathbf{H}_t = \mathbf{D}_t\mathbf{R}_t\mathbf{D}_t, \\
&\mathbf{R}_t = diag\{\mathbf{Q}_t\}^{-1/2}\mathbf{Q}_t diag\{\mathbf{Q}_t\}^{-1/2}, \\
&\mathbf{Q}_t = \mathbf{S}\circ\left(\boldsymbol{\iota}\boldsymbol{\iota}' - \mathbf{A} - \mathbf{B}\right) + \mathbf{A}\circ\mathbf{Z}_{t-1}\mathbf{Z}_{t-1} + \mathbf{B}\circ\mathbf{Q}_{t-1}, \mathbf{Z}_t = \mathbf{D}_t^{-1}\times\mathbf{r}_t,
\end{aligned} \tag{74.25}$$

where $\boldsymbol{\iota}$ is a vector of ones and $\circ$ is the Hadamard product of two identically sized matrices which are computed simply by element multiplication. $\mathbf{Q}_t$ and $\mathbf{S}$ are, respectively, the conditional and unconditional covariance matrices of the standardized residual vector $\mathbf{Z}_t$ that came from GARCH. For the CARR case, the standardized residual vector $\mathbf{Z}_t^*$ is calculated from the adjusted conditional range.[3] $\mathbf{A}$ and $\mathbf{B}$ are estimated parameter matrices. Most cases, however, set them as scalars. In a word, DCC differs from CCC by only allowing $\mathbf{R}$ to be time varying.

It is difficult to introduce the exogenous variables into the DCC model because of the technical limitations for the mean reverting process. Chou and Cai (2009) proposed a double smooth transition conditional correlation CARR (DSTCC-CARR) model.[4] In addition to the multi-asset CARR part, the DSTCC-CARR model builds the smooth transition correlation structure through the standardized residuals $\mathbf{Z}_t^*$ of the rescaled range.

$$E\left[\mathbf{Z}_t^*\mathbf{Z}'^*_t|\Omega_{t-1}\right] = P_t, \tag{74.26}$$

$$\begin{aligned}
P_t = &(1-G_{2t})\big((1-G_{1t})P_{(11)} + G_{1t}P_{(21)}\big) \\
&+ G_{2t}\big((1-G_{1t})P_{(12)} + G_{1t}P_{(22)}\big),
\end{aligned} \tag{74.27}$$

where the transition logistic functions are $G_{jt} = \left(1 + e^{-\gamma_j(s_{jt}-c_j)}\right)^{-1}, \gamma_j > 0, j = 1, 2.$ The symbols c_j and γ_j in the transition function are location and speed parameters, respectively. Please see Chou and Cai (2009) for the details. Base on this framework, Cai et al. (2009) used CPI and VIX as transition variables to investigate the correlations among six international stock indices.

74.3.9 Other Model Extensions

In addition to the classification of range models, Harris et al. (2011) developed a cyclical volatility model which employs the range to investigate the short and long

---

[3]For asset i, $z_{i,t}^* = r_{i,t}/\lambda_{i,t}^*$, where $\lambda_{i,t}^* = adj_i \times \lambda_{i,t}$ and $adj_i = \frac{\overline{\sigma}_i}{\overline{\hat{\lambda}}_i}$. The scaled expected range $\lambda_{i,t}^*$ is computed by a product of $\lambda_{i,t}$ and the adjusted coefficient adj_i which is the ratio of the unconditional standard deviations $\overline{\sigma}_i$ for the return series to the sample mean $\overline{\hat{\lambda}}_i$ of the estimated conditional range.

[4]Silvennoien and Terasvirta (2008, 2009) proposed the smooth transition conditional correlation GARCH (STCC-GARCH) model and the double smooth transition conditional correlation GARCH (DSTCC-GARCH) model.

dynamics of exchange rate volatility. Their results indicated that the cyclical volatility model performed better than the range-based EGARCH and FIEGARCH models in computational efficiency and out-of-sample forecast. In contrast to modeling range directly, some studies put the range into existing models to increase the explanatory power. For example, Lin and Rozeff (1994) put the estimated range process into the GARCH model and showed that the range estimator was still useful in explaining the conditional variance.

74.4 The Realized Range Volatility

There has been much research investigating the measurement of volatility due to the use of high-frequency data. In particular, the realized volatility, calculated by the sum of squared intraday returns, provides a more efficient estimate for volatility. The review of realized volatility has been discussed in Andersen et al. (2001), Andersen et al. (2003), Barndorff-Nielsen and Shephard (2005), Andersen et al. (2006, 2007), and McAleer and Mederos (2008). Martens and van Dijk (2007) and Christensen and Podolskij (2007) replaced the squared intraday return with the high/low range to get a new estimator called realized range.

Initially, we assumed that the asset price P_t follows the geometric Brownian motion:

$$dP_t = \mu P_t dt + \sigma P_t dz_t, \tag{74.28}$$

where μ is the drift term, σ is the constant volatility, and z_t is a Brownian motion. There are τ equal-length intervals divided into a trading day. The daily realized volatility RV_t at time t can be expressed by

$$RV_t = \sum_{i=1}^{\tau} \left(\ln P_{t,i} - \ln P_{t,i-1}\right)^2, \tag{74.29}$$

where $P_{t,i}$ is the price for the time $i \times \Delta$ on the trading day t and Δ is the time interval. Then, $\tau \times \Delta$ is the trading time length in a trading day. Moreover, the realized range RR_t is

$$RR_t = \frac{1}{4\ln 2} \sum_{i=1}^{\tau} \left(\ln H_{t,i} - \ln L_{t,i-1}\right)^2, \tag{74.30}$$

where $H_{t,i}$ and $L_{t,i}$ are the highest price and the lowest price of the ith interval on the tth trading day, respectively.

As mentioned before, several studies suggest improving efficiency by using the open and close prices, like Garman and Klass (1980). Furthermore, assuming that P_t follows a continuous sample path, Martingale, Christensen, and Podolskij (2007) proposed integrated volatility and showed this range estimator remains consistent in the presence of stochastic volatility.

$$\ln P_t = \ln P_0 + \int_0^t \mu_s ds + \int_0^t \sigma_{s-} dz_t, \;\; for \;\; 0 \le t < \infty. \tag{74.31}$$

The obvious and important question is that the realized range should be seriously affected by microstructure noise. Martens and van Dijk (2007) considered a bias-adjustment procedure, which scales the realized range by using the ratio of the average level of the daily range and the average level of the realized range. Christensen et al. (2009) provided another bias correlation for the realized range to help divide the high-frequency data to minimize its asymptotic conditional variance. Both found that the scaled realized ranges perform better than the (scaled) realized volatility. Todorova (2012) also showed that the adjusted realized ranges perform better than the daily range for the DAX 30 index.

It is interesting to note that the realized range can be extended to estimate covariance. Bannouh et al. (2009) used the concept of Brandt and Diebold's (2006) non-arbitrage portfolio to propose a realized co-range estimator:

$$RCR_t = \frac{1}{2\lambda_1\lambda_2}\left(RR_{P,t} - \lambda_1^2 RR_{1,t} - \lambda_2^2 RR_{2,t}\right), \tag{74.32}$$

where $RR_{p,t}$, $RR_{1,t}$, and $RR_{2,t}$ are the realized ranges of the portfolio P, asset 1, and asset 2. λ_1 and λ_1 are the weights of two assets in the portfolio ($\lambda_1 + \lambda_2 = 1$).

74.5 The Financial Applications of Range Volatility

The range mentioned in this chapter is a measure of volatility. From the theoretical points of view, it indeed provides a more efficient estimator of volatility than the return. It is intuitively reasonable due to the further information provided by the range data. In addition, the return volatility neglects the price fluctuation, especially when existing a short distance between the closing prices of the two trading days. We can therefore conclude that the high/low range volatility should contain some additional information compared with the close-to-close volatility. Moreover, the range is readily available, which has low cost. Hence, most research related to volatility may be applied to the range. Bollerslev et al. (1992) and Poon and Granger (2003) provided extensive discussions on the application of volatilities in the financial markets.

Before the range was adapted by the dynamic structures, however, its application was very limited.[5] Based on the SV framework, Gallant et al. (1999) and Alizadeh, Brandt, and Diebold incorporated the range into the equilibrium asset pricing models. Chou (2005) and Brandt and Jones (2006), on the other hand, filled the

[5]In some studies, range was used as one of the estimators to improve the explanatory power of models.

gap between a discrete-time dynamic model and range. Their work creates many opportunities for future research. In the following sections, we will give a classified review for the financial applications of range volatility.

74.5.1 Value at Risk

Value at Risk (VaR) is designed to measure the potential loss of the asset. It is widely used in the financial markets. We can calculate it by $\mathrm{VaR}_t^{\alpha} = \mu_t + f_{\alpha}\sigma_t$, where α is the given significant level and f_{α} is the left quantile of the distribution F at α. Asai and Brugal (2012) used an asymmetric heterogeneous ARMA model to fit range for estimating VaR and showed that the one-step-ahead VaR forecast of the log range performed well during the global financial crisis. Chen et al. (2012) proposed a range-based threshold conditional VaR (CAViaR) model which also outperformed other models during the crisis. Moreover, Brownlees and Gallo (2010) showed that the daily range performed as well as the ultra-high-frequency data (UHFD) volatility measure they proposed for VaR prediction. In addition, Shao, Lian and Yin (2009) used the CARR model to model the realized range in estimating VaR, but it only performed the same with the realized volatility model. However, Louzis et al. (2012) found that the adjusted realized range can generate superior VaR estimates.

74.5.2 Hedge

With the development of conditional volatility models, there has been a dramatic increase in future hedging. From the calculation of minimum variance hedging, the optimal dynamic hedge ratio can be expressed as $h_t = \rho\sigma_{S,t}/\sigma_{F,t}$, where ρ is the correlation of spot and futures returns and $\sigma_{s,t}$ and $\sigma_{F,t}$ are the standard deviations of spot and futures returns, respectively. Within frameworks of a constant conditional correlation (CCC) model and a dynamic conditional correlation (DCC) model, Chou and Liu (2011) showed that the range-based multivariate volatility model has more efficiency gain than the return-based approaches.

74.5.3 Volatility Spillover

Volatility spillover can reflect the information flow among financial markets. Most studies analyze the behavior of volatility spillover by estimating the conditional variance and covariance. Gallo and Otranto (2008) estimated the weekly range through a new Markov Switching bivariate model to show the relevant role of Hong Kong as a dominant market. Engle et al. (2012) applied daily range to a multivariate MEM approach which is used to discuss the volatility transmission across East Asian markets. Chiang and Wang (2011) combined copula functions with a time-varying logarithmic CARR (TVLCARR) model to investigate the volatility contagion for the G7 stock markets.

74.5.4 Portfolio Management

Covariance process plays an important role in asset allocation. Based on the conditional mean-variance framework, Chou and Liu (2010) used Chou et al. (2009) method to show that the economic value of volatility timing for the range is significant in comparison to the return. Wu and Liang (2011) did similar work by incorporating dynamic copulas into an asymmetric CARR model. The results imply that the range volatility can be extended to practical applications.

74.5.5 Microstructure Issues

In recent years, there has been a shift in attention to high-frequency data for some financial assets. It must be noted that the microstructure analysis often accompanies different levels of microstructure noise.[6] Martens and van Dijk (2007) claimed that the realized range also cannot avoid the bias caused from the microstructure noise.[7] However, Akay et al. (2010) showed that the alternative range-based volatility estimates are relatively efficient and removed the upward bias caused by the microstructure noise. Kalev and Duong (2008) utilized Martens and van Dijk's (2007) realized range to test the Samuelson Hypothesis for the futures contract.[8]

74.5.6 Other Financial Applications

As mentioned above, range is available and can easily be applied to volatility issues. Chou et al. (2103) adopted the CARR model to investigate the long-term impact of terrorist attacks on the maturity, volume, and open interest effects for the S&P 500 index futures. Corrado and Truong (2007) reported that the range estimator has similar forecasting ability of volatility compared with the implied volatility. However, the implied volatilities are not available for many assets and the option markets are insufficient in many developed countries. In such cases, the range is more practical. Besides, range is often used as one of volatility proxies. Please refer to Liu and Hung (2010), Patton (2011), Liu et al. (2012), Chen and Wu (2009), Karanasos and Kartsaklas (2009), and Gallo and Otranto (2008).

In contrast to the range itself, some studies pay more attention to the high and low prices. Cheung et al. (2009) employed a vector error correlation model (VECM)

[6]In general, it mainly comes from bid-ask bounce and varies with the sampling frequency.

[7]For low-frequency data, Alizadeh et al. (2002) showed that the range estimator is efficient and free of microstructure noise.

[8]The Samuelson (time-to-delivery) effect means that volatility increases when a futures contract approaches its delivery date. Ripple and Moosa (2009) also used the realized range to test the effect of maturity, trading volume, and open interest on crude oil futures. In contrast, Karali and Thurman (2010) just used the daily range to prove the Samuelson effect.

to model the dynamic relationship between the high and low prices of stock indices. Please also see He and Wan (2009) and He et al. (2010) for the relevant applications.

74.6 Conclusion and Limitations

Volatility plays a central role in many areas of finance. In view of the theoretical and practical studies, the price range provides an intuitive and efficient estimator of volatility. In this study, we began our discussion by reviewing the price range estimators. There has been a dramatic increase in the number of publications on this work since Parkinson (1980) introduced the high/low range. From then on, some new range estimators have been considered with opening and closing prices. The new price range estimators are distributed feasible weights according to the differences among the highest, lowest, opening, and closing. Through the analysis, we can gain a better understanding of the nature of range.

Some dynamic volatility models combined with price range are also introduced in this chapter. They led to broad applications in finance, especially the CARR model, which incorporates both the superiority of range in forecasting volatility and the elasticity of the GARCH model. In addition, the range-based volatility models contribute significantly to the financial applications. Last, the realized range replaced the squared intraday return of realized volatility with the high/low range to obtain a more efficient estimator. Although the financial applications of range volatility are still in its infancy, the possible areas such as risk management, investment, and microstructure issues are explained in this chapter. Future studies are obviously required for this topic.

The range estimator undoubtedly has some inherent shortcomings. It is well known that the financial asset price is very volatile and is easily influenced by instantaneous information. In statistics, the range is very sensitive to the outliers. Chou (2005) provided an answer by using the quantile range to get a robust measure of price range. For example, the new range estimator can be calculated by the difference between the top and the bottom 5 % observations on average. Also see Yeh et al. (2009) for further discussion.

In theory, many range estimators in previous sections depended on the assumption of continuous-time geometric Brownian motion. The range estimators derived from Parkinson (1980) and Garman and Klass (1980) required a geometric Brownian motion with zero drift. Rogers and Satchell (1991) allowed a nonzero drift, and Yang and Zhang (2000) further allowed overnight price jumps. Moreover, only finite observations can be used to build the range. It means the range will appear with some unexpected bias, especially for the assets with lower liquidity and finite transaction volume. Garman and Klass (1980) pointed out that this will produce the later opening and early closing. They also said the difference between the observed highs and lows will be less than that between the actual highs and lows. It means that the calculated high/low estimator should be downward biased. In addition, Beckers (1983) pointed that disadvantaged buyers and sellers may trade the highest and lowest prices, so the range values might be less representative for measuring volatility. Because of the limitations involved and the importance of range volatility measure, range-based volatility modeling will continue to be a specialist subject and studied vigorously.

References

Akay, O., Griffiths, M. D., & Winters, D. B. (2010). On the robustness of range-based volatility estimators. *Journal of Financial Research, 33*, 179–199.

Alizadeh, S., Brandt, M., & Diebold, F. (2002). Range-based estimation of stochastic volatility models. *Journal of Finance, 57*, 1047–1091.

Andersen, T., Bollerslev, T., Diebold, F., & Ebens, H. (2001). The distribution of realized stock return volatility. *Journal of Financial Economics, 61*, 43–76.

Andersen, T., Bollerslev, T., Diebold, F., & Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica, 71*, 579–625.

Andersen, T., Bollerslev, T., Christoffersen, P., & Diebold, F. (2006). Volatility and correlation forecasting. In G. C. Elliott, W. J. Granger, & A. Timmermann (Eds.), *Handbook of economic forecasting* (pp. 778–878). Amsterdam: North-Holland.

Andersen, T., Bollerslev, T., Christoffersen, P., & Diebold, F. (2007). Practical volatility and correlation modeling for financial market risk management. In M. Carey & R. Stulz (Eds.), *The risks of financial institutions* (pp. 513–548). Chicago: University of Chicago Press for NBER.

Asai, M., & Brugal, I. (2012). Forecasting volatility using range data: Analysis for emerging equity markets in Latin America. *Applied Financial Economics, 22*, 461–470.

Asai, M., & Unite, A. (2010). General asymmetric stochastic volatility models using range data: Estimation and empirical evidence from emerging equity markets. *Applied Financial Economics, 20*, 1041–1049.

Ball, C., & Torous, W. (1984). The maximum likelihood estimation of security price volatility: Theory, evidence, and application to option pricing. *Journal of Business, 57*, 97–112.

Bannouh, K., van Dijk, D., & Martens, M. (2009). Range-based covariance estimation using high-frequency data: The realized co-range. *Journal of Financial Econometrics, 7*, 341–372.

Barndorff-Nielsen, O. E., & Shephard, N. (2005). How accurate is the asymptotic approximation to the distribution of realized variance. *Identification and inference for econometric models. A Festschrift in honour of TJ Rothenberg*, 306–311.

Bauwens, L., Laurent, S., & Rombouts, J. V. K. (2006). Multivariate GARCH models: A survey. *Journal of Applied Econometrics, 21*, 79–109.

Beckers, S. (1983). Variances of security price returns based on highest, lowest, and closing prices. *Journal of Business, 56*, 97–112.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics, 31*, 307–327.

Bollerslev, T. (1990). Modeling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. *Review of Economics and Statistics, 72*, 498–505.

Bollerslev, T., Chou, R. Y., & Kroner, K. (1992). ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics, 52*, 5–59.

Brandt, M., & Diebold, F. (2006). A no-arbitrage approach to range-based estimation of return covariances and correlations. *Journal of Business, 79*, 61–74.

Brandt, M., & Jones, C. (2006). Volatility forecasting with range-based EGARCH models. *Journal of Business and Economic Statistics, 24*, 470–486.

Brownlees, C. T., & Gallo, G. M. (2010). Comparison of volatility measures: A risk management perspective. *Journal of Financial Econometrics, 8*, 29–56.

Cai, Y., Chou, R. Y., & Li, D. (2009). Explaining international stock correlations with CPI fluctuations and market volatility. *Journal of Banking and Finance, 33*, 2026–2035.

Cappiello, L., Engle, R., & Sheppard, K. (2006). Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial Econometrics, 4*, 537–572.

Chen, C., & Wu, C. (2009). Small trades and volatility increases after stock splits. *International Review of Economics and Finance, 18*, 592–610.

Chen, C. W. S., Gerlach, R., & Lin, E. M. H. (2008). Volatility forecasting using threshold heteroskedastic models of the intra-day range. *Computational Statistics and Data Analysis, 52*, 2990–3010.

Chen, C. W. S., Gerlach, R., Hwang, B. B. K., & McAleer, M. (2012). Forecasting value-at-risk using nonlinear regression quantiles and the intra-day range. *International Journal of Forecasting, 28*, 557–574.

Cheung, Y. (2007). An empirical model of daily highs and lows. *International Journal of Finance and Economics, 12*, 1–20.

Cheung, Y., Cheung, Y., & Wan, A. (2009). A high–low model of daily stock price ranges. *Journal of Forecasting, 28*, 103–119.

Chiang, M., & Wang, L. (2011). Volatility contagion: A range-based volatility approach. *Journal of Econometrics, 165*, 175–189.

Chou, R. (2005). Forecasting financial volatilities with extreme values: The conditional autoregressive range (CARR) model. *Journal of Money, Credit, and Banking, 37*, 561–582.

Chou, R. (2006). Modeling the asymmetry of stock movements using price ranges. *Advances in Econometrics, 20A*, 231–257.

Chou, R. Y., & Cai, Y. (2009). Range-based multivariate volatility model with double smooth transition in conditional correlation. *Global Finance Journal, 20*, 137–152.

Chou, R. Y., & Liu, N. (2010). The economic value of volatility timing using a range-based volatility model. *Journal of Economic Dynamics and Control, 34*, 2288–2301.

Chou, R. Y., & Liu, N. (2011). Estimating time-varying hedge ratios with a range-based multivariate volatility model, Working paper, Academia Sinica.

Chou, H., & Wang, D. (2007). Forecasting volatility of UK stock market: A test of conditional autoregressive range (CARR) model. *International Research Journal of Finance and Economics, 10*, 7–13.

Chou, R. Y., Wu, C., & Liu, N. (2009). Forecasting time-varying covariance with a range-based dynamic conditional correlation model. *Review of Quantitative Finance and Accounting, 33*, 327–345.

Chou, H., Zaabar, R., & Wang, D. (2103). Measuring and testing the long-term impact of terrorist attacks on the US futures market. *Applied Economics, 45*, 225–238.

Christensen, K., & Podolskij, M. (2007). Realized range-based estimation of integrated variance. *Journal of Econometrics, 141*, 323–349.

Christensen, K. M., Christensen, K., Podolskij, M., & Vetter, M. (2009). Bias-correcting the realized range-based variance in the presence of market microstructure noise. *Finance and Stochastics, 13*, 239–268.

Christoffersen, P. F. (2002). *Elements of financial risk management*. San Diego: Academic.

Cipollini, F., Engle, R. F., & Gallo, G. M. (2009). A model for multivariate non-negative valued processes in financial econometrics, Working paper.

Corrado, C., & Truong, C. (2007). Forecasting stock index volatility: Comparing implied volatility and the intraday high-low price range. *Journal of Financial Research, 30*, 201–215.

Cox, J., & Rubinstein, M. (1985). *Options markets*. Englewood Cliffs: Prentice-Hall.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of the variance of United Kingdom inflation. *Econometrica, 50*, 987–1007.

Engle, R. (2002a). Dynamic conditional correlation: A simple class of multivariate GARCH models. *Journal of Business and Economic Statistics, 20*, 339–350.

Engle, R. (2002b). New frontiers for ARCH models. *Journal of Applied Econometrics, 17*, 425–446.

Engle, R., & Gallo, G. M. (2006). A multiple indicators model for volatility using intra-daily data. *Journal of Econometrics, 131*, 3–27.

Engle, R., & Russell, J. R. (1998). Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica, 66*, 1127–1162.

Engle, R., & Sheppard, K. (2001). Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH, NBER Working paper 8554.

Engle, R., Gallo, G. M., & Velucchi, M. (2012). Volatility spillovers in East Asian financial markets: A MEM-based approach. *Review of Economics and Statistics, 94*, 222–223.

Fernandes, M., Mota, B., & Rocha, G. (2005). A multivariate conditional autoregressive range model. *Economics Letters, 86*, 435–440.

Floros, C. (2009). Modelling volatility using high, low, open and closing prices: Evidence from four S&P indices. *International Research Journal of Finance and Economics, 28*, 198–206.

Gallant, R., Hsu, C., & Tauchen, G. (1999). Using daily range data to calibrate volatility diffusions and extracting integrated volatility. *Review of Economics and Statistics, 81*, 617–631.

Gallo, G. M., & Otranto, E. (2008). Volatility spillovers, interdependence and comovements: A Markov switching approach. *Computational Statistics and Data Analysis, 52*, 3011–3026.

Garman, M., & Klass, M. (1980). On the estimation of security price volatilities from historical data. *Journal of Business, 53*, 67–78.

Grammatikos, T., & Saunders, A. (1986). Futures price variability: A test of maturity and volume effects. *Journal of Business, 59*, 319–330.

Hanke, J., & Wichern, D. (2005). *Business forecasting* (8th ed.). New York: Prentice Hall.

Harris, R. D. F., & Yilmaz, F. (2010). Estimation of the conditional variance–covariance matrix of returns using the intraday range. *International Journal of Forecasting, 26*, 180–194.

Harris, R. D. F., Stoja, E., & Yilmaz, F. (2011). A cyclical model of exchange rate volatility. *Journal of Banking and Finance, 35*, 3055–3064.

Harvey, A., Ruiz, E., & Shephard, N. (1994). Multivariate stochastic variance models. *Review of Economic Studies, 61*, 247–264.

He, A., & Wan, A. (2009). Predicting daily highs and lows of exchange rates: A cointegration analysis. *Journal of Applied Statistics, 36*, 1191–1204.

He, A., Kwok, J., & Wan, A. (2010). An empirical model of daily highs and lows of West Texas intermediate crude oil prices. *Energy Economics, 32*, 1499–1506.

Hsieh, D. A. (1991). Chaos and nonlinear dynamics: Application to financial markets. *Journal of Finance, 46*, 1839–1877.

Hsieh, D. A. (1993). Implications of nonlinear dynamics for financial risk management. *Journal of Financial and Quantitative Analysis, 28*, 41–64.

Kalev, P. S., & Duong, H. N. (2008). A test of the Samuelson hypothesis using realized range. *Journal of Futures Markets, 28*, 680–696.

Karali, B., & Thurman, W. N. (2010). Components of grain futures price volatility. *Journal of Agricultural and Resource Economics, 35*, 167–182.

Karanasos, M., & Kartsaklas, A. (2009). Dual long-memory, structural breaks and the link between turnover and the range-based volatility. *Journal of Empirical Finance, 16*, 838–851.

Kunitomo, N. (1992). Improving the Parkinson method of estimating security price volatilities. *Journal of Business, 65*, 295–302.

Lee, O., & Shin, D. W. (2008). Geometric ergodicity and β-mixing property for a multivariate CARR model. *Economics Letters, 100*, 111–114.

Li, H., & Hong, Y. (2011). Financial volatility forecasting with range-based autoregressive volatility model. *Finance Research Letters, 8*, 69–76.

Lin, J., & Rozeff, M. S. (1994). Variance, return, and high-low price spreads. *Journal of Financial Research, 17*, 301–319.

Lin, E. M. H., Chen, C. W. S., & Gerlach, R. (2012). Forecasting volatility with asymmetric smooth transition dynamic range models. *International Journal of Forecasting, 28*, 384–399.

Liu, H., & Hung, J. (2010). Forecasting volatility and capturing downside risk of the Taiwanese futures markets under the financial tsunami. *Managerial Finance, 36*, 860–875.

Liu, H., Chiang, S., & Cheng, N. (2012). Forecasting the volatility of S&P depositary receipts using GARCH-type models under intraday range-based and return-based proxy measures. *International Review of Economics and Finance, 22*, 78–91.

Louzis, D. P., Xanthopoulos-Sisinis, S., & Refenes, A. N. (2012). The role of high frequency intra-daily data, daily range and implied volatility in multi-period value-at-risk forecasting, Working paper.

Mandelbrot, B. (1971). When can price be arbitraged efficiently? A limit to validity of the random walk and martingale models. *Review of Economics and Statistics, 53*, 225–236.

Manganelli, S. (2005). Duration, volume and volatility impact of trades. *Journal of Financial Markets, 8*, 377–399.

Martens, M., & van Dijk, D. (2007). Measuring volatility with the realized range. *Journal of Econometrics, 138*, 181–207.

McAleer, M., & Medeiros, M. (2008). Realized volatility: A review. *Econometric Reviews, 27*, 10–45.

Molnár, P. (2012). Properties of range-based volatility estimators. *International Review of Financial Analysis, 23*, 20–29.

Nelson, D. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica, 59*, 347–370.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business, 53*, 61–65.

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies. *Journal of Econometrics, 160*, 246–256.

Poon, S. H., & Granger, C. W. J. (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature, 41*, 478–639.

Ripple, R. D., & Moosa, I. A. (2009). The effect of maturity, trading volume, and open interest on crude oil futures price range-based volatility. *Global Finance Journal, 20*, 209–219.

Rogers, L., & Satchell, S. (1991). Estimating variance from high, low and closing prices. *Annals of Applied Probability, 1*, 504–512.

Rogers, L., & Zhou, F. (2008). Estimating correlation from high, low, opening and closing prices. *The Annals of Applied Probability, 18*, 813–823.

Rogers, L., Satchell, S., & Yoon, Y. (1994). Estimating the volatility of stock prices: A comparison of methods that use high and low prices. *Applied Financial Economics, 4*, 241–247.

Shao, X., Lian, Y., & Yin, L. (2009). Forecasting value-at-risk using high frequency data: The realized range model. *Global Finance Journal, 20*, 128–136.

Shu, J. H., & Zhang, J. E. (2006). Testing range estimators of historical volatility. *Journal of Futures Markets, 26*, 297–313.

Silvennoinen, A., & Teräsvirta, T. (2008). Multivariate autoregressive conditional heteroskedasticity with smooth transition in conditional correlations, SSE/EFI Working paper series in Economics and Finance.

Silvennoinen, A., & Teräsvirta, T. (2009). Modeling multivariate autoregressive conditional heteroskedasticity with the double smooth transition conditional correlations GARCH model. *Journal of Financial Econometrics, 7*, 373–411.

Taylor, S. J. (1986). *Modelling financial time series*. Chichester: Wiley.

Todorova, N. (2012). Volatility estimators based on daily price ranges versus the realized range. *Applied Financial Economics, 22*, 215–229.

Wiggins, J. (1991). Empirical tests of the bias and efficiency of the extreme-value variance estimator for common stocks. *Journal of Business, 64*, 417–432.

Wu, C., & Liang, S. (2011). The economic value of range-based covariance between stock and bond returns with dynamic copulas. *Journal of Empirical Finance, 18*, 711–727.

Yang, D., & Zhang, Q. (2000). Drift-independent volatility estimation based on high, low, open, and closing prices. *Journal of Business, 73*, 477–491.

Yeh, J.,Wang, J., & Kuan, C. (2009). A noise-robust estimator of volatility based on interquantile ranges, Working paper.

Business Models: Applications to Capital Budgeting, Equity Value, and Return Attribution

75

Thomas S. Y. Ho and Sang Bin Lee

Contents

Abstract

This chapter describes a business model in a contingent claim modeling framework. The model defines a "primitive firm" as the underlying risky asset of a firm. The firm's revenue is generated from a fixed capital asset and the firm incurs both fixed operating costs and variable costs. In this context, the shareholders hold a retention option (paying the fixed operating costs) on the core capital asset with a series of growth options on capital investments. In this framework of two interacting options, we derive the firm value.

The chapter then provides three applications of the business model. Firstly, the chapter determines the optimal capital budgeting decision in the presence of fixed operating costs and shows how the fixed operating cost should be

T.S.Y. Ho (✉)
Thomas Ho Company Ltd, New York, NY, USA
e-mail: Tom.ho@thomasho.com

S.B. Lee
Hanyang University, Seong-Dong-Ku, Seoul, Korea
e-mail: leesb@hanyang.ac.kr

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_75,

accounted by in an NPV calculation. Secondly, the chapter determines the values of equity value, the growth option, and the retention option as the building blocks of primitive firm value. Using a sample of firms, the chapter illustrates a method in comparing the equity values of firms in the same business sector. Thirdly, the chapter relates the change in revenue to the change in equity value, showing how the combined operating leverage and financial leverage may affect the firm valuation and risks.

Keywords

Bottom-up capital budgeting • Business model • Capital budgeting • Contingent claim model • Equity value • Financial leverage • Fixed operating cost • Gross return on investment (GRI) • Growth option • Market performance measure • NPV • Operating leverage • Relative value of equity • Retention option • Return attribution • Top-down capital budgeting • Wealth transfer

75.1 Introduction

A "business model" often simply describes "ways that a firm makes money." It is a general description of the business environment, forecasts of earnings, and the proposed business strategies, and it often lacks the rigorous specification of a financial model. Despite the ambiguity, business models are important to corporate finance, investments, portfolio management, and many aspects of financial businesses.

They provide a framework to determine a firm's value, to evaluate corporate strategies, and to distinguish one firm from another firm or one business sector to another. The prevalent use of business models cannot be understated. Yet, despite the tremendous growth in applications of financial modeling in capital markets, the use of financial principles in developing business models is largely unexplored.

An early example of financial modeling of a business is pioneered by Stoll (1976). He presents a business model of a market maker. Demsetz (1968) suggests that market makers are in the business of providing liquidity to a market and they are compensated by the market via their bid-ask spreads. Stoll then develops the optimal dealer's bid-ask prices within Demsetz's business environment and shows precisely how a trader should set their bid-ask quotes, their trading strategies relating to their inventory positions, and finally the profits to the traders under competition. In short, Stoll provides the business model of a trading firm, leading to the subsequent growth of the microstructure theory. The successful use of the "dealer's business model" in microstructure theory demonstrates the importance of insights gained in modeling a business. For example, Ho and Marcis (1984) extend the business model to incorporate the fixed operating costs of a market making firm to determine the equilibrium number of market makers in the AMEX market.

However, to date, few rigorous business models have been proposed in the corporate finance literature. There are some examples. Trigeorgis (1993) values projects as multiple real options on the underlying asset value. Botteron et al. (2003) use barrier options to model the flexibility in production and sales of multinational enterprises under exchange rate uncertainties. Brennan and Schwartz (1985) determine the growth model of a mining firm. Cortazar et al. (2001) develop a real option model for valuing natural resource exploration investments such as oil and copper when there is joint price and geological-technical uncertainty. Gilroy and Lukas (2006) formalize the choice of market entry strategy for an individual multinational enterprise from a real option perspective. Fontes (2008) addresses investment decisions in production systems by using real options. Sodal et al. (2008) value the option to switch between the dry bulk market and wet bulk market for a combination carrier. Villani (2008) combines the real option approach with the game theory to examine an interaction between two firms that invest in R&D. Wirl (2008) investigates optimal maintenance of equipment under uncertainty and the options of scrapping versus keeping the equipment as a backup while paying the keeping cost. These models explore the use of contingent claim models in various corporate financial decisions ranging from abandoning or increasing the mining capabilities to scrapping versus maintenance of equipment. In reality, real option approach is in three different corporate uses. There are a strategic way of thinking, an analytical valuation tool, and an organization-wide process for evaluation, monitoring, and managing capital investment according to Triantis and Borison (2001).

This chapter extends the real option literature to describe the business models in a more general context. The purpose of the chapter is twofold: firstly we propose the use of real option approach to describe a business and secondly we show how such a business model can be used in some applications.

Our model is a discrete time, multi-period, contingent claim model. We assume that a firm is subjected to a business risk. The revenues are generated from a capital asset. It has to incur a fixed operating cost, making a fixed payment continually (a perpetual payment) to stay in business, and has the options to invest in future projects. That is, the firm must pay an exercise price continually to retain the option in business and at the same time maintains the growth options of Myers (1984). Such retention options and the growth options cannot be separated.

The model is applicable to many business sectors, including the retail chain companies, airlines, software companies, and other businesses whose revenues are generated from a core capital investment and whose expense structure consists of both fixed operating costs and variable costs. The inputs to the business model can be drawn from the published financial statements and market data, and therefore, the model is empirically testable.

We then show how the business model can be used in three important areas in corporate finance: (1) to determine the optimal capital budgeting decision given a fixed operating cost, (2) to relative value firms in the same business sector with different business models, and (3) to relate the change of the firm's revenue to the change of the equity value.

The first application deals with what Myers (1984) describes as two approaches in capital budgeting, the discounted cash flow and strategic planning approaches, as "two cultures and one problem" in the valuation of a firm. The bottom-up method (the discounted cash flow approach) determines the net present value of a project, and the manager accepts the project if the net present value is positive and rejects it otherwise. The top-down method (the strategic planning approach) considers all future state-dependent investments simultaneously and determines the optimal investment strategies that maximize the firm value.

While the two approaches are related by the valuation of a firm, corporate finance literature has not described how they are related to each other explicitly. Specifically, in the presence of fixed operating costs, how should the "expenses" of a project, at the margin, be incorporated in the NPV calculation? We show that the standard one period model cannot describe the relationships between the cost of the project to the future inflow of the project and the outflows of the project cost as well as the firm's fixed operating costs. In this chapter, we show how the NPV method is related to the top-down method via the implied fixed-cost measure. Relating to this issue, McDonald (2006) argues that the discounted cash flow and the real option valuation should provide the same answer when the methods are used correctly. However, he further argues that to the extent that the managers who use the real option valuation have effectively adopted a different business model, there is a real and important difference between the discounted cash flow and the real option valuation.

The second application of a business model deals with measuring the impact of the growth option, the debt, and the fixed operating cost on the observed equity value. These results are illustrated by applying the model to a sample of retail chain companies.

The third application focuses on the relationship between the firm's revenues and the stock valuation. We show how the operating leverage and the financial leverage together affect the change in the equity value with a change in the revenue. The model provides a return attribution of the equity returns based on the firm's business model. This approach enables corporate managers to evaluate the impact of the firm's operating leverage and the financial leverage on the risk of the firm's earnings.

The chapter proceeds as follows. Section 75.2 provides the business model of a firm. Section 75.3 provides the numerical simulations of the capital budgeting problem, comparing the optimal capital budgeting decisions based on the bottom-up and top-down decisions. Section 75.4 provides the application of the business model to the equity value decomposition. Section 75.5 describes return attribution results using the business model. Finally, Sect. 75.6 contains the conclusions.

75.2 The Model Assumptions

Many retail chain stores must incur significant fixed operating costs in setting up the distribution system, producing or buying the products, managing the core business processes. At the same time, the retail chain store invests in new distribution centers, and each investment is a capital budgeting decision. Each product development is a capital budgeting decision, which should increase the firm value at the margin. These are some of many examples where capital budgeting decision on each project is part of the business model of the firm, which must include the management of a significant fixed operating cost.

The model uses the standard basic assumptions in real option literature. We assume a multi-period discrete time model where all agents make their decisions at specific regular intervals; the one period interest rate is R_F, a constant; the firm seeks to maximize the shareholders' wealth; and the market is efficient. We assume a frictionless market with no corporate taxes and personal taxes, and therefore, the capital structure is irrelevant to the maximization of shareholders value. The following assumptions describe the model of the firm of this chapter.

Assumption 1. The Business Risk of the Firm (GRI) In this model, unlike many standard real option models, we use the sales (or revenue) of the firm as the risk driver and not the operating profits, as commonly used. The sales represent the business risk of the firm, while the operating profits are affected by the business model of the firm. We assume that the firm is endowed with a capital asset (CA). For example, the capital asset can be a factory that produces goods and services resulting in sales. The sales generated by one unit of the capital asset are called the *gross return on investment* (*GRI*). GRI is the risk driver of the model, and the risk represents the uncertain demand for the products. Therefore, the sales are stochastic, given by the following equation.

$$\text{Sales} = \text{GRI} \times \text{CA} \tag{75.1}$$

When the GRI increases, the firm would increase its sales for the same capital asset. When there is a down turn in GRI, the sales would fall. Extending the model to multiple risk sources should provide a more realistic model but may obscure the basic insights that the model provides.

We assume that GRI follows a binomial lattice process that is lognormal with no drift. The upstate and downstate are given by a proportional increase of $\exp(-\sigma)$ with probability q and a proportional decrease of $\exp(-\sigma)$ with probability (1 – q). σ is the volatility assumed to be constant. The market probability q is chosen so that the expected value of GRI over one period is the observed GRI at the beginning of each step. That is, the risk follows a martingale process.

While GRI follows a recombining binomial process, we will use a non-recombining tree notation. Specifically, we let $n = 0, 1, 2,\ldots$, and for each time n, we let the index i denote the state variable $i = 0, 1, 2,\ldots, 2^n - 1$. Then at any node of the tree (n, i), the binomial upstate and downstate nodes of the following step are denoted by (n, 2i + 1) and (n, 2i), respectively. Then the martingale process is specified by the following equation:

$$\begin{aligned} &GRI(n,i) = q \times GRI(n+1, 2i+1) + (1-q) \times GRI(n+1, 2i), \\ &q = \frac{1 - e^{-\sigma}}{e^{\sigma} - e^{-\sigma}}, \text{where } \sigma \text{ is the volatility of GRI} \end{aligned} \tag{75.2}$$

for $n = 0, 1, 2,\ldots$ and $i = 0, 1, 2,\ldots, 2^n - 1$

Let ρ be the cost of capital for the business risk, which is the required rate of return for that business risk, GRI. Note that the standard cost of capital of a firm reflects the risk of the firm's free cash flows taking the operating leverage into account, not the firm's sales risk as we do here. Since the firm risk is the same at each node, the cost of capital ρ is constant in all states and time periods on the lattice.

Assumption 2. The Primitive Firm (V_p) To apply the contingent claim valuation approach to value the firm, we begin with the definition of the primitive firm as the "underlying security." The primitive firm is a simple corporate entity which has no debt or claims other than the common shares, which are publicly traded.

The firm has one unit capital asset. The capital asset does not depreciate, and the value does not change. We can think of the capital assets as the distribution centers of a retail chain store. Let m be the gross profit margin. For simplicity, in this section, we assume that there are no costs associated in generating the sales, and that m equals unity. And therefore, the firm's sales are the profits, which are distributed to all the shareholders. Equation 75.2 presents the sales risk, and the $GRI(n,i)$ and ρ are the sales and cost of capital of the primitive firm at each node (n,i) on the lattice.

Note that the sales (and hence the profits) are always positive, because GRI follows a multiplicative process. By the definition of the cost of capital, the primitive firm value at each node point on the binomial lattice is

$$V_P(n,i) = \frac{GRI(n,i) \times m \times CA}{\rho}, \text{where m} = 1 \tag{75.3}$$

for $n = 0, 1, 2, \ldots$ and $i = 0, 1, 2, \ldots, 2^n - 1$.

Given the binomial process of the primitive firm, which we will use as the "underlying security," we can derive the risk-neutral probabilities, $p(n, i)$, at time n and state i. The derivation is given in Appendix 1.

$$p = \frac{A - e^{-\sigma}}{e^{\sigma} - e^{-\sigma}}, \text{where } A = \frac{1 + R_F}{1 + \rho} \tag{75.4}$$

When the cost of capital equals the risk-free rate and when the volatility σ is sufficiently small, the risk-neutral probability is approximately 0.5. That is, when σ is small, the upward movement is approximately the same as the downward movement, and therefore, the expected value with the binomial probability of 0.5 shows that the GRI must follow a martingale process. When the cost of capital is high relative to the risk-free rate, the risk-neutral probability would assign a lower weight to the upward movement, according to Eq. 75.4, to balance the use of the risk-free rate, a lower rate than the cost of capital, for discounting the future value. The use of the risk-neutral probability ensures the valuation method is consistent with that of the market valuation of Eq. 75.3.

Note that as long as the volatility and the cost of capital are independent of the time n and state i, the risk-neutral probability is also independent of the state and time and is the same at each node point on the binomial lattice. We will value our firm relative to the primitive firm. Therefore, using the standard relative valuation argument, we may assume that the primitive firm follows a drift at the risk-free rate. The market probability q is relevant only to the extent of determining the cost of capital ρ, but q is not used explicitly in the model.

The use of the risk-neutral probabilities enables us to discount all cash flows of our firm by the risk-free rate in all states of the world. The primitive firm value V_p specifies the stochastic process of the "underlying security," and Eq. 75.4 is the standard assumption made in the contingent claim valuation model.

Assumption 3. The Firm's Cash Flows (CF) and Value (V) *V* is the value of a firm that has fixed operating costs, fixed expenditures for operating purposes. The fixed operating cost (FC) is independent of the units of the goods sold and is paid at the end of each period. Payments to the vendors and suppliers and the employees' salaries and benefits are some examples of the fixed operating costs, and they may constitute a significant part of the firm's cash outflow.

The net profit of the firm, using Eq. 75.1, is given by

$$CF(n,i) = (GRI(n,i) \times CA(n,i) - FC) \tag{75.5}$$

for n = 0, 1, 2, . . . and i = 0, 1, 2,. . ., $2^n - 1$

Note that GRI is the only source of risk to the firm's net income. The firm pays all the net income as dividends. In the case of negative income, the firm issues equity to finance the short fall of cash for simplicity. Therefore, the firm's net income is the free cash flows, and the present value of which is the firm value V.

Assumption 4. The Planning Horizon (T) and the Terminal Conditions We assume that there is a strategic planning time horizon T. We will value the firm at each node at planning horizon T. Conditional on the firm not defaulted before reaching the horizon T, we can determine the firm value at time T.

Without the loss of generality, we make some simplifying assumptions at the terminal date. In this model, we assume that all future fixed operating cost is capitalized at time T to be a constant FC(T). After the horizon date T, the firm may default on the fixed operating cost. And therefore, the capitalized value of the fixed operating cost should depend on the primitive firm value at time T. The value of this capitalized value is also a contingent claim. The value, based on Merton (1973), is provided in Appendix 2 and result will be used later. For clarity of the presentation at this point, we keep the model simple without affecting the main results. Therefore, at the terminal date T, the firm value is given by

$$V(T,i) = Max\begin{bmatrix} V_p - FC(T) + (GRI(T,i) \times CA(T,i) - FC), \\ V_p - FC(T) + (GRI(T,i) \times (CA(T,i)) + I) - FC - I), 0 \end{bmatrix} \tag{75.6}$$

where $V_p = \frac{GRI(T,i) \cdot CA(T,i)}{\rho}$ for i = 0, 1, 2,. . ., 2n − 1.

That is, the firm value at time T is the primitive firm value with the capital asset CA, plus the cash flow of the firm over the final period, net of the capitalized fixed costs. The limited liability of a corporation is assumed in this model, and therefore, the firm value is bounded from being negative.

Assumption 5. Investment Decisions The firm has an option to make $I capital investment at each node, (n, i) every year over the planning horizon. For simplicity, we assume that the decisions are not reversible in that the firm cannot undo the investments in any future state of the world.

The increase of the capital investment leads to a direct increase in the firm's capital assets. And, we have

$$CA(n+1, 2i) = CA(n+1, 2i+1) = CA(n, i) + I(n, i) \tag{75.7}$$

The GRI of the business is not affected by the increase in the firm size, and hence, the business risk is independent of the capital investment. However, the sales are affected by the capital budgeting decisions. The marginal increase in sales to the firm with the investment at the node (n,i) is given by

$$Sales(n, i) = I(n, i) \times GRI(n, i) \tag{75.8}$$

Finally, the investment decisions are made at all the nodes such that the firm value is maximized. It is important to note that since the firm can decide on the investment at each state of the world, $CA(n,i)$ at each node depends on the path to that node. Therefore, the model is a path-dependent model.

These assumptions complete the description of the model. Assumption (1) describes the risk class of the business. Assumption (2) introduces the primitive firm enabling us to relate the cost of capital ρ to the risk-neutral valuation framework. Assumption (3) specifies the business model of the firm, identifying the firm's free cash flows as the residual of all the claims, like the fixed costs, on the firm's sales. We use the simplest business model in this chapter, but this assumption can be generalized to study different business models, which can be specified by different cost and sales structure. Assumption (4) specifies the terminal condition, following the standard assumptions made on the horizon in strategic planning. Assumption (5) specifies the marginal investment $I(n,i)$ that increases the capital asset $CA(n,i)$ The marginal returns of the investments can be generalized, even though, we choose the simplest relationship here. Given the above assumptions, we can now determine the maximum value of the firm based on the optimal capital investment decisions.

Let us use the following numerical example to illustrate the model in Table 75.1. Consider a particular scenario over two periods, where n = 0, 1, 2.

The stochastic variable is GRI. Given the investment schedule on line 3, the capital asset over time is given by line 4. Sales are determined by GRI and CA. The fixed cost is constant over time. The free cash flow is then determined following the standard income statements.

Table 75.1 The numerical example for scenario

Time	0	1	2
GRI	0.05	0.1	0.15
Investment		1	0
CA	30	31	31
Sales		3.1	4.65
FC		3	3
Investment		1	0
Free cashflows		−0.9	1.65

75.3 Simulation Results of the Capital Budgeting Decisions

The firm seeks to maximize the firm value by using the control variables, which are the capital investments, at all the node points along the scenario paths in the binomial lattice to the horizon date. There are $\sum_{n=1}^{T} 2^n$ capital investment decisions.

We use the backward substitution method based on the non-recombining tree. We first assume a set of investment decisions, $I(n, i)$, at each node along all paths. Note that $I(n, i)$ can equal to 0 in some of the nodes. At the horizon date, we can determine the firm value at each node. Then we use the risk-neutral probability and determine the firm value at time $T - 1$, such that the firm value at that node point has a risk-free return based on the risk-neutral probability, V*. Specifically,

$$V^*(T-1, i) = \frac{(p \cdot V(T, 2i+1) + (1-p)V(T, 2i))}{1 + R_F} \tag{75.9}$$

for i = 0, 1, . . .$2^{T-1} - 1$.

Note that we are rolling back a non-recombining tree and not a recombining lattice. Therefore, the state i here denotes a state along a scenario path of a tree, and the states (2i + 1) and 2i refer to the binary states of the subsequent period.

If the firm value is less than the value FC + I, which is the cash outflow, the firm declares bankrupt and has value zero; otherwise, the firm value is $V^* - FC - I$. That is,

$$V(T-1, i) = max\left[\begin{array}{l}(V^*(T-1, i) + \\ GRI \cdot CA(T-1, i) - FC - I(T-1, i), 0\end{array}\right] \tag{75.10}$$

Note that we have assumed that the firm has decided on all the investment decisions at the beginning of the period. Therefore, at each node, the firm is obligated to invest I(n,i), a nonnegative value. We continue with this process recursively, rolling back one period at a time. We then determine the firm value. That is, we recursively apply the following Eq. 75.11 till n = 0.

$$V(n-1,i) = \max\left[\begin{array}{ll}\dfrac{(p\cdot V(n,2i+1)+(1-p)V(n,2i))}{1+R_F}+ & \\ GRI(n-1,i)\times CA(n-1,i)-FC-I(n-1,i), & 0\end{array}\right] \tag{75.11}$$

for n = 0, 1, ...T − 1 and i = 0, 1, ..., $2^n - 1$.

We now seek a set of investment decisions $I(n, i)$ along all the paths to determine the highest value of the firm. This search can be accomplished by a nonlinear optimization procedure.

For clarity of the exposition, we have chosen to use a non-recombining tree and a nonlinear optimization to determine the firm value. However, the model can be specified using a recombining binomial lattice, and the firm value can be determined using the standard roll back method, without the use of any numerical nonlinear optimization method. The model is presented in Appendix 3. We have shown that the two approaches are equivalent.

We simulate the model with the following inputs: the risk-free rate of 10 %; the cost of capital ρ of 10 % with volatility σ of 30 %; a risk-neutral probability p of 0.425557; an initial capital asset CA of \$30 million; and the capitalized fixed cost of FC/0.1, where the capitalized fixed cost is assumed to present value of the perpetual fixed-cost payment discounted at the risk-free rate. We consider the problem over 5 years where the firm can invest \$1 million on a new distribution center at each node on the binomial lattice. The optimal investment decisions (top-down capital budgeting decisions) are determined by a nonlinear optimal search algorithm[1], where the investment decisions are the choice variables with the objective function being the firm value.

The optimal decision can be related to the capital budgeting decisions. When the capital investment is made, the free cash flow (CF) is given by

$$CF = GRI \times CA - FC \tag{75.12}$$

Investment decisions should be made at the margin, and therefore, one may argue that the fixed operating cost is not needed to be considered. Given that the cost of capital is ρ, then the net present value of the project is

$$NPV = \frac{GRI \times I}{\rho} + GRI \times I - I \tag{75.13}$$

The capital investment is made when NPV > 0.[2] This capital budgeting decision can be called a bottom-up approach. In this approach, line managers deal with the capital budgeting decisions, maximizing the net present value of each project which

[1]We use an optimization subroutine, GlobalSearch, written in Mathematica. The description of the procedure is provided at www.loehleenterprises.com.

[2]For clarity of the exposition, let the NPV be defined by Eq. 75.13. To be precise, the expected cash flow may not be perpetual in the presence of default. We will explain the implication of default on the free cash flow later in this section.

they manage. The capital budgeting decisions are made from the line manager's point of view rather than the headquarters' overall view, in the sense that the line manager accepts or rejects a project by focusing on the project's cash flows. If we assume that there is no fixed operating cost, the bottom-up approach can be shown to be the same as the top-down approach. In sum, the net present value maximization at the local level should lead to the global optimization for the firm that has no fixed operating costs.

Figure 75.1 shows the capital budgeting decisions using the bottom-up approach, where 1 and 0 denote the acceptance and rejection decision, respectively. For example, 0 at the top node represents that the firm rejects the project at time 0 and state 0. Since we assume the non-recombining tree, we have 2^n states at period n.

However, if we optimize the capital budgeting decisions using the top-down method, we have different optimal decisions shown in Fig. 75.2, where the shaded nodes represent the states where the capital budgeting decisions differ between the bottom-up and top-down methods.

When we compare Figs. 75.1 and 75.2, the results show that the firm accepts projects using the bottom-up approach that are rejected by the top-down approach. That means many NPV positive projects may have negative impact to firm value.

Note that our model is consistent to that of Myers (1977). By viewing the fixed operating cost as claims to the value of the primitive firms, positive net present value project may not be accepted by the global optimization to maximize the firm value. Our model interprets the result to suggest that portion of the fixed operating costs should be incorporated in the calculation of the net present value of the project. Therefore, our valuation framework provides a model to adjust for the presence of fixed operating costs in capital budgeting in a multi-period context, something that the Myers model does not cover.

Specifically, we define the marginal present value $MPV(n, i)$ at any node point (n, i) to be the marginal increase in the firm value in accepting a project at node (n, i) based on the top-down optimized solution. It is computed

$$MPV(n,i) = FV^{*}(n,i) - FV(n,i) \tag{75.14}$$

where $FV^{*}(n, i)$ and $FV(n, i)$ are the firm values at the node (n, i) with the investment and without the investment, respectively, while holding all the investment decisions in other nodes constant. This definition of the marginal change in the firm value isolates the effect of the investment at a specific node from the growth options at the other nodes.

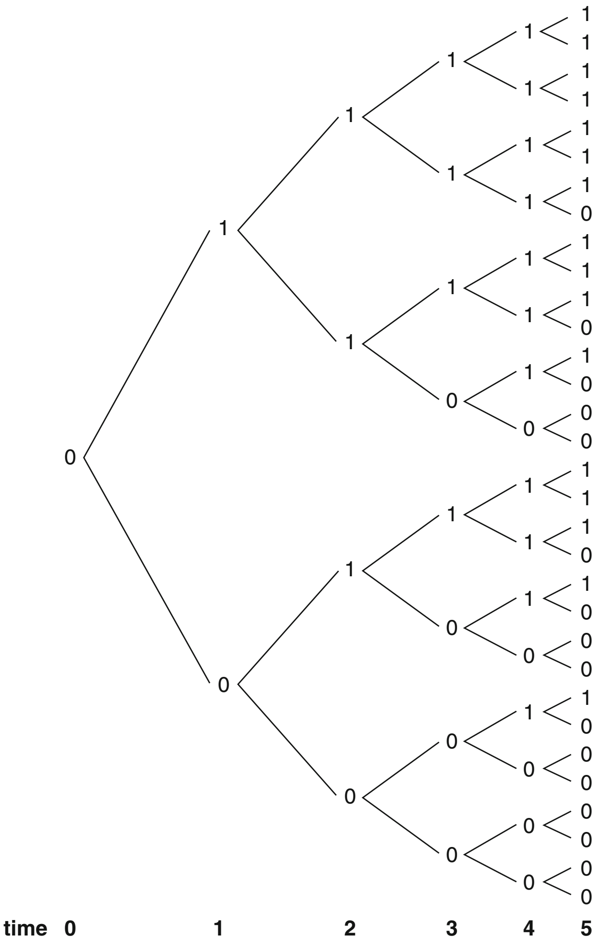
Let $NPV(n, i)$ be the net present value of the project at node (n, i) such that

$$NPV = PV - I \tag{75.15}$$

and let the present value of the fixed cost as a function of the primitive firm value at node (n, i) be

$$F(n,i) = F\left(V_p\right) \tag{75.16}$$

Fig. 75.1 Bottom-up capital budgeting decisions



Then the "wealth transfer," WT, the loss of the shareholders value in taking the project in the presence of the fixed cost, is given by

$$WT(n,i) = \frac{dF}{dV_p} PV(n,i) \tag{75.17}$$

It follows that the marginal change in the firm value is the NPV net of the wealth transfer effect.

$$MPV(n,i) = NPV(n,i) - WT(n,i) \tag{75.18}$$

Rearranging Eqs. 75.15, 75.17, and 75.18, we have

$$MPV(n,i) = PV(n,i) \cdot D - I \tag{75.19}$$

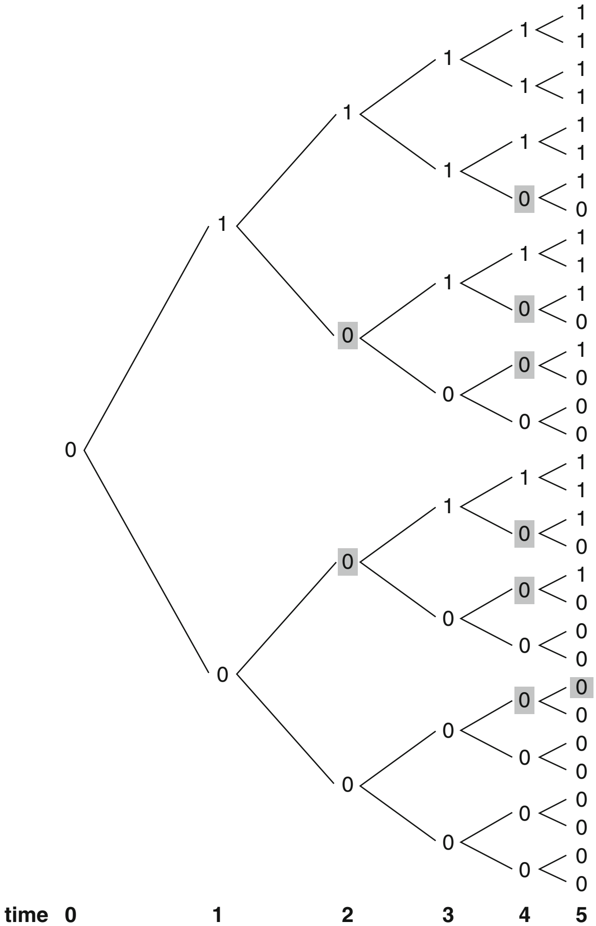


Fig. 75.2 Top-down capital budgeting decisions

$$D = 1 - \frac{dF}{dV_p} \tag{75.20}$$

Note that D as a function of the firm value is determined by the fixed-cost structure. Since we assume that the fixed cost is a fixed cash flow at any time and state, the function is the same at any node. Equation 75.19 can be interpreted intuitively. In the presence of a fixed cost, the capital budgeting decision depends on the fixed-cost factor. Portion of the present value of the incoming cash flow should first be adjusted by the fixed-cost factor and the project is taken (rejected) if MPV > (<) 0.

D can be derived in our model, as the fixed cost can be valued. The plot of the fixed-cost factor as a function of the percentage change in the firm value is provided in Fig. 75.3 below.

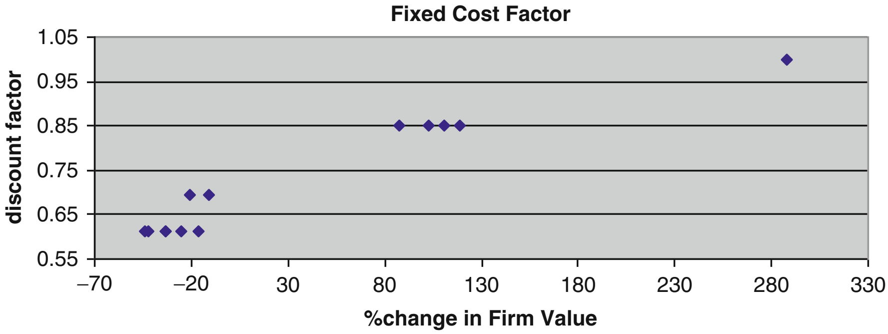


Fig. 75.3 Fixed-cost factor

As expected, the fall in the firm value would lead to a lower discount factor, resulting in more positive NPV projects being rejected. When the firm value is significantly high, the fixed-cost factor is one, and then the NPV and the top-down approach are the same. In general, given a firm's business model, the fixed-cost factor function can be derived. And this function provides the link between the bottom-up and top-down capital budgeting problem.

The model assumes that all the projects are independent of each other in the sense that the capital budgeting decision of one project is independent of the other projects. Using the bottom-up method, the independence of projects would lead to independence in the capital budgeting decisions across the projects. Yet in the presence of the fixed operating cost, it is straightforward to show that optimal decisions of the projects are related. For example, referring to Fig. 75.2 in the top-down capital budgeting decision, we would optimally invest in the upstate for period 1 and would again optimally invest in period 2 in both the upstate and downstate. However, if we do not invest in period 1, then the top-down optimal solution would lead to no investment in the subsequent downstate in the second period. Therefore, the model shows that these projects are not independent in the capital budgeting decision, contrary to the bottom-up capital budgeting decision rule.

75.4 Relative Valuation of Equity

In this section, we decompose the value of the primitive firm into its components. We recognize that the firm's capitalization is a compound option on the underlying business risk. These embedded options are options on the financial leverage, operating leverage and the strategic value. We estimate these option values using a sample of retail chain stores, and we show that such decomposition can provide us useful insights into the valuation the firms' equities.

In deriving the value of the firm, we assume that the firm pays out all the free cash flows and we construct a recombining lattice from the tree described in

Table 75.2 Inputs to the business model

	Target	Lowe's	Wal-Mart	Darden
GRI	2.9769	2.5807	4.8082	2.2823
Gross profit margin(m)	0.3169	0.2880	0.2274	0.2222
Fixed cost/total asset(FC/CA)	0.6564	0.4684	0.7907	0.2293
Capital investment (I/CA)	0.1563	0.2381	0.1530	0.1130
Leverage(CA/E)	1.7218	1.2965	1.3033	1.7190

GRI = Sales/Capital assets
Gross profit margin = (Sales − cost of goods sold)/Sales
I/A = Capital investment/Capital assets
Leverage = Capital assets/Equity

the previous section, such that the firm value is derived by the rolling back procedure. The recombining lattice is described in Appendix 3.

We consider the following retail chain stores: Wal-Mart, Target, Lowe's and Darden. We will describe these firms in brief. Wal-Mart (WMT) is now the largest retailer in North America. The company is operating approximately 4,000 stores worldwide. WMT is also a leader in developing and implementing retail information technology. Target (TGT) is the fourth largest US general merchandise retailer. It has approximately 1,000 stores, with much of their revenue derived from its discount stores. Lowe's Companies (LOW) is the second largest US home improvement retailer, selling retail building materials and supplies through more than 600 stores. Darden Restaurants, Inc (DRI) has over 600 restaurants and is a leader in the casual dining sector. The details of implementing the model using the observed data are provided in Appendix 4.

In this sample of firms, they all share the basic business model of retail chain business. They focus on their core production of their products and they sell the products through their distribution networks. The turnover, which is the sales to the total asset, depends on consumer spending. In times of recession, consumers may lower their spending on merchandizing, dining, and expenditures. As such, we may consider these firms as belonging to the similar risk class with similar cost of capital for the business.

We have shown that the inputs to the business model are profit margin m, fixed operating costs FC, turnover x, capital investment rate I, and leverage l. All these inputs can be derived or observed from the financial statements. The business model then derives the market value of equity, which is also observed in the market. We can calibrate the cost of capital of the business and the business volatility such that the equity value and the model inputs best fit the observed values.

Specifically, we use the data below, based on January 31, 2002 financial statements, as input to the business model in Table 75.2.

We then determine the implied volatility of the business risk driver (volatility) and the implied cost of capital of the business by minimizing the sum of squares of the observed market performance measures and the corresponding model value. The results are presented in Table 75.3.

Table 75.3 Reported and estimated market performance measures

	Target		Lowe's		Wal-Mart		Darden	
	Reported	Estimated	Reported	Estimated	Reported	Estimated	Reported	Estimated
S/E	5.1009	5.3231	5.3543	5.4717	7.6893	7.8420	3.1623	3.2699
S/V	0.8321	0.8500	0.9054	0.9089	0.9351	0.9363	0.8634	0.8779
Cost of capital	0.0860		0.0821		0.0702		0.1036	
Volatility	0.2776		0.3908		0.3149		0.3772	

The cost of capital and volatility are used to calibrate the model such that the sum of squares of the difference between reported and the estimated S/E, p/e and S/V is minimized

The results show that the cost of capital implied from the firm's equity value for Wal-Mart is particularly low when compared with the other retail chain stores. Darden has the highest cost of capital, which is 12.42 %. Target and Lowe's has similar cost of capital of about 9 %.

We can now decompose the primitive firm value of each firm into it building blocks of value. We have shown that the market equity value is a compound option of three options. Equity is an option on the firm value, which has an embedded real option net of the "perpetual risky coupon debt" of the fixed costs. Or the market equity value can be built from the underlying firm value. Starting from the underlying firm value, we can add the real option and net of the perpetual debt. Then all equity firms with a real option (which is an all equity growth firm) are the underlying risky assets, whose European call option is the market value of equity.

Let V_p be the value of the firm without debt, growth, or fixed costs, which we called the primitive firm. It can be calculated by using the valuation model assuming that the fixed cost and capital investment rate are zero. Let V_{fc} be the value of the firm without debt and growth, but has the fixed costs, which we call the fixed-cost firm. F is the market value of the fixed costs, which is defined as

$$F = V_p - V_{fc} \tag{75.21}$$

Let V be the value of the firm without debt, but with optimal capital investment strategy and fixed cost. Then G is the value of the growth option, which can be calculated as the difference between the firm value V and the firm without growth, V_{fc}.

$$G = V - V_{fc}. \tag{75.22}$$

D is the market value of the debt, relatively valued to the firm value V. Then the market capitalization of the firm (market value of the equity) is the underlying firm with the growth option net of the fixed costs and the debt.

$$S = V - D \tag{75.23}$$

Or the equity value can be reexpressed as

Table 75.4 The value decomposition of the capitalization value S

	Target	Lowe's	Wal-Mart	Darden
Mkt equity (S)	41,840	36,520	275,270	3,385
Primitive firm (V*)	161,313	84,728	762,427	9,612
Mkt value fixed cost (F)	129,766	60,269	568,304	6,438
Growth option (G)	17,674	15,722	99,878	682
Mkt value of debt (D)	7,382	3,661	18,732	471
Book equity (E)	7,860	6,674	35,102	1,035
Estimated (S/E)	**5.3231**	**5.4717**	**7.8420**	**3.2699**

Table 75.5 Determinants of the market multiples

	Target	Lowe's	Wal-Mart	Darden
$\mathbf{V_p/CA}$	11.9200	9.7913	16.6651	5.4017
$\mathbf{V_{fc}/V_p}$	0.1956	0.2887	0.2546	0.3302
$\mathbf{V/V_{fc}}$	1.5602	1.6428	1.5145	1.2148
CA/E	1.7218	1.2965	1.3033	1.7190
S/V	0.8500	0.9089	0.9363	0.8779
S/E	**5.3231**	**5.4717**	**7.8420**	**3.2699**
Book value(E/shares)	8.7062	8.6044	7.8004	5.8818

$$S = V_p - F + G - D \tag{75.24}$$

The decomposition of the value is summarized in Table 75.4.

To compare the results across the firms, we can normalize the equity value by the firm's book equity value, by considering the market-to-book multiples (S/E), as reported in the last row of Table 75.4. The results show that Wal-Mart has the highest multiple of 7.8420.

Now, Table 75.5 provides insights into the determinants of the market multiples of the firms. We can use the above results and derive the values in proportions as reported below.

Note that the multiple (S/E) is the product of all the ratios presented in the rows above. And therefore, the table provides a decomposition of the equity multiples. The result shows that the firms have significant fixed operating costs. For example, Target's fixed operating cost is 80.44 % of the primitive firm value. However, the market assigns a significant growth value to Target. In fact, the firm with growth option is a multiple of 1.5602 to the firm without growth option. The results also show that Wal-Mart attains the high multiple because of its high value of the primitive firm value to its total asset. As we have shown above, the high multiple value is mainly the result of a market low cost of capital to the firm business.

The business model provides a systematic approach to determine the building blocks of value to the market observed equity value. And therefore, this approach provides us insight into the determinants of the market value of equity.

Table 75.6 Stock return decomposition by firm values

	Target	Lowe's	Wal-Mart	Darden
$\ln(V_p/CA)$	0.0100	0.0100	0.0100	0.0100
$\ln(V_{fc}/V_p)$	0.0138	0.0085	0.0107	0.0070
$\ln(V/V_{fc})$	–0.0022	–0.0014	–0.0022	0.0003
ln(S/V)	0.0027	0.0010	0.0013	0.0016
ln(CA/E)	0.0000	0.0000	0.0000	0.0000
ln(Book value(E/shares))	0.0000	0.0000	0.0000	0.0000
ln(Stock price)	**0.0244**	**0.0180**	**0.0197**	**0.0189**

75.5 Equity Return Attribution

The business model can also provide insights into the relationship between the equity value and the firm's revenue. In this section, we use the business model to determine the impact of a 1 % increase in the gross investment return on the stock returns. And in the process, we determine impact of the operating leverage, financial leverage, and the growth option on the equity returns.

First note that the stock price multiple to the book value can be expressed as follows.

$$\frac{S}{E} = \frac{V_p}{CA} \times \frac{V_{fc}}{V_p} \times \frac{V}{V_{fc}} \times \frac{S}{V} \times \frac{CA}{E} \tag{75.25}$$

It follows that

$$\ln\frac{S}{E} = \ln\frac{V_p}{CA} + \ln\frac{V_{fc}}{V_p} + \ln\frac{V}{V_{fc}} + \ln\frac{S}{V} + \ln\frac{CA}{E} \tag{75.26}$$

Given in proportional increase of GRI by 1 %, the change of the equity to book multiple is given by

$$\Delta\ln\frac{S}{E} = \Delta\ln\frac{V_p}{CA} + \Delta\ln\frac{V_{fc}}{V_p} + \Delta\ln\frac{V}{V_{fc}} + \Delta\ln\frac{S}{V} + \Delta\ln\frac{CA}{E} \tag{75.27}$$

This equation provides an attribution of the proportional change in the stock multiple. The changes of the components are simulated and are provided in Table 75.6.

Note that the sum of the rows equal to the stock price change (the last row). For example, consider Wal-Mart; 1 % increase in the gross return on investment, and hence 1 % increase in sales, would lead to 1.97 % increase in the equity value. The return attribution shows that a significant portion of this return comes from the increase in the primitive firm value (1 %) and the effect of the operating leverage (1.07 %). The % increase of the primitive firm value is directly proportional to the % increase in revenue, by definition. The increase in the growth option value is impacted less by the revenue change, resulting in a negative contribution of the equity returns (–0.22 %). The financial leverage has an insignificant impact (0.13 %) because of the relatively low financial leverage of Wal-Mart measured in market value. The capital asset and book equity value are not affected by the change in revenues. This result seems to apply approximately to other stores in this sample of retail chain stores, showing that these stores are quite similar in the sense that they are industry leaders in their specific businesses. However, for retail chain stores with higher operating leverage relative to the firm's value, then the relationships are more complex.

The analysis shows that the business model enables us to identify how the operating leverage and financial leverage affect the equity returns and thus provides useful insights into the relationship between the market valuation and the profitability of the business. The use of the contingent claim approach to formulate the business risk enables us to incorporate the risk of the business (the volatility of the gross return on investment) to the debt structure and the operating leverage of the firm, something that the traditional financial ratio approach cannot capture.

75.6 Conclusions

This chapter provides a parsimonious model of a firm. The model enables us to value the firm as a contingent claim on the business risks. Using a contingent claim valuation framework, we can then relate the firm maximization to the capital budgeting rule, the fixed operating costs, and the cost of capital of the project as well as that of the firm. The model enables us to determine the impact of the fixed costs on the NPV valuation of a project. The business model also enables us to gain insight into the building blocks of value for the firm's equity and the relationship of the equity returns to its revenues.

Specifically, we have shown that the top-down and bottom-up decisions are related by the fixed-cost factor, which is a function of the firm value. This function can be specified given the business model of the firm. The lower the firm value is, the deeper the discount on the present value of the project is. Therefore, this may lead to a rejection of a positive NPV project. This result has several implications in corporate finance. For some firms with high operating leverage, for example, communication companies, seeking to acquire other firms, the model suggests that the acquisition analysis should focus not only on the synergic effect in the capital budgeting decision but the fixed-cost factor opposing effect. For a start-up company, the extensive use of the operating cost substituting the variable costs would adversely affect its capital budgeting decisions.

While we use retail chain stores to describe the business model, other businesses also share a similar model. Also, the model can be generalized to incorporate multiple risk sources or perpetual fixed operating costs with more complex fixed-cost schedules. These and other extensions of the model are not expected to change the key insights provided in the chapter.

Furthermore, the model assumptions can be relaxed to further investigate other corporate finance issues. For example, the model can analyze the impact of the fixed operating costs on the debt structure. Debts can be viewed as junior debt to the "perpetual debt," the fixed operating cost. The underlying security in this contingent claim valuation is the primitive firm. The impact of the fixed operating costs on the firm's debt may explain the bond behavior observed in the market, as the bond would behave like a junior debt (Ho and Lee 2004b).

Appendix 1: Derivation of the Risk-Neutral Probability

The risk-neutral probabilities $p(n,i)$ can be calculated from the binomial tree of V_p. Let $V_p(n,i)$ be the firm value at node (n,i). In the upstate, the firm value is

$$V_p(n,i) = \frac{GRI(n,i) \times CA}{\rho}$$

By the definition of the binomial process of the gross return on investment,

$$V_p(n+1, 2i+1) = V_p(n,i)e^{\sigma}$$

Further, the firm pays a cash dividend of $C_u = V_p(n,i) \times \rho \times e^{\sigma}$. Therefore, the total value of the firm V_p^u, an instant before the dividend payment in the upstate, is

$$V_p^u = V_p \times (1+\rho) \times e^{\sigma} \tag{75.28}$$

Similarly, the total value of the firm V_p^d, an instant before the dividend payment in the downstate, is

$$V_p^d = V_p \times (1+\rho) \times e^{-\sigma} \tag{75.29}$$

Then the risk-neutral probability p is defined as the probability that ensures the expected total return is the risk-free return.

$$p \times V_p^u + (1-p) \times V_p^d = (1+R_F) \times V_p \tag{75.30}$$

Substituting V_p, V_p^u, V_p^d into equation above and solve for p, we have

$$p = \frac{A - e^{-\sigma}}{e^{\sigma} - e^{-\sigma}} \tag{75.31}$$

Where $A = \frac{1+R_F}{1+\rho}$.

Appendix 2: The Model for the Fixed Operating Cost at Time T

When the firm may default the fixed operating cost, the fixed operating cost can be viewed as a perpetual debt of a risk bond. The valuation formula of the perpetual debt is given by Merton (1973).

$$\Phi(V,\infty) = \frac{FC}{r_f}\left\{1 - \frac{\left(\frac{2FC}{\sigma^2 V}\right)^{\frac{2r_f}{\sigma^2}}}{\Gamma\left(2+\frac{2r_f}{\sigma^2}\right)} M\left(\frac{2r_f}{\sigma^2}, 2+\frac{2r_f}{\sigma^2}, \frac{-2FC}{\sigma^2 V}\right)\right\},$$

Where

V = the primitive firm value

FC = fixed cost per year

r_f = risk free rate

Γ = the gamma function (defined in the footnote)

σ = the standard deviation of $\widetilde{GRI}$

M (•) = the confluent hypergeometric function (defined in the footnote)

$$M\left(a, 2+a, -\frac{2FC}{\sigma^2 V}\right)$$
$$= \frac{1}{br_f} e^{\frac{b}{V}} \left[-(a+a)bFC\left(\frac{b}{V}\right)^a + e^{\frac{b}{V}} FC(aV\Gamma(2+a) + (1+a)(b-aV)\Gamma\left(1+a, \frac{b}{V}\right) \right]$$

where

$$a = \frac{2r_f}{\sigma^2}, b = \frac{2FC}{\sigma^2}, \Gamma(x) = \int_0^\infty t^{x-1}e^{-t}dt,$$
$$and\ \Gamma(a,x) = \int_0^\infty t^{x-1}e^{-t}dt.$$

Appendix 3: The Valuation Model Using the Recombining Lattice

In this model specification, we assume that the GRI stochastic process follows a recombining binomial lattice:

$$GRI(n,j) = q \times GRI(n+1,j+1) + (1-q) \times GRI(n+1,j),$$
$$q = \frac{1-e^{-\sigma}}{e^{\sigma}-e^{-\sigma}}, \text{where } \sigma \text{ is the volatility of GRI},$$

where n = 0, 1,... T and j = 0, ..., n.

At time T, the horizon date, consider the node (T, j); j is the state on a recombining lattice. Suppose that the firm has made k investments in the period T, where $0 \leq k \leq T - 1$. The firm value is given by Eq. 75.32:

$$V(T,j;k) = Max \begin{bmatrix} \dfrac{GRI(T,j) \cdot (CA + (k+1)I)}{\rho} - FC(T) + \\ (GRI(T,j) \times (CA + (k+1)I) - FC - I), \\ \dfrac{GRI(T,j) \cdot (CA + kI)}{\rho} - FC(T) + \\ (GRI(T,j) \times (CA + kI) - FC), 0 \end{bmatrix}, \tag{75.32}$$
$$\text{for k} = 0, \ldots, \text{T and j} = 0, \ldots, \text{T}$$

and CA is the initial capital asset.

Now we roll back one period. We then compare the firm value with or without making an investment I. Given that the firm at the end of the period T − 1 has already invested k times and would not invest at time T − 1, the firm value is

$$\begin{aligned} &\frac{p \cdot V(T,j+1;k) + (1-p) \cdot V(T,j;k)}{1+R_F} \\ &+GRI(T-1,j) \cdot (CA + kI) - FC \end{aligned} \tag{75.33}$$

If the firm at that time invests in the capital asset, then the firm value is

$$\begin{aligned} &\frac{p \cdot V(T,j+1;k) + (1-p) \cdot V(T,j;k)}{1+R_F} \\ &+GRI(T-1,j) \cdot (CA + (k+1)I) - FC - I \end{aligned} \tag{75.34}$$

Optimal decision is to maximize the values of the firm under three possible scenarios: taking the investment, not taking the investment, or defaulting. Therefore, the value of the firm at the node (T − 1, j) with k investments is

$$\begin{aligned} v &= V(T-1,j;k) \\ &= Max \begin{bmatrix} \dfrac{p \cdot V(T,j+1;k+1) + (1-p) \cdot V(T,j;k+1)}{1+R_F} + \\ GRI(T-1,j) \cdot (CA + (k+1)I) - FC - I, \\ \dfrac{p \cdot V(T,j+1;k) + (1-p) \cdot V(T,j;k)}{1+R_F} + \\ GRI(T-1,j) \cdot (CA + kI) - FC, 0 \end{bmatrix}, \\ &for = 0,1, \ldots, T-1, and\ j = 0,1, \ldots, T-1. \end{aligned} \tag{75.35}$$

Now we can determine the firm value recursively for each n, n = T − 1, T − 2,. . .1.

At the initial period,

$$\begin{aligned} &V(T,j;k) = \\ &Max\left[\begin{array}{l}\dfrac{p\cdot V(T,j+1;k)+(1-p)\cdot V(T,j;k)}{1+R_F}+\\ GRI(T,j)\cdot(CA+kI)-FC,0\end{array}\right], \\ &for\ T=j=k=0. \end{aligned} \tag{75.36}$$

The firm value at the initial time can then be derived by recursively rolling back the firm value to the initial point, where $n = 0$. We follow the method of the fiber bundle modeling approach in Ho and Lee (2004a).

To illustrate, we use a simple numerical example. Following the previous numerical example, we assume that the GRI is 0.1, the capital asset CA is 30, the risk-free rate and the cost of capital are both 10 %, the risk-neutral probability is 0.425557, the volatility 30 %, the fixed cost FC is 3, and finally the investment is 1.

Given the above assumption, the binomial process is presented below.

The binomial lattice of GRI

Time	0	1	2	3
j	GRI			
3				0.245960311
2			0.18221188	0.134985881
1		0.134985881	0.1	0.074081822
0	0.1	0.074081822	0.054881164	0.040656966

Given the GRI binomial lattice, we can now derive the firm value lattices. The values are derived by backward substitution. The firm value depends on the capital asset level CA, the state j, and the time n.

		Firm value	V(n,j,CA)		
State j					CA
3				55.2836	32
2				14.9999	32
1				0.0000	32
0				0.0000	32
j					
3				52.5780	31
2			31.0516	13.5150	31
1			5.3286	0.0000	31
0			0.0000	0.0000	31
j					
3				49.8725	30
2			29.0473	12.0302	30
1		14.9802	4.6541	0.0000	30
0	6.329610	1.0230	0.0000	0.0000	30
Time n	**0**	**1**	**2**	**3**	

At time 3, the firm values are derived by Eq. 75.32 for each level of outstanding capital asset level at time 3, an instant before the investment decision. Then the firm values for time 2 are derived by Eq. 75.34. Once again, the firm value depends on the outstanding CA level. The firm value at time 0 does not involve any investment decision, and therefore, it is derived by rolling back from the firm values where the CA level is 30.

Appendix 4: Input Data of the Model

The input data of the model are derived from the balance sheets and income statements of the firms.

IS	Target	Lowe's	Wal-Mart	Darden
Revenue	39,888	22,111.1	217,799	4,021.2
Costs of sales	27,246	15,744.2	168,272	3,127.7
Gross profit	12,642	6,366.9	49,527	893.5
Gross profit margin (m)[a]	0.3169	0.2880	0.2274	0.2222
Fixed cost	8,883	4,053.2	36,173	407.7
Depreciation	1,079	534.1	3,290	153.9
Interest cost	464	180	1,326	31.5
Other incomes	0	24.7	2,013	0.9
Pretax incomes	2,216	1,624.3	10,751	301.3
Tax	842	601	3,897	104.2
Effective tax ratio (τ)[b]	0.3800	0.3700	0.3625	0.3458

[a]Gross profit/Revenue
[b]Tax/Pretax incomes

Balance sheet	Target	Lowe's	Wal-Mart	Darden
Capital assets	13,533	8,653.4	45,750	1,779.5
Gross return on invest (GRI)[a]	2.9475	2.5552	4.7606	2.2597
LTD[b]	8,088	3,734	18,732	517.9
Book equity	7,860	6,674.4	35,102	1,035.2

[a]Initial GRI without the investment = Revenue/Capital assets
[b]We assume that the firms have only one bond. This assumption can be relaxed using the information of the debt structure of a firm

Market information	Target	Lowe's	Wal-Mart	Darden
Shares[a]	902.8	775.7	4,500	176
Stock price[a]	44.41	46.07	59.98	18.6
Market capitalization (equity)[a]	40,093	35,736	269,910	3,274

(*continued*)

Market information	Target	Lowe's	Wal-Mart	Darden
Risk free rate (R_f)[b]	0.06	0.06	0.06	0.06
Coupon rate[b]	0.06	0.06	0.06	0.06
Max invest[c]	2,115	2,060.5	7,000	201

[a]market data
[b]We assume that the risk free rate and coupon rate are 6 %
[c]We use the capital expenditure in cash flow statements as the Max invest

References

Botteron, P., Chesney, M., & Gibson-Anser, R. (2003). Analyzing firms strategic investment decisions in a real options framework. *International Financial Markets, Institutions and Money, 13*, 451–479.

Brennan, M. J., & Schwartz, E. S. (1985). Evaluating natural resource investments. *Journal of Business, 58*, 135–157.

Cortazar, G., Schwartz, E. S., & Casassus, J. (2001). Optimal exploration investments under price and geological-technical uncertainty: A real options model. *R&D Management, 31*, 181–190.

Demsetz, H. (1968). The cost of transacting. *Quarterly Journal of Economics, 82*, 33–53.

Fontes, D. B. M. M. (2008). Fixed versus flexible production systems: A real option analysis. *European Journal of Operational Research, 188*, 169–184.

Gilroy, B. M., & Lukas, E. (2006). The choice between greenfield investment and cross-border acquisition: A real option approach. *The Quarterly Review of Economics and Finance, 46*, 447–465.

Ho, T. S. Y., & Lee, S. B. (2004a). *The Oxford guide to financial modeling*. New York: Oxford University Press.

Ho, T. S. Y., & Lee, S. B. (2004b). Valuing high yield bonds: A business modeling approach. *Journal of Investment Management, 2*, 1–12.

Ho, T. S. Y., & Marcis, R. (1984). Dealer bid-ask quotes and transaction prices: An empirical study of some AMAX options. *Journal of Finance, 39*, 23–45.

McDonald, R. L. (2006). The role of real options in capital budgeting: Theory and practice. *Journal of Applied Corporate Finance, 18*, 28–39.

Merton, R. (1973). On the pricing of corporate debt: The risk structure of interest rates. *Journal of Finance, 29*, 449–470.

Myers, S. C. (1977). Determinants of corporate borrowing. *Journal of Financial Economics, 5*, 147–175.

Myers, S. C. (1984). Finance theory and financial strategy. *Interface, 14*, 126–137, reprinted in 1987, Midland Corporate Finance Journal 5, 6–13.

Sodal, S., Koekebakker, S., & Aadland, R. (2008). Market switching in shipping: A real option model applied to the valuation of combination carriers. *Review of Financial Economics, 17*, 183–203.

Stoll, H. R. (1976). Dealer inventory behavior: An empirical investigation of NASDAQ stocks. *Journal of Financial and Quantitative Analysis, 11*, 356–380.

Triantis, A., & Borison, A. (2001). Real options: State of the practice. *Journal of Applied Corporate Finance, 14*, 8–24.

Trigeorgis, L. (1993). The nature of option interactions and the valuation of investments with multiple real options. *Journal of Financial and Quantitative Analysis, 28*, 1–20.

Villani, G. (2008). An R&D investment game under uncertainty in real option analysis. *Computational Economics, 32*, 199–219.

Wirl, F. (2008). Optimal maintenance and scrapping versus the value of back ups. *Computational Management Science, 5*, 379–392.

VAR Models: Estimation, Inferences, and Applications

76

Yangru Wu and Xing Zhou

Contents

Abstract

Vector autoregression (VAR) models have been used extensively in finance and economic analysis. This paper provides a brief overview of the basic VAR approach by focusing on model estimation and statistical inferences. Applications of VAR models in some finance areas are discussed, including asset pricing, international finance, and market microstructure. It is shown that such approach provides a powerful tool to study financial market efficiency, stock return predictability, exchange rate dynamics, and information content of stock trades and market quality.

Y. Wu (✉) • X. Zhou
Rutgers Business School – Newark and New Brunswick, Rutgers, The State University of New Jersey, New Brunswick, NJ, USA
e-mail: yangruwu@business.rutgers.edu; xingzhou@business.rutgers.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_76,

Keywords
VAR • Granger-causality test • Impulse response • Variance decomposition • Cointegration • Asset return predictability • Market quality • Information content of trades • Informational efficiency

76.1 Introduction

Following seminar work by Sims (1980), the vector autoregression (VAR) approach has been developed as a powerful modeling tool for studying the interactions among economic and financial variables and for forecasting. As basic VAR models focus on the statistical representation of the dynamic behavior of time-series data but without much restriction on the underlying economic structure and can be easily estimated, they have gained increasing popularity in both economics and finance. In this chapter, we provide an overview of basic VAR models and some of their most popular applications in financial economics.

76.2 A Brief Discussion of VAR Models

A basic p-lag VAR model has the following form:

$$X_t = A + \Phi_1 X_{t-1} + \Phi_2 X_{t-2} + \ldots + \Phi_p X_{t-p} + \varepsilon_t, t = 1, 2, \ldots T, \tag{76.1}$$

where $X_t = (x_{1,t}, x_{2,t}, \ldots, x_{n,t})'$, and it is an $(n \times 1)$ vector of economic time series. In addition, Φ are coefficient matrices and ε_t is an $(n \times 1)$ vector of residuals. The residual vector is assumed to have zero mean, zero autocorrelation, and time invariant covariance matrix Ω. For example, a p-lag bivariate VAR model can be expressed as

$$\begin{aligned}\begin{pmatrix} x_{1,t} \\ x_{2,t} \end{pmatrix} &= \begin{pmatrix} a_1 \\ a_2 \end{pmatrix} + \begin{pmatrix} \phi_{1,1}^1 & \phi_{1,2}^1 \\ \phi_{2,1}^1 & \phi_{2,2}^1 \end{pmatrix} \begin{pmatrix} x_{1,t-1} \\ x_{2,t-1} \end{pmatrix} \\ &\quad + \cdots + \begin{pmatrix} \phi_{1,1}^p & \phi_{1,2}^p \\ \phi_{2,1}^p & \phi_{2,2}^p \end{pmatrix} \begin{pmatrix} x_{1,t-p} \\ x_{2,t-p} \end{pmatrix} + \begin{pmatrix} \varepsilon_{1,t} \\ \varepsilon_{2,t} \end{pmatrix},\end{aligned} \tag{76.2}$$

where $\text{cov}(\varepsilon_{1,t}, \varepsilon_{2,s}) = \sigma_{12}$ if $t = s$ and $\text{cov}(\varepsilon_{1,t}, \varepsilon_{2,s}) = 0$ otherwise.

76.2.1 Estimation

In order to estimate a VAR model, the number of lags of endogenous variables has to be determined first, as the results and hence inferences from estimation can be very sensitive to the lag choice. The general rule is to choose the lag p to minimize some model selection criteria. Two most common information criteria include the Akaike (AIC) and the Schwarz-Bayesian (SBC):

$$\mathrm{AIC}(p) = \ln\left|\bar{\sum}(p)\right| + \frac{2}{T}pn^2 \tag{76.3}$$

and

$$\mathrm{SBC}(p) = \ln\left|\bar{\sum}(p)\right| + \frac{\ln T}{T}pn^2, \tag{76.4}$$

where $\bar{\sum}(p) = T^{-1}\sum_{t=1}^{T}\hat{\varepsilon}_t\hat{\varepsilon}_t{}'$, and p, T, and n represent the lag, sample size, and number of variables, respectively. Once the lag is determined, the VAR model can be estimated using either the ordinary least squares (OLS) method or the maximum likelihood method.[1]

76.2.2 Inferences

Since a p-lag VAR model contains many parameters, interpreting the estimation results can be difficult especially when p is large. Instead of focusing on interpreting each individual parameter estimated, some summary measures are usually employed to provide useful information on the dynamics among variables included in the model. In this section, we focus on three major kinds of summary analysis.

76.2.2.1 Granger-Causality Tests

One important application of the VAR approach is forecasting. Specifically, this approach can be used to address such questions as whether some variables contain valuable information about the future dynamics of other variables in the model. Granger (1969) proposes a test on the forecasting relationship between two variables, which was further developed in Sims (1972).

According to Granger (1969), in a bivariate VAR model as discussed above, if x_1 helps predict x_2, then x_1 *Granger-causes* x_2. Otherwise, x_1 *fails to Granger-cause* x_2. More formally, x_1 *fails to Granger-cause* x_2 if for all $s > 0$, the mean squared error (MSE) of a forecast of $x_{2,t+s}$ based on $(x_{2,t}, x_{2,t-1}, \ldots)$ is the same as the MSE of a forecast of $x_{2,t+s}$ based on $(x_{1,t}, x_{1,t-1}, \ldots)$ and $(x_{2,t}, x_{2,t-1}, \ldots)$.

To test whether x_1 Granger-causes x_2, we can estimate the following model:

[1]For further discussions on VAR techniques, see Watson (1994); Lutkepohl (1991), and Stock and Watson (2001), among others.

$$x_{2,t} = a_2 + \phi^1_{2,1}\, x_{1,t-1} + \ldots + \phi^p_{2,1}\, x_{1,t-p} + \phi^1_{2,2}\, x_{2,t-1} + \ldots + \phi^p_{2,2}\, x_{2,t-p} + \varepsilon_{2,t}. \tag{76.5}$$

Then we can conduct an F-test on the null hypothesis:

$$\mathrm{H}_0\colon \phi^1_{2,1} = \ldots = \phi^p_{2,1} = 0.$$

If the null hypothesis is rejected, we can conclude that x_1 *Granger-causes* x_2.

While Granger-causality tests provide useful information on the forecasting ability of the model variables, the predicative power of one variable is not equivalent to its true causality ability. Therefore, results from Granger-causality tests need to be interpreted with great caution.

76.2.2.2 Impulse Response Analysis

Another approach to examine the interactions among variables in the VAR model is to use impulse response functions, which represent the reactions of model variables to shocks hitting the system. In order to estimate the impulse response function, we first transform the VAR model into vector moving average form:

$$X_t = \mu + \varepsilon_t + \Psi_1\varepsilon_{t-1} + \Psi_2\varepsilon_{t-2} + \Psi_3\varepsilon_{t-3} + \ldots, \tag{76.6}$$

where $\frac{\partial X_{t+s}}{\partial \varepsilon_t'} = \Psi_s$. Therefore, the (i, j) component of Ψ_s, $\psi_s^{i,j}$ captures the effects of a one-unit shock to variable j at time t (ε_t^j) on variable i at time $t + s$ $(x_{i,t+s})$ and hence is interpreted as the impulse response function.

However, such interpretation is only possible if the elements of ε_t are not correlated, i.e., the variance-covariance matrix of ε_t is a diagonal matrix. Sims (1980) proposes a *recursive causal ordering* to solve this problem. For example, if $X_t = (x_{1,t}, x_{2,t}, x_{3,t})$, then the variables in the X_t can be ordered in a way so that $x_{1,t}$ affects $x_{2,t}$ and $x_{3,t}$ but not vice versa and $x_{2,t}$ affects $x_{3,t}$ but not vice versa. Which particular order to use should be determined by the specific context and economics models being examined. With such recursive causal reordering, the VAR model can be rewritten as

$$\begin{aligned}
x_{1,t} &= a_1 + \Pi_{1,1}X_{t-1} + \Pi_{1,2}X_{t-2} + \ldots + \Pi_{1,p}X_{t-p} + \eta_{1,t} \\
x_{2,t} &= a_2 + \beta_{2,1}x_{1,t} + \Pi_{2,1}X_{t-1} + \Pi_{2,2}X_{t-2} + \ldots + \Pi_{2,p}X_{t-p} + \eta_{2,t} \\
&\vdots \\
x_{n,t} &= a_n + \beta_{n,1}x_{1,t} + \ldots + \beta_{n,n-1}x_{n-1,t} + \Pi_{n,1}X_{t-1} + \Pi_{n,2}X_{t-2} \\
&\quad + \ldots + \Pi_{n,p}X_{t-p} + \eta_{n,t}.
\end{aligned} \tag{76.7}$$

Transform this VAR model into the Wold representation:

$$X_t = \mu + \Theta_0\eta_t + \Theta_1\eta_{t-1} + \Theta_2\eta_{t-2} + \Theta_3\eta_{t-3} + \ldots, \tag{76.8}$$

where Λ_0 is a lower triangular matrix. The impulse response function can be estimated as

$$\frac{\partial x_{i,\,t+s}}{\partial \eta_{j,\,t}} = \theta^s_{i,j} \tag{76.9}$$

where $i, j = 1, 2, \ldots, n$ and $s > 0$.

76.2.2.3 Variance Decomposition

A related issue on the influence of a one-unit shock to the system is its contribution to the variance of the forecast error in predicting the future value of each variable in the system. Specifically, if we write the forecast error of X_{t+s} at time t as

$$X_{t+s} - \hat{X}_{t+s|t} = \varepsilon_t + \Lambda_1 \varepsilon_{t+s-1} + \Lambda_2 \varepsilon_{t+s-2} + \cdots + \Lambda_{s-1} \varepsilon_{t+1}, \tag{76.10}$$

the MSE of this forecast is

$$\begin{aligned} \mathrm{MSE}\left(\hat{X}_{t+s|t}\right) &= E\left[\left(X_{t+s} - \hat{X}_{t+s|t}\right)\left(X_{t+s} - \hat{X}_{t+s|t}\right)'\right] \\ &= \Omega + \Lambda_1 \Omega \Lambda_1{}' + \cdots + \Lambda_{s-1} \Omega \Lambda_{s-1}{}'. \end{aligned} \tag{76.11}$$

For one variable in the model, x_1, the forecast error has the form:

$$x_{i,\,t+s} - \hat{x}_{i,\,t+s|t} = \sum_{h=0}^{s-1} \rho^h_{i,\,1} \eta_{1,\,T+s-h} + \cdots + \sum_{h=0}^{s-1} \rho^h_{i,\,n} \eta_{n,\,T+s-h}. \tag{76.12}$$

With the same *recursive causal ordering* as discussed in the previous section, the variance and covariance matrix is orthogonal. Therefore, the variance of the forecast error for $x_{i,t+s}$ is

$$\mathrm{var}\left(x_{i,\,t+s} - \hat{x}_{i,\,t+s|t}\right) = \sigma^2_{\eta_1} \sum_{h=0}^{s-1} \left(\rho^h_{i,\,1}\right)^2 + \cdots + \sigma^2_{\eta_n} \sum_{h=0}^{s-1} \left(\rho^h_{i,\,n}\right)^2. \tag{76.13}$$

76.3 Applications of VARs in Finance

The VAR models have been used very broadly in the finance literature. In this chapter, we focus on a few important applications to illustrate the advantages and some potential concerns when applying this methodology in empirical studies.

76.3.1 Stock Return Predictability and Optimal Asset Allocation

The VAR has been widely regarded as a useful model to study stock return predictability. Campbell and Shiller (1987) propose the following VAR model for asset returns:

$$\begin{bmatrix} \Delta y_t \\ S_t \end{bmatrix} = \begin{bmatrix} a(L) & b(L) \\ c(L) & d(L) \end{bmatrix} \begin{bmatrix} \Delta y_{t-1} \\ S_{t-1} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix}. \tag{76.14}$$

where in the case of the present value model of stock prices, y_t is dividend and S_t is the difference between stock price and a multiple of dividends; and in the case of the term structure of interest rates, y_t is one-period interest rate and S_t is the spread between long-term and short-term interest rates.

For stock market data, they use real annual prices and real dividends on a broad stock index from 1981 to 1986. For term structure data, they use monthly US Treasury 20-year yield series and 1-month Treasury bill rate. They find that y_t is nonstationary and its first difference Δy_t is stationary, while S_t is in general stationary in both the stock market and term structure cases. There is weak evidence of cointegration between stock prices and dividends, although the cointegration relationship in the term structure is more significant. They find that the yield spread Granger-causes short-rate changes and that excess returns on long-term bonds are significantly predictable, although the expectation hypothesis of the term structure is formally rejected. As for stock market, dividend changes are found to be highly predictable and the dividend-price spreads Granger-cause dividend changes. The present value model of stock price is however rejected.

Using a log-linear approximation for stock prices and dividends, Campbell and Shiller (1989) estimate the following bivariate VAR:

$$\begin{bmatrix} \delta_t \\ r_{t-1} - \Delta d_{t-1} \end{bmatrix} = \begin{bmatrix} a(L) & b(L) \\ c(L) & d(L) \end{bmatrix} \begin{bmatrix} \delta_{t-1} \\ r_{t-2} - \Delta d_{t-2} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \end{bmatrix} \tag{76.15}$$

where d_t is log dividends, δ_t is log dividend-price ratio, r_t is ex post discount rate, and $r_t - \Delta d_t$ is interpreted as the growth-adjusted discount rate. Campbell and Shiller (1989) find that stock returns are somewhat predictable. The lagged log dividend-price ratio has a positive impact on stock returns, and lagged real dividend growth rate has a negative impact. The findings that dividend-price ratio has predicting power for future stock returns are consistent with earlier studies (e.g., Shiller 1984; Flood et al. 1986). Log dividend-price ratio Granger-causes real dividend growth. Short-term discount rates are not helpful in explaining stock-price movements.

Hodrick (1992) studies the statistical properties of three models for predicting stock returns on long horizons. The three model specifications are (1) OLS regression of stock return on past dividend yield originally proposed by Fama and French (1988), (2) a reorganization of the first specification where log stock return is regressed by cumulative past dividend yields, and (3) a first-order VAR in three variables: log stock returns, dividend yields, and 1-month Treasury bill return relative to its previous 12-month moving average. Using simulations, Hodrick (1992) finds that the VAR alternative has the correct size, and provides unbiased long-horizon statistics, and therefore is the preferred technique to the other two model specifications in the prediction of stock return over long horizons. Empirically, the VAR tests provide strong evidence of predictive power of the 1-month-ahead return. The VAR approach is a useful alternative way to properly calculate various statistics in long-horizon predictions, including the implied regression slope coefficients, implied R^2, and variance ratios.

Several authors report that the log of earnings, dividends, and stock prices are integrated of order one processes, and both dividend payout and dividend yields are stationary (e.g., Cochrane 1992; Mankiw et al. 1991; Lee 1996). That is, log earnings and dividends series are cointegrated of order one. Log stock prices and log dividends are also cointegrated of order one. Lee (1998) employs a tri-variate structural VAR model to identify the various components of stock price and examine the response of stock prices to different types of shocks: permanent and temporary changes in earnings and dividends and changes in discount factors and nonfundamental factors. The variables in the structural VAR are the change in log earnings, the spread between log dividends and earnings, and the spread between log stock prices and dividends. Lee (1998) finds that about half of the stock-price variation is unrelated to earnings or dividend changes. Time-varying excess stock returns account for much of the remaining deviation of stock prices. Deviation of stock prices from fundamentals does not persist for a long time, suggesting a fad interpretation of stock market movements.

The VAR has been employed to study optimal consumption and portfolio decisions over long horizons when asset returns are not i.i.d. For example, Campbell and Viceira (1999) use a restricted VAR consisting of the excess stock market returns over the risk-free rate and the log dividend-price ratio. The VAR parsimoniously describes the changing investment opportunity over time for the long-horizon investor. The predictability of asset returns over time significantly increases the complexity of the optimal consumption and portfolio solution. Using a log-linear approximation, Campbell and Viceira (1999) nicely obtain an analytic solution. Optimal portfolio consists of two components. The first component is the myopic demand which is what one would obtain in the absence of asset return predictability. This is the classical result on optimal asset allocation. The second component is the hedging demand which arises due to the serial correlation of asset returns. Campbell et al. (2003) extend the model to allow multiple assets in the VAR. Viceira (2001) examines long-horizon consumption and portfolio decisions with nontradable labor income. These studies demonstrate how the VAR can be conveniently employed to yield interesting and important insight in the consumption and portfolio optimization.

76.3.2 Exchange Rate Prediction

The VAR is widely used to study exchange rate dynamics. We summarize several papers on exchange rate predictability. There is a large literature on exchange rates forecasting, started from the seminal work of Meese and Rogoff (1983), who report that most structural and time-series exchange rate models, including an unconstrained VAR consisting of exchange rate and six key macroeconomic variables, cannot produce a better forecast of future spot exchange rates than the naïve random walk model. Subsequent authors develop various models and techniques to better understand exchange rate dynamics, hoping to produce more accurate forecast of exchange rates than the naïve random walk.

Bekaert and Hodrick (1992) consider the following VAR model for two countries (e.g., the United States which is called country 1 and Japan which is called country 2):

$$Y_t = A_0 + A_1 Y_{t-1} + u_{t+1} \tag{76.16}$$

where
$Y_t = [r_{1t}, r_{2t}, rs_{2t}, dy_{1t}, dy_{2t}, fp_{2t}]'$,
r_{jt} is the excess equity market return in country j,
rs_{jt} is the excess dollar rate of return on a currency j money market investment,
dy_{jt} is the dividend yield in country j,
fp_{jt} is the forward premium on currency j in terms of the US dollars.

They find that dividend yields that are known to have predictive power for equity returns can predict excess returns in foreign exchange market. Similarly, forward premiums that are known to predict excess returns in the foreign exchange market have predictive power for equity excess returns. They also find that excess returns in the foreign exchange market have strong positive persistence.

Baillie and Bollerslev (1989) employ a VAR to study the long-run and short-run dynamics of spot and forward exchange rates for seven major currencies against the US dollar. They report that spot and forward exchange rates can be characterized as unit-root processes, and spot and forward exchange rates are cointegrated for each currency pair. Furthermore, one common unit root, or stochastic trend, is detected in the multivariate time-series models for the seven spot and forward exchange rates. These findings suggest that the seven exchange rates possess one long-run relationship and that the disequilibrium error around the long-run relationship can partly explain the subsequent short-run movements in the exchange rates.

Diebold et al. (1994) examine one immediate implication of the Baillie and Bollerslev (1989) finding that spot and forward exchange rates are cointegrated. That is, cointegration implies an error-correction representation of spot and forward exchange rates that should better forecast the change in future spot exchange rate than the naïve random walk model. Empirical findings of Diebold et al. (1994) are however quite negative. The VAR model with the cointegration restriction does not produce more accurate out-of-sample forecast of exchanges rates than the simple martingale model in short horizons. Mark (1995) and Mark and Sul (2001) however find substantial predictability of exchange rates over longer horizons (3–4 years) when monetary fundamental variables are employed.

Numerous papers study the information contents of the forward foreign exchange rate premium in predicting future spot exchange rate changes. According to the uncovered interest rate parity, when the forward exchange rate is traded at a 1 % premium relative to the current spot exchange rate, the future spot exchange rate should be expected to depreciate by 1 %. However, empirically the uncovered interest rate parity is strongly rejected. Furthermore, the forward premium often forecasts the change in future spot exchange rate in the opposite direction. Hai et al. (1997) propose a permanent-transitory components model for the spot and forward exchange rates. The permanent component which is shared by the spot and forward exchange rates is a random walk without drift, while the transitory components of spot and forward exchange rates are assumed to follow a vector autoregressive moving average process. The model is estimated using the Kalman filter. They report that this simple parametric model is useful in understanding why the forward rate may be an unbiased predictor of future spot rate even though an increase in forward premium predicts a dollar appreciation. The estimates of the expected excess return on short-term dollar-denominated assets are persistent and reasonable in magnitude.

Mark and Wu (1998) find that a vector error correction model for the spot and forward exchange rates can account for many salient features of the data. In particular, the estimated risk premium series is highly persistent, the risk premium and the expected future depreciation of the spot exchange rate alternate between positive and negative values and change sign infrequently. They show that standard intertemporal asset pricing model is not capable of generating reasonable foreign exchange risk premiums and therefore does not help to explain why the forward exchange premium forecasts future spot exchange rate changes in the wrong direction. On the other hand, a noise-trader model along the line of De Long et al. (1990) is potentially promising in explaining the anomaly.

76.3.3 Measuring Market Quality and Informational Content of Stock Trades

The VAR approach has also been applied to high-frequency data to measure informational content of stock trades and the quality of security markets. For example, Hasbrouck (1991) uses the following vector autoregressive system to model the interaction of trades and quote revisions:

$$r_t = \sum_{i-1}^{I} a_i r_{t-i} + \sum_{i=0}^{I} b_i x_{t-i}^0 + v_{1,t}, \tag{76.17}$$

and

$$x_t^0 = \sum_{i-1}^{I} c_i r_{t-i} + \sum_{i=1}^{I} d_i x_{t-i}^0 + v_{2,t}, \tag{76.18}$$

where t indexes transactions and r_t and x_t^0 represent price changes and trade direction, respectively. Within this framework, the information content of a trade is measured as the ultimate impact of the trade innovation on the stock price and can be inferred from the impulse respond function of price changes to a shock to x_t^0. Specifically, if an order arrives at time t, its cumulative impact on quote revisions through step m is represented by

$$\alpha_m\left(v_{2,t}\right) = \sum_{i=0}^{m} E\left[r_{t+i}|v_{2,t}\right]. \tag{76.19}$$

As m increases, $\alpha_m(v_{2,t})$ can be shown to converge to the revisions in the efficient price resulting from the shock $v_{2,t}$.

Notably, this model differs from usual VAR specification since the contemporaneous net trading volume appears as one of the independent variables in explaining quotes returns. This is largely due to the trades/quote timing convention assumed in the study. According to Hasbrouck (1991), trades (x_t) take place after market makers post bid and ask quotes (r_{t-1}). Based on the trades occurred, market makers revise their quotes (r_t) and more trades follow. Therefore, even though x_t and r_t carry the same subscript t, they are not determined simultaneously. In fact, r_t takes place after x_t and hence cannot influence x_t. Such *recursive causal ordering* is also necessary in order to estimate the impulse response function.

In another influential paper, Hasbrouck (1993) uses the VAR specification to discuss a new measure of market quality. In this model, security transaction prices are decomposed into a random walk component, identified as the efficient price, and a residual stationary component, termed the pricing error. Specifically, the (logarithm of) the actual transaction price at t, p_t, is modeled as

$$p_t = m_t + s_t \tag{76.20}$$

where m_t is the efficient price which follows a random walk and s_t is the pricing error which is a zero-mean covariance-stationary stochastic process.

As discussed in Hasbrouck (1993), the pricing error consists of two components, an information uncorrelated component (e.g., arising from inventory effects) and an information correlated component (which arises from adverse selection effects). This dispersion of the pricing error, denoted σ_s, measures how closely the real transaction price tracks the efficient price and hence measures the market quality: the smaller the σ_s, the higher the market quality.

To empirically determine σ_s, Hasbrouck (1993) estimates a vector autoregressive (VAR) model over a set of four price and trade variables $\{r_t, x_t^0, x_t, x_t^{1/2}\}$, where $r_t = p_t - p_{t-1}$, x_t^0 is the trade direction (which is equal to 1 for buys and -1 for sells), and x_t and $x_t^{1/2}$ represent signed size and signed square root of size, respectively. The σ_s is then calculated for each firm as a function of the estimated variance-covariance matrix and the coefficient matrix of the VAR model under the Beveridge-Nelson identification restriction (1981).[2] Applying this measure to a sample of NYSE firms, Hasbrouck (1993) finds that the average σ_s is about 0.33 % of the stock price, and it exhibits elevation at both the beginning and end of the trading session.

76.3.4 Relative Informational Efficiency

The relative informational efficiency of related asset markets has been of great interests to finance scholars for decades. Numerous studies have employed the VAR approach to address substantial questions such as whether prices/trading activities in one market reflect new information faster than that in the other markets. Following seminal work by Black (1975), there has been a huge literature studying inter-market relationships between the stock and option markets. The underlying rationale is that any information related to the value of a firm's equity security should also be reflected in its derivative contracts, and which market moves first in response to the arrival of new information is determined to a large degree by informed traders' preferences for trading. According to Black (1975), the option market might be more attractive to informed traders than the market for the underlying stock because options offer higher financial leverage, and the option market is characterized by less stringent margin requirements and no uptick rule for short selling.

Whether the option market is indeed leading the stock market in reflecting new information has been directly examined in numerous empirical studies. One approach is to analyze the dynamic relationship between the stock and option markets using a VAR model. For example, Chan et al. (2002) propose the following multivariate VAR model on the trades and quotes revisions in options and their underlying stocks:

$$r_t = a_1 r_{t-1} + \cdots + a_p r_{t-p} + b_0 x_t + b_1 x_{t-1} + \cdots b_p x_{t-p} + \varepsilon_{1,t}, \tag{76.21}$$

and

$$x_t = c_1 r_{t-1} + \cdots + c_p r_{t-p} + d_1 x_{t-1} + \cdots d_p x_{t-p} + \varepsilon_{2,t}. \tag{76.22}$$

In this model, $r_t = [r_t^s, r_t^c, r_t^p]'$ and $x_t = [x_t^s, x_t^c, x_t^p]'$, where r_t^s, r_t^c, and r_t^p denote returns calculated using quote midpoints and x_t^s, x_t^c, and x_t^p represent net trading volume (buyer-initiated volume minus seller-initiated volume) in the stock, call, and put option markets during time interval t, respectively. Further, the error terms $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ are assumed to have zero means and are jointly and serially uncorrelated. This specification is in the same spirit as in Hasbrouck (1991) since the

[2]For more estimation details, refer to Hasbrouck (1993).

contemporaneous net trading volume appears as one of the explanatory variables for quote returns but not vice versa.

The quote returns and net trading volumes are calculated on a 5-min intervals for a sample of 14 most actively traded NYSE stocks with options traded on CBOE for a total of 231 trading days. The authors find that stock net trading volume, but not option trading volume, carries explanatory power for both contemporaneous and subsequent stock and option quote revisions. On the other hand, both stock and option quote revisions predict subsequent quote changes in both markets. These findings suggest that informed traders only initiate trades in the stock market. However, they also trade in the option market by submitting limit orders. This is consistent with the notion that the less liquidity in the option market might discourage informed traders to trade using market orders, despite the high financial leverage it offers.[3]

The VAR approach has also been used in a recent literature on the relative informational efficiency of stocks and corporate bonds. As both stocks and corporate bonds are claims on the issuing firms' assets, any information on the value of these assets will affect their prices. The analysis on the lead-lag relationships between the stock and bond price movements can then be used to address the question as to whether stocks lead bonds in reflecting firm-specific information:

$$R_{s,t} = \alpha_s + \sum_{i=1}^{I} \beta^i_{s,s} R_{s,t-i} + \sum_{i=1}^{L} \gamma^i_{s,b} R_{b,t-i} + \varepsilon_{s,t}, \tag{76.23}$$

and

$$R_{b,t} = \alpha_b + \sum_{i=1}^{I} \beta^i_{b,s} R_{s,t-i} + \sum_{i=1}^{L} \gamma^i_{b,b} R_{b,t-i} + \varepsilon_{b,t}, \tag{76.24}$$

[3]Whether options lead stocks in the price discovery process is still a question open to debate. Early studies in this literature present strong evidence in favor of the option lead in prices. For example, Latane and Rendleman (1976) and Beckers (1981) show that the volatility implied in option prices predicts future stock-price volatility. Consistently, Manaster and Renleman (1982) find that the option-implied stock prices contain valuable information about the equilibrium prices of the underlying stocks that has not been revealed in the stock market. However, Vijh (1988) questions the results in Manaster and Rendleman (1982), since using daily closing prices introduces a bias associated with the bid-ask spread and nonsynchronous trading. After purging the effects of bid-ask spreads, Stephan and Whaley (1990) find that the stock market leads the option market. Nevertheless, Chan et al. (1993) argue that the stock lead is due to the relative smaller stock tick. If the average of the bid and ask is used instead of transaction prices, neither market leads the other. Latter studies on the stock option lead-lag analysis have been focused more on the trading volume. Easley et al. (1998) show that "positive news option volumes" and "negative news option volumes" have predictive power for future stock-price changes. See also Pan and Poteshman (2006) and Cao et al. (2003). By measuring the relative share of price discovery occurring in the stock and option markets, Chakravarty et al. (2004) conclude that informed trading takes place in both stock and option markets, suggesting an important role for option volume.

where $R_{s,t}$ ($R_{b,t}$) represents the return on the stock (bond) at time t. Within this framework, Granger-causality tests are conducted to identify the lead-lag relationships.

Hotchkiss and Ronen (2002) examine pre-TRACE (FIPS) transaction price summaries for 55 high-yield bonds and find that stocks do not lead bonds, consistent with Zhou (2009) who finds that lagged TRACE 50 high-yield bond prices contain valuable information about current stock returns, and that they serve an important role in disseminating firm-specific information.[4] In contrast, Downing et al. (2009) conclude that stock returns do lead nonconvertible bond returns at the hourly level in times of financial distress and that therefore the corporate bond market is less informationally efficient.[5] Gurun et al. (2011) also find significant stock leads for daily bond indices, which diminish with certain information releases. The findings of less informational efficiency of corporate bonds seem to be consistent with the notion that the liquidity (transaction costs) for corporate bonds is much lower (higher) than that for stocks. However, in a recent study by Ronen and Zhou (2012), the authors question the ability of VAR in capturing the information flow between stocks and bonds. Since for any given firm there typically is multiplicity of bond issues (in contrast to a single or very few stock issues), VAR analysis on pair-wise comparisons of each bond with the issuer's stock can be misleading. In fact, it cannot reveal the information most desired by a single trader: whether there exists at least one bond constituting an information-based trading venue. For more discussions, see Ronen and Zhou (2012).

76.4 Summary

This chapter endeavors to summarize the basic structure of a VAR model and some related estimation methods and structural analysis. It also offers a quick overview of some important applications of the VAR approach in both economics and finance. Despite all the advantages for the VAR models, it is important to realize that VAR models can only be applied to stationary time-series data. Therefore, in order to properly implement this approach in empirical studies, certain tests on the stationarity of the time-series data have to be conducted first. If null hypothesis of stationarity is rejected, a vector error correction model (VECM) should be used instead. For more details on VAR and VECM models, please refer to Hamilton (1994).

[4]Zhou's (2009) analysis incorporates the option market trades in addition to the stock and bond market transactions. Chava and Tookes (2005) also incorporate the option markets and examine the volatility reaction of stock, bond, and options to macroeconomic and firm specific information, finding significant effects near announcements. Overall, they find that corporate bond and option trades have information content for future stock price movements.

[5]For convertible bonds, Downing et al. (2009) find that these results also hold, but only for those with conversion options deep in the money. Kwan (1996) first examines the relation between individual stocks and bonds using weekly quote data and also finds evidence in support of stock leads.

References

Baillie, R., & Bollwerslev, T. (1989). Common stochastic trends in a system of exchange rates. *Journal of Finance, 44*, 167–181.

Beckers, S. (1981). Standard deviations implied in option prices as predictors of future stock price variability. *Journal of Banking and Finance, 5*, 363–381.

Bekaert, G., & Hodrick, R. (1992). Characterizing predictable components in excess returns on equity and foreign exchange markets. *Journal of Finance, 47*, 467–509.

Black, F. (1975). Fact and fantasy in the use of options. *Financial Analysts Journal, 31*(36–41), 61–72.

Campbell, J., & Shiller, R. (1987). Cointegration and tests of present value models. *Journal of Political Economy, 95*, 1062–1088.

Campbell, J., & Shiller, R. (1989). The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies, 1*, 195–228.

Campbell, J., & Viceira, L. (1999). Consumption and portfolio decisions when expected returns are time varying. *Quarterly Journal of Economics, 114*, 433–495.

Campbell, J., Chan, Y. L., & Viceira, L. (2003). A multivariate model of strategic asset allocation. *Journal of Financial Economics, 67*, 41–80.

Cao, C., Chen, Z., & Griffin, J. M. (2003). The informational content of option volume prior to takeovers. *Journal of Business, 78*, 1071–1109.

Chakravarty, S., Gulen, H., & Mayhew, S. (2004). Informed trading in stock and option markets. *Journal of Finance, 59*, 1235–1258.

Chan, K., Chung, P., & Johnson, H. (1993). Why option prices lag stock prices: A trading-based explanation. *Journal of Finance, 48*, 1957–1967.

Chan, K., Chung, Y. P., & Fong, W.-M. (2002). The informational role of stock and option volume. *Review of Financial Studies, 15*, 1049–1075.

Chava, S., & Tookes, H. (2005). *Where (and how) does news impact trading*? (Unpublished working paper). Texas A&M University/Yale School of Management.

Cochrane, J. (1992). Explaining the variance of price-dividend ratios. *Review of Financial Studies, 5*, 243–280.

De Long, B., Shleifer, A., Summers, L., & Waldman, R. (1990). Noise trader risk in financial markets. *Journal of Political Economy, 98*, 703–738.

Diebold, F., Gardeazabai, J., & Yilmaz, K. (1994). On cointegration and exchange rate dynamics. *Journal of Finance, 49*, 727–735.

Downing, C., Underwood, S., & Xing, Y. (2009). An intraday analysis of the relative informational efficiency of stocks and bonds. *Journal of Financial and Quantitative Analysis, 44*, 1081–1102.

Easley, D., O'Hara, M., & Srinivas, P. S. (1998). Option volume and stock prices: Evidence on where informed traders trade. *Journal of Finance, 53*, 431–466.

Fama, E., & French, K. (1988). Dividend yields and expected stock returns. *Journal of Financial Economics, 22*, 3–26.

Flood, R., Hodrick, R., & Kaplan, P. (1986). *An evaluation of recent evidence on stock market bubbles* (NBER working paper).

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross spectral methods. *Econometrica, 37*, 424–438.

Gurun, U. G., Johnston, R., & Markov, S. (2011). *Sell-side debt analysts and market efficiency* (Working paper).

Hai, W., Mark, N., & Wu, Y. (1997). Understanding spot and forward exchange rate regressions. *Journal of Applied Econometrics, 12*, 715–734.

Hamilton, J. D. (1994). *Time series analysis*. Princeton: Princeton University Press.

Hasbrouck, J. (1991). Measuring the information content of stock trades. *Journal of Finance, 46*, 179–207.

Hasbrouck, J. (1993). Assessing the quality of a security market: A new approach to transaction-cost measurement. *Review of Financial Studies, 6*, 191–212.

Hodrick, R. (1992). Dividend yields and expected stock returns: Alternative procedures for inference and measurement. *Review of Financial Studies, 5*, 357–386.

Hotchkiss, E. S., & Ronen, T. (2002). The informational efficiency of the corporate bond market: An intraday analysis. *Review of Financial Studies, 15*, 1325–1354.

Kwan, S. H. (1996). Firm-specific information and the correlation between individual stocks and bonds. *Journal of Financial Economics, 40*, 63–80.

Latane, H., & Rendleman, R. J. (1976). Standard deviation of stock price ratios implied in option prices. *Journal of Finance, 31*, 369–381.

Lee, B.-S. (1996). Time-series implications of aggregate dividend behavior. *Review of Financial Studies, 9*, 589–618.

Lee, B.-S. (1998). Permanent, temporary and non-fundamental components of stock prices. *Journal of Financial and Quantitative Analysis, 33*, 1–32.

Lutkepohl, H. (1991). *Introduction to multiple time series analysis*. Berlin: Springer.

Manaster, S., & Rendelman, R. J. (1982). Option prices as predictors of equilibrium stock prices. *Journal of Finance, 37*, 1043–1057.

Mankiw, N. G., Romer, D., & Shapiro, M. D. (1991). Stock market forecastability and volatility: A statistical appraisal. *Review of Economic Studies, 58*, 455–477.

Mark, N. (1995). Exchange rates and fundamentals: Evidence on long-horizon predictability. *American Economic Review, 85*, 201–218.

Mark, N., & Sul, D. (2001). Nominal exchange rates and monetary fundamentals: Evidence from a small post-Bretton Woods panel. *Journal of International Economics, 53*, 29–52.

Mark, N., & Wu, Y. (1998). Rethinking deviations from uncovered interest parity: The role of covariance risk and noise. *Economic Journal, 108*, 1686–1706.

Meese, R. A., & Rogoff, K. (1983). Empirical exchange rate models of the seventies: Do they fit out of sample? *Journal of International Economics, 14*, 3–24.

Pan, J., & Poteshman, A. (2006). The information in option volume for stock prices. *Review of Financial Studies, 19*, 871–908.

Ronen, T., & Zhou, X. (2012). Trade and information in the corporate bond market. *Journal of Financial Markets, 16*, 61–103.

Shiller, R. (1984). Stock prices and social dynamics. *Brookings Papers on Economic Activity, 2*, 457–498.

Sims, C. A. (1972). Money, income, and causality. *American Economic Review, 62*, 540–552.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica, 48*, 1–48.

Stephan, J., & Whaley, R. (1990). Intraday price changes and trading volume relations in the stock and stock option markets. *Journal of Finance, 45*, 191–220.

Stock, J., & Watson, M. (2001). Vector autoregressions. *Journal of Economic Perspectives, 15*, 101–115.

Viceira, L. (2001). Optimal portfolio choice for long-horizon investors with nontradable labor income. *Journal of Finance, 56*, 433–470.

Vijh, A. M. (1988). Potential biases from using only trade prices of related securities on different exchanges. *Journal of Finance, 43*, 1049–1055.

Watson, M. (1994). Vector autoregressions and cointegration. In R. F. Engle & D. McFadden (Eds.), *Handbook of econometrics* (Vol. IV). Amsterdam: Elsevier.

Zhou, X. (2009). *Information-based trading in the junk bond market* (Working paper), Cornell University.

Model Selection for High-Dimensional Problems

77

Jing-Zhi Huang, Zhan Shi, and Wei Zhong

Contents

Abstract

High-dimensional data analysis is becoming more and more important to both academics and practitioners in finance and economics but is also very challenging because the number of variables or parameters in connection with such data can be larger than the sample size. Recently, several variable selection approaches

J.-Z. Huang • Z. Shi
Smeal College of Business, Penn State University, University Park, PA, USA
e-mail: jxh56@psu.edu; zus116@psu.edu

W. Zhong (✉)
Wang Yanan Institute for Studies in Economics and Department of Statistics, School of Economics, Xiamen University, Xiamen, China
e-mail: wzhong@xmu.edu.cn

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_77,

have been developed and used to help us select significant variables and construct a parsimonious model simultaneously. In this chapter, we first provide an overview of model selection approaches in the context of penalized least squares. We then review independence screening, a recently developed method for analyzing ultrahigh-dimensional data where the number of variables or parameters can be exponentially larger than the sample size. Finally, we discuss and advocate multistage procedures that combine independence screening and variable selection and that may be especially suitable for analyzing high-frequency financial data.

Penalized least squares seek to keep important predictors in a model while penalizing coefficients associated with irrelevant predictors. As such, under certain conditions, penalized least squares can lead to a sparse solution for linear models and achieve asymptotic consistency in separating relevant variables from irrelevant ones. Independence screening selects relevant variables based on certain measures of marginal correlations between candidate variables and the response.

Keywords

Model selection • Variable selection • Dimension reduction • Independence screening • High-dimensional data • Ultrahigh-dimensional data • Generalized correlations • Penalized least squares • Shrinkage • Statistical learning • LASSO • SCAD penalty • Oracle property

77.1 Introduction

High-dimensional data analysis has now become increasingly frequent and necessary in various research fields, such as finance, genetics, computer science, and geography. For example, the price of a stock may depend on a huge number of variables, such as the company's dividend yields and price-to-earnings ratios, past values of the stock, related bond and derivative prices, other relevant companies' stock prices, overall market information, and some other macroeconomic factors. How to select significant variables for the price of this stock is important and yet challenging given the high-dimensional space of potential predictors. Another example is the estimation of covariance matrix in portfolio problem. Suppose we want to construct an optimal portfolio using, say, 200 stocks. In this case, there are 20,100 parameters in the covariance matrix to estimate. As the number of stocks to be included in the analysis increases, the number of such parameters to be estimated will increase dramatically. It is also challenging to estimate this large covariance matrix using the traditional statistical methods. For more examples, see Fan and Li (2006), Hastie et al. (2009), and Fan et al. (2011b).

One method often used in the literature to make high-dimensional problems tractable is to assume the sparsity condition – namely, only a small number of predictors are assumed to contribute to the response variable. Under the sparsity principle, variable selection approaches play a fundamental role in model selection. For instance, traditional subset selection methods combined with variable selection criteria, such as the Akaike (1973) information criterion (AIC) and the Bayesian information criterion (Schwartz 1978), may perform well in the relatively low dimensional space. However, these methods suffer from expensive computational cost and model instability for high-dimensional data (Breiman 1996). Recently, various regularization methods have been proposed for variable selection in high-dimensional data analysis. Examples include the least absolute shrinkage and selection operator (LASSO) of Tibshirani (1996), the smoothly clipped absolute deviation (SCAD) (Fan and Li 2001; Zou and Li 2008), the least-angle regression (LARS) algorithm (Efron et al. 2004), the elastic net (Zou and Hastie 2005; Zou and Zhang 2009), the adaptive LASSO (Zou 2006), and the Dantzig selector (Candes and Tao 2007), among others. With the sparsity assumption, regularization methods can improve both the prediction accuracy and the model interpretability in the high-dimensional setting. Some of these newly proposed variable selection methods have been used in economics and finance recently, in areas such as macroeconomic forecasting (Bai and Ng 2008), term structure of interest rates (Huang and Shi 2010), and portfolio choice (Goto and Xu 2010).

Nonetheless, one very important challenge facing both researchers and practitioners nowadays is to better understand ultrahigh-dimensional data, where the number of predictors, say, p, is usually much larger than the sample size, say, n ($p \gg n$) – namely, $\log p = O(n^{\alpha})$ for some $\alpha > 0$. Issues such as computational expediency, statistical accuracy, and algorithmic stability call for new statistical modeling techniques for dimension reduction and model selection (Fan et al. 2009). One such idea is a two-stage procedure: In the first stage (the "screening" stage), some kind of screening is used to reduce the dimension of the original data set; in the next stage (the "cleaning" stage), well-established variable selection methods are used to simultaneously select significant variables and estimate statistical effects of those selected variables. Ji and Jin (2012) show that under certain settings, the two-stage approach has advantages over the one-stage LASSO and subset selection in identifying important predictors in the ultrahigh-dimensional analysis.

One way to implement the first step is to use the independence screening framework developed recently in the theoretical statistics literature. Specifically, at the screening stage, independence screening procedures are used to remove as many irrelevant variables as possible in order to reduce the dimensionality from ultrahigh p to a relatively large-scale d that may be less than n. As such, the independence screening procedures can dramatically reduce the computational complexity. Fan and Lv (2008) propose the sure independence screening (SIS) via Pearson correlation ranking to effectively narrow down the ultrahigh dimensionality to a moderate scale, in which the aforementioned regularization methods can be applied. The authors further show that the SIS enjoys a sure screening property within the context of linear regressions with Gaussian predictors and responses. That is, all truly important predictors can be selected with probability approaching one as the sample size goes to the infinity. Hall and Miller (2009) use the generalized correlation learning to capture nonlinear dependence between predictors and the response. Fan et al. (2009) and Fan and Song (2010) consider a more general version of independent learning that ranks the maximum marginal likelihood estimators or the maximum marginal likelihood for generalized linear models. Some other studies are based on nonparametric methods. For instance, Fan et al. (2011a) propose independence screening based on spline approximation in sparse ultrahigh-dimensional additive models; Zhu et al. (2011) use a model-free sure independent ranking and screening (SIRS) to select important predictors in a general class of multi-index models; and Li et al. (2012) study a unified screening framework via distance correlation learning.

It is also worth mentioning that these screening procedures can possess some favorable theoretical properties, such as the sure screening property (Fan and Lv 2008) and the ranking consistency property (Zhu et al. 2011).

We should also note that two important methods in the literature albeit not focused in this review are factor models and principal component analysis (PCA). Factor models are widely used in unsupervised learning as a means of dimension reduction. The factor structure provides a sequence of best linear approximations to multivariate data set, by seeking nonredundant components that are as statistically independent as possible. In the Gaussian setting, it suffices to find linear orthogonal components, as high-order cross dependence is determined by second moments alone. Consequently, PCA becomes a standard procedure in multivariate analysis and is described in any textbooks on multivariate analysis (see, e.g., Mardia et al. 1980). Extensions of PCA include principal curves and surfaces (Hastie 1994), curvilinear component analysis (Demartines and Hérault 1997), and independent component analysis (Comon 1994). The factor approach has been used widely in finance and economics. For instance, it has been applied to macroeconomic forecasting (Stock and Watson 2002) and asset volatility prediction (Anderson and Vahid 2007). See Campbell et al. (1997) for a more complete reference.

The rest of this chapter is organized as follows. Section 77.2 reviews well-established variable selection methods within the framework of penalized least squares. In Sect. 77.3, we provide a brief review of the existing independence screening methods and call for application in financial studies. Section 77.4 discusses model selection versus dimension reduction. Section 77.5 concludes.

77.2 Variable Selection Approaches

Variable selection techniques play an increasingly important role in the high-dimensional problems. Here, the high dimensionality means that $p = O(n^{\alpha})$ with $0 < \alpha < 1$. In the early stages of statistical modeling, it is typical to include as many as potential influential predictors into a model in order to reduce possible model bias. Nonetheless, it is natural to assume that only a subset of predictors contribute to the response in the true model. Under this sparsity assumption, variable selection can improve both the prediction accuracy and the interpretability of the fitted model.

Consider the following linear regression model:

$$y = \mathbf{X}\beta + \varepsilon, \tag{77.1}$$

where $\mathrm{y} = (Y_1, \ldots, Y_n)^T$ is an $n \times 1$ vector of responses, $\mathbf{X} = (\mathrm{x}_1, \ldots, \mathrm{x}_n)^T$ is an $n \times p$ random design matrix, $\beta = (\beta_1, \ldots, \beta_p)^T$ is a $p \times 1$ vector of parameters, and $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)^T$ is an $n \times 1$ vector of independent and identically distributed (i.i.d.) random errors.

When the dimension parameter p is large, it is natural to assume that the model is sparse – namely, only a small subset of predictors, say, true predictors $\{X_j : \beta_j \neq 0, j = 1, \ldots, p\}$, contribute to the response y. We can use variable selection techniques to help identify true predictors.

77.2.1 Classic Variable Selection Criteria

A variable selection criterion is a statistic of a fitted model that measures the goodness of fit. The literature on this issue is extensive. Two well-known examples are the Akaike (1973) information criterion (*AIC*) and the Bayesian information criterion (*BIC*) developed by Schwartz (1978). Another example is the generalized cross-validation statistic (*GCV*) proposed by Craven and Wahba (1979). Also, Shao (1997) discusses the consistency and efficiency of variable selection, and Miller (2002) provides a comprehensive review of the subset selection in regression.

Below we review some widely used variable selection criteria:

- *Residual sum of squares* (*RSS*). For the linear regression model (77.1), many variable selection criteria are built on the residual sum of squares (*RSS*) that is defined as follows:

$$RSS = \left\| \mathrm{y} - \mathrm{X}\hat{\beta} \right\|^2 = \sum_{i=1}^{n} \left(Y_i - \mathrm{x}_i \hat{\beta} \right)^2, \tag{77.2}$$

where $\hat{\beta}$ is an estimate of β. Because $\mathrm{x}_i\hat{\beta}$ is the fitted value of the ith observation Y_i, *RSS* can measure the goodness of model fit.

- R^2 *and adjusted* R^2. R^2 is a commonly used statistic for model fitting and is related to *RSS*. Specifically, we have

$$R_d^2 = 1 - \frac{RSS_d}{RSS_0}, \tag{77.3}$$

where RSS_d is the residual sum of squares when an intercept and d predictors are fitted in the model, where $1 \leq d \leq p$, and RSS_0 is the *RSS* with only the intercept fitted. R^2 can measure how well the fitting of the d predictors is. However, it is known that R^2 increases with the number of predictors used in the model. Therefore, R^2 cannot serve as a variable selection criterion. The adjusted R^2 addresses this concern of R^2 by penalizing the number of predictors used. The more predictors used, the higher the penalty.

The adjusted R^2 is also called Fisher's A-statistic. The latter is defined as follows:

$$\begin{aligned} A_d &= 1 - \left(1 - R_d^2\right)\frac{n-1}{n-d} \\ &= 1 - \frac{RSS_d/(n-d)}{RSS_0/(n-1)}. \end{aligned} \tag{77.4}$$

Fisher's A-statistic A_d does not necessarily increase when a new predictor is added to the model. Therefore, it can serve as a variable selection criterion for model fitting.

- *Prediction sum of squares* (*PRESS*). This is a prediction-based variable selection criterion proposed by Allen (1974). When a model includes d predictors, $PRESS_d$ is defined as follows:

$$PRESS_d = \sum_{i=1}^{n} \left(Y_i - \widehat{Y}_{id}\right)^2, \tag{77.5}$$

  where $\widehat{Y}_{id}$ is the predicted value of Y_i from the fitted model using all observations but the ith one.
- *Cross-validation* (*CV*). Cross-validation can be considered to be a generalization of *PRESS*. The idea of *CV* is that we randomly set a small number of observations (the testing set) aside, use the remaining observations (the training set) to fit the model, predict the testing data set, and then summarize the performance of the prediction.

  If we set only one observation aside each time, we obtain the so-called leave-one-out *CV*, that is, essentially the *PRESS*. In practice, we can partition the data set into K equivalent subsets, leave one subset out each time, and predict this subset using the remaining $K - 1$ ones. Both *CV* and *PRESS* can be used to estimate the prediction errors of the fitted model and provide a good measure of how well the prediction of the proposed model is.

  A related statistic is the generalized cross-validation statistic (*GCV*) for linear regression models proposed by Craven and Wahba (1979) and is defined by the following:

$$GCV = \frac{RSS_d/n}{(1 - d/n)^2}. \tag{77.6}$$

  It is shown that under the mild conditions, if n is much larger than d, the *PRESS* statistic can be asymptotically approximated by

$$\begin{aligned} PRESS_d &\approx \frac{n^2}{(n-d)^2} RSS_d \\ &= \frac{RSS_d}{(1 - d/n)^2} = nGCV. \end{aligned} \tag{77.7}$$

Therefore, GCV is also a widely used variable selection criterion.

- *Akaike information criterion* (AIC). Proposed by Akaike (1973), AIC is related to the Kullback-Leibler mean information and is defined as follows:

$$AIC = RSS_d + 2d\sigma^2. \tag{77.8}$$

It is known that AIC is equivalent to Mallows' C_p (Mallows 1973) measure in linear regression models. For such a model with d predictors, Mallows' C_p of the model is given by the following:

$$C_p = \frac{RSS_d}{\sigma^2} - (n - 2d). \tag{77.9}$$

- *Bayesian information criterion* (BIC). The BIC, suggested by Schwartz (1978), is defined by the following:

$$BIC = RSS_d + \log(n)d\sigma^2. \tag{77.10}$$

In practice, we choose a model with the smallest information criterion to achieve variable selection. It can be shown that the BIC is a consistent criterion. That is, when assuming there exists a true model with finite parameters, the BIC can determine the true model as the sample size approaching the infinity. However, the AIC may provide an overfitted model. On the other hand, the AIC is an asymptotically less efficient criterion (Shao 1997), but the BIC is not.

77.2.2 Penalized Least Squares

Although the best subset selection with classic variable selection criteria can perform well in practice, the method suffers from the highly expensive computational cost, especially for the high-dimensional regression models. Furthermore, the subset selection approaches lack stability, and their theoretical properties are difficult to examine (Breiman 1996).

For linear regression models (77.1), penalized least squares (PLS) methods provide one alternative that can overcome the aforementioned limitations of the subset selection approaches. Consider the following objective function $Q(\beta)$:

$$Q(\beta) = \frac{1}{2}\|\mathrm{y} - \mathbf{X}\beta\|^2 + n\sum_{j=1}^{p} p_\lambda\left(|\beta_j|\right), \tag{77.11}$$

where $p_\lambda(\cdot)$ is the penalty function and λ is the regularity parameter that controls the size of the penalty. We can obtain the PLS estimate $\hat{\beta}^{PLS}$ by minimizing the objective function $Q(\beta)$. If there is no penalty, then we recover the objective function of ordinary least squares and obtain the OLS estimate $\hat{\beta}^{OLS}$.

In the balance of this subsection, we will discuss some well-known penalty functions as well as how to choose a good penalty function. Moreover, we will provide connections between the penalized least squares (77.11) and the classic best subset selection and the ridge regression.

77.2.2.1 L_q Penalties with $0 \leq q \leq 2$

- **L_0 Penalty: Best Subset Selection**

The best subset selection with classic variable selection criteria can be written as the form of PLS with some L_0 penalty functions. Notice that choosing a model with the minimum C_p is equivalent to minimizing the following PLS objective function:

$$Q(\beta) = \frac{1}{2}\|\mathrm{y} - \mathbf{X}\beta\|^2 + \sigma^2 \sum_{j=1}^{p} I\left(\left|\beta_j\right| \neq 0\right). \tag{77.12}$$

Motivated by (77.12), the best subset selection with classic variable selection criteria is equivalent to minimizing the objective function (77.11) with the following L_0 penalty function:

$$p_\lambda\left(\left|\beta_j\right|\right) = \frac{\lambda^2}{2} I\left(\left|\beta_j\right| \neq 0\right) \tag{77.13}$$

with a different tuning parameter λ.

For example, *AIC* is asymptotically equivalent to the following PLS:

$$Q(\beta) = \frac{1}{2}\|\mathrm{y} - \mathbf{X}\beta\|^2 + n\frac{\left(\sigma\sqrt{2/n}\right)^2}{2} \sum_{j=1}^{p} I\left(\left|\beta_j\right| \neq 0\right) \tag{77.14}$$

with $\lambda = \sigma\sqrt{2/n}$. Another example is *BIC*, which is asymptotically equivalent to the PLS as follows:

$$Q(\beta) = \frac{1}{2}\|\mathrm{y} - \mathbf{X}\beta\|^2 + n\frac{\left(\sigma\sqrt{\log(n)/n}\right)^2}{2} \sum_{j=1}^{p} I\left(\left|\beta_j\right| \neq 0\right) \tag{77.15}$$

with $\lambda = \sigma\sqrt{\log(n)/n}$.

- **L_2 Penalty: Ridge Regression**

The well-known ridge regression is proposed by Hoerl and Kennard (1970) in order to deal with the collinearity problem in predictors. Although ridge regression cannot possess the variable selection feature, it is a solution of penalized least squares (77.11) with an L_2 penalty – namely, $p_\lambda\left(\left|\beta_j\right|\right) = \frac{\lambda}{2}\left|\beta_j\right|^2$. Therefore, the ridge regression estimates can be obtained by minimizing the following PLS objective function:

$$Q(\beta) = \frac{1}{2}\|\mathrm{y} - \mathbf{X}\beta\| + \frac{n\lambda}{2}\sum_{j=1}^{p}|\beta_j|^2. \tag{77.16}$$

Like the ordinary least squares, the ridge regression also has the explicit solution:

$$\hat{\beta} = \left(\mathbf{X}^{\mathrm{T}}\mathbf{X} + n\lambda I_p\right)^{-1}\mathbf{X}^{\mathrm{T}}\mathrm{y}, \tag{77.17}$$

where I_p is a $p \times p$ identity matrix.

- ***L_q* Penalty: Bridge Regression**

This regression is proposed by Frank and Friedman (1993) and corresponds to the following objective function:

$$Q(\beta) = \frac{1}{2}\|\mathrm{y} - \mathbf{X}\beta\|^2 + \frac{n\lambda}{q}\sum_{j=1}^{p}|\beta_j|^q, \tag{77.18}$$

where $0 < q < 2$. The penalty function used here, the L_q penalty, bridges the L_0 and L_2 penalties. The current literature focuses on bridge regressions with $0 < q < 1$ (Friedman 2008; Zou and Li 2008), as they perform well in the presence of noised surrogates correlated with a true explanatory variable. In these situations, the penalty function is concave over $(0, \infty)$. In the appendix, we discuss the thresholding operator induced by the $L_{0.5}$ penalty and investigate its implementation based on local linear approximation.

- ***L_1* Penalty: LASSO**

The least absolute shrinkage and selection operator (LASSO) is first proposed by Tibshirani (1996). Specifically, the LASSO estimate is the solution to the following problem:

$$\begin{gathered}\min_{\beta}\frac{1}{2}\|\mathrm{y} - \mathbf{X}\beta\|^2,\\ \text{subject to}\ \sum_{j=1}^{p}|\beta_j| \le s,\end{gathered} \tag{77.19}$$

where the tuning parameter s controls the regularization size. The main advantage of LASSO is that it allows us to shrink regression coefficients and select significant predictors simultaneously.

The LASSO solutions are shown to be equivalent to the penalized least squares with the L_1 penalty as follows:

$$Q(\beta) = \frac{1}{2}\|\mathrm{y} - \mathbf{X}\beta\|^2 + n\lambda\sum_{j=1}^{p}|\beta_j|. \tag{77.20}$$

Therefore, the LASSO is a special case of the bridge regression with $q = 1$. It can be shown that the LASSO can exactly shrink some coefficients to zero and hence gives a sparse model, which enhances the model interpretability.

Although penalized L_0 regressions can be used to conduct variable selection, the computation involved is usually extensive, and the set of selected predictors can be unstable. Penalized L_2 regressions (ridge regressions) can shrink the estimated coefficients and make the model stable, but these regressions do not possess the variable selection feature. Penalized L_1 regressions (LASSO) can provide both shrinkage estimation and variable selection, but the estimators are biased even for large true coefficients. The natural question is then: "What kind of penalty functions is good for variable selection and parameter estimation?"

77.2.2.2 The SCAD Penalty

Fan and Li (2001) argue that good penalty functions should provide the estimators with the following three properties in the high-dimensional regression problems:

1. **Unbiasedness**: The penalized estimator should be nearly unbiased to reduce model bias, especially for the large true coefficients.
2. **Sparsity**: The penalized estimator can automatically set small estimated coefficients to zero to achieve variable selection and reduce model complexity.
3. **Continuity**: The penalized estimator is continuous in the data in the sense that it can avoid instability in model prediction.

To proceed, Fan and Li (2001) consider the linear regression model (77.1) with the design matrix $\mathbf{X}$ satisfying $\mathbf{X}^{\mathbf{T}}\mathbf{X} = nI_p$, where I_p is a $p \times p$ identity matrix. It then follows that (77.11) can be rewritten as the following:

$$\begin{aligned} Q(\beta) &= \frac{1}{2}\left\|\mathrm{y} - \mathrm{X}\hat{\beta}_0\right\|^2 + \frac{n}{2}\left\|\hat{\beta}_0 - \beta\right\|^2 + n\sum_{j=1}^{d} p_\lambda(|\beta_j|), \\ &= \frac{1}{2}\left\|\mathrm{y} - \mathbf{X}\hat{\beta}_0\right\|^2 + n\sum_{j=1}^{d}\left\{\frac{1}{2}\left(\hat{\beta}_{0j} - \beta_j\right)^2 + p\lambda(|\beta_j|)\right\}, \end{aligned} \tag{77.21}$$

where $\hat{\beta}_0 = \mathbf{X}^{\mathrm{T}}\mathrm{y}/n$ is the ordinary least squares estimate. The first term of (77.21) is constant with respect to β, so minimizing the object $Q(\beta)$ reduces to a componentwise regression problem. Consider the univariate minimization problem:

$$\hat{\theta}(z) = \arg\min_{\theta\in\mathrm{R}}\left\{\frac{1}{2}(z-\theta)^2 + p\lambda(|\theta|)\right\}. \tag{77.22}$$

Antoniadis and Fan (2001) and Fan and Li (2001) examine the conditions under which the univariate penalized estimator $\hat{\theta}(z)$ can possess the above three properties:

1. **Unbiasedness** if $p'_\lambda(\theta) = 0$ for large θ
2. **Sparsity** if $\min_{\theta \geq 0}\{\theta + p'_\lambda(\theta)\} > 0$
3. **Continuity** if and only if $\arg\min_{\theta \geq 0}\{\theta + p'_\lambda(\theta)\} = 0$

where $p_\lambda(\theta)$ is nondecreasing and continuously differentiable on $[0, \infty)$ and $p'_\lambda(0)$ means $p'_\lambda(0+)$ here. In general, a good penalty function $p_\lambda(\theta)$ should be singular at the origin to generate sparse estimators in variable selection and concave when θ is large to reduce the model bias.

It has been shown in the literate that an L_q penalty does not satisfy the unbiasedness condition. As such, such a penalty function increases the model bias. To construct a penalty function satisfying all the three conditions mentioned above, Fan and Li (2001) introduce the smoothly clipped absolute deviation (SCAD) penalty, whose first derivative is given by the following:

$$p'_\lambda(\theta) = \lambda\left\{I(\theta \le \lambda) + \frac{(a\lambda - \theta)_+}{(a-1)\lambda} I(\theta > \lambda)\right\}, \tag{77.23}$$

where $\theta > 0$, $p_\lambda(0) = 0$, and $a > 2$. Here, $a = 3.7$ is often suggested by a Bayesian argument. Another penalty function, called minimax concave penalty (MCP) by Zhang (2010), has the same spirit as SCAD, and its first derivative is given by the following:

$$p'_\lambda(\theta) = \frac{(a\lambda - \theta)_+}{a}. \tag{77.24}$$

For more theoretical and numerical studies of SCAD and MCP, see Fan and Li (2001), Zou and Li (2008), and Zhang (2010).

77.2.2.3 The Oracle Property

In addition to the aforementioned properties, another equally important property established in Theorem 2 of Fan and Li (2001) is the oracle property. Ensuring the optimal asymptotic performance, the oracle property entitles statistician the power to work as well as if they had an oracle.

Specifically, assume that the true parameter vector $\beta_0 = [\beta_{10}^T, \beta_{20}^T]^T$ and $\beta_{20}^T = 0$, and under some regularity conditions, the $\sqrt{n}$-consistent local minimizer $\hat\beta = \left[\hat\beta_1^T, \hat\beta_2^T\right]^T$ of the above $Q(\beta)$ with SCAD penalty satisfies the following conditions:

- Sparsity: $\hat\beta_2 = 0$.
- Asymptotic normality:

$$\sqrt{n}(I_1(\beta_{10}) + \Sigma)\left\{\hat\beta_1 - \beta_{10} + (I_1(\beta_{10}) + \Sigma)^{-1}\mathrm{b}\right\} \to N(0, I_1(\beta_{10})),$$

in distribution, where $\Sigma = \mathrm{diag}\{p''_\lambda(|\beta_{10}|)\}$, $I_1(\beta_{10}) = I_1(\beta_{10}, 0)$ is the Fisher information with knowing $\beta_{20} = 0$ in advance, and $\mathrm{b} = \mathrm{sgn}(\beta_{10}) \circ p'_\lambda(|\beta_{10}|)$ with $\circ$ denoting the componentwise product.

As shown in Fan and Li (2001), hard thresholding and SCAD have oracle property when $\lambda_n \to 0$ and $\sqrt{n}\lambda_n \to \infty$, and are thus more efficient than least squares directly applied to the full model if the model is sparse. The authors further conjecture that the LASSO may not have the oracle property because of the bias problem. This conjecture is recently proven in Zou (2006), who also proposes a variation of the LASSO – the adaptive LASSO – to make oracle property attainable for the L_1 penalty.

The basic idea underlying the adaptive LASSO is adding the adaptive weights for penalizing different coefficients in the L_1 penalty. Specifically, the objective function of the adaptive LASSO is given as follows:

$$Q(\beta) = \frac{1}{2}\|\mathrm{y} - \mathbf{X}\beta\|^2 + n\lambda\sum_{j=1}^{p}\hat{w}_j|\beta_j|, \tag{77.25}$$

where $\{\hat{w}_j\}_{j=1}^{p}$ can be estimated by $\hat{w}_j = \left(\left|\widetilde{\beta}_j\right|\right)^{-\gamma}$ with any root-n-consistent estimates $\widetilde{\beta}$. Zou (2006) shows that with an appropriately chosen λ, the adaptive LASSO performs as well as the oracle.

Another issue with L_1 penalty concerns its difficulty in dealing with highly correlated variables in the predictor set. In their simulation study, Zou and Hastie (2005) show that the LASSO solution paths are unstable in the presence of multicollinearity. To fix this problem, these authors propose another generalization of the LASSO via combining L_1 penalty and L_2 penalty, called elastic net (EN), that is given by the following:

$$Q(\beta) = \frac{1}{2}\|\mathrm{y} - \mathbf{X}\beta\|^2 + n\lambda_1\sum_{j=1}^{p}|\beta_j| + n\lambda_2\sum_{j=1}^{p}\beta_j^2. \tag{77.26}$$

This method can be used to encourage a grouping effect. Furthermore, Zou and Hastie (2005) suggest that the estimator $\hat{\beta}^{enet} = \left(1 + \frac{\lambda_2}{n}\right)$ argmin $Q(\beta)$ be used in order to correct for the double-shrinkage problem in the implementation of the elastic net method.

Notice that in Eq. 77.26, the L_1 part performs automatic variable selection, while the L_2 part stabilizes the solution paths and hence improves the prediction performance. In order for the elastic net to have the oracle property, Zou and Zhang (2009) propose the adaptive elastic net as a combination of the L_2 penalty and the adaptive L_1 penalty. They show that the adaptive elastic net enjoys the oracle property under the assumption that the model dimension diverges with the sample size.

77.2.3 Computational Algorithms

So far we have reviewed a number of representative model selection methods used in the literature. In this subsection, we discuss the implementation of some of these methods:

- **The LQA algorithm**. When the convex penalty function (e.g., the L_1 penalty) is used, the objective function (77.11) is convex, and hence convex optimization algorithms can be applied. However, some penalty functions (e.g., the SCAD penalty) are used, and then the object is not convex anymore. Fan and Li (2001) propose a unified and effective local quadratic approximation (LQA) algorithm for optimizing a nonconvex penalized objective function.

The idea is to use the quadratic curve to locally approximate the objective function. To be specific, for a given initial value $\beta_0 = (\beta_0, \ldots, \beta_{p0})^T$ which is not close to 0, the penalty function $p_\lambda(\cdot)$ can be locally approximated by a quadratic function as follows:

$$[p_\lambda(|\beta_j|)]' = p'_\lambda(|\beta_j|)\mathrm{sgn}(\beta_j) \approx \{p'_\lambda(|\beta_{j0}|)/|\beta_{j0}|\}\beta_j, \quad \text{for} \quad \beta_j \approx \beta_{j0}. \tag{77.27}$$

In another word, we have

$$p_\lambda(|\beta_j|) \approx p_\lambda(|\beta_{j0}|) + \frac{1}{2}\frac{p'_\lambda(|\beta_{j0}|)}{|\beta_{j0}|}\left(\beta_j^2 - \beta_{j0}^2\right), \quad \text{for} \quad \beta_j \approx \beta_{j0}. \tag{77.28}$$

With the LQA, the objective function (77.11) with nonconvex penalty becomes a convex function and admits a closed-form solution. The LQA algorithm sets the sufficiently small coefficients to zero and hence produces a sparse model. However, a drawback of this algorithm is that once a coefficient is shrunken to zero, it will remain to be zero in subsequent iterations.

- **The LLA algorithm**. Instead of using LQA, Zou and Li (2008) suggest a better approximation by using the local linear approximation (LLA) as follows:

$$p_\lambda(|\beta_j|) \approx p_\lambda(|\beta_{j0}|) + p'_\lambda(|\beta_{j0}|)(|\beta_j| - |\beta_{j0}|), \quad \text{for} \quad \beta_j \approx \beta_{j0}. \tag{77.29}$$

The LLA is the minimum convex majorant of the concave function on $[0, \infty)$. With the LLA, the objective function (77.11) with a nonconvex penalty becomes an iteratively reweighted penalized L_1 regression. Zou and Li (2008) show that the one-step LLA estimator naturally adopts a sparse representation and enjoys the oracle properties. In the appendix, we would illustrate the execution of this one-step sparse estimator with the $L_{0.5}$ and SCAD penalty.

- **LARS algorithm**. Efron et al. (2004) develop the least-angle regression (LARS) algorithm for penalized variable selection. This fast and efficient algorithm can produce the entire LASSO solution path $\left\{\hat{\beta}(\lambda), \lambda > 0\right\}$, which is piecewise linear in λ. See Efron et al. (2004) for details.

77.3 Independence Screening Procedures

Modern technology for data collection allows researchers to collect ultrahigh-dimensional data at relatively low cost in economics, finance, and scientific fields. Here, ultrahigh dimensionality means that the number of predictors (p) is highly greater than the number of observations (n). Specifically, $p = O(\exp(\alpha n))$ with $\alpha > 0$.

That is, the dimensionality p is allowed to increase as the sample size n at the exponential rate. However, for such ultrahigh-dimensional data, the aforementioned regularization methods may fail because of issues on computational cost, statistical accuracy, and algorithmic stability (Fan et al. 2009). To solve these problems, statisticians and econometricians proposed a two-stage screening and cleaning approach. That is, an independence screening procedure is applied first and then followed by regularization methods. Ji and Jin (2012) further theoretically demonstrate that under some regularity conditions, the two-stage approach can outperform the one-stage LASSO and subset selection. In this section, we briefly review the recent developments of independence screening procedures for ultrahigh-dimensional data.

77.3.1 Sure Independence Screening

For ultrahigh-dimensional linear regression model, Fan and Lv (2008) propose the sure independence screening (SIS) via Pearson correlation learning to reduce the ultrahigh dimension down to a relative large scale.

Consider the following linear regression model, as defined earlier in Eq. (77.1):

$$\mathrm{y} = \mathbf{X}\beta + \varepsilon.$$

Under the sparsity assumption, denote the true model as $\mathcal{M}_* = \{1 \le j \le p : \beta_j \ne 0\}$ with the model size $s = |\mathcal{M}_*|$, where $|\mathcal{M}_*|$ represents the number of elements in the set $\mathcal{M}_*$. Then denote the standardized columnwise design matrix as $\mathbf{X}_s$ and define $\omega = (\omega_1, \ldots, \omega_p)^{\mathrm{T}}$ as follows:

$$\omega = \mathbf{X}_s^{\mathrm{T}} \mathrm{y}. \tag{77.30}$$

Note that ω_j is the marginal Pearson correlation between the jth predictor X_j and the response Y scaled by its standard deviation. On the other hand, ω_j can also be considered as the least squares estimated coefficient for standardized X_j in the marginal regression $\mathrm{y} = X_j\beta_j + \varepsilon$. Therefore, $|\omega_j|$ can characterize the magnitude of marginal relationship between the predictor X_j and the response Y.

The SIS ranks the importance of all predictors according to $|\omega_j|$ and removes those predictors weakly correlated with the response Y, i.e., ones with small absolute values of ω_j. To be specific, for any given $\gamma \in (0,1)$, the SIS selects predictors with the first $[\gamma n]$ largest $|\omega_j|$ and defines the submodel

$$\hat{\mathcal{M}}_\gamma = \left\{1 \le j \le p : \left|\omega_j\right| \text{ is among the first } [\gamma n] \text{ largest of all}\right\},$$

where $[\gamma n]$ denotes the integer part of γn. Under some regularity conditions, Fan and Lv (2008) show that

$$P\left(\mathcal{M}_* \subseteq \hat{\mathcal{M}}_\gamma\right) \to 1, \quad \text{as } n \to \infty. \tag{77.31}$$

That is, the proposed SIS can efficiently shrink the ultrahigh dimension p down to a relatively large-scale $d = [\gamma n]$, while all truly important predictors can be selected into the submodel $\hat{\mathcal{M}}_\gamma$ with probability approaching one as the sample size tends to the infinity. This desirable theoretical property is called sure screening property by Fan and Lv (2008).

77.3.2 Generalized Correlation Ranking

Sure independence screening via Pearson correlation learning can perform well in the ultrahigh-dimensional linear regression model. However, Pearson correlation can only capture the linear relationship between each predictor X_j and the response Y. When Pearson correlation $\rho(X_j, Y)$ is zero, it only means that the response Y is linearly uncorrelated with the predictor X_j. If the predictor X_j is nonlinearly but not linearly influential to the response Y, the SIS is most likely to miss this important predictor. In order to capture the nonlinearity in the ultrahigh-dimensional problems, Hall and Miller (2009) suggest techniques based on ranking generalized empirical correlation between the response Y and each predictor X_j, which can capture both linearity and nonlinearity.

Hall and Miller (2009) define the generalized correlation between two random variables X and Y as follows:

$$\rho_g(X, Y) = \sup_{h \in \mathcal{H}} \frac{\operatorname{cov}\{h(X), Y\}}{\sqrt{\operatorname{var}\{h(X)\}\operatorname{var}(Y)}},$$

where $\mathcal{H}$ is a class of functions including all linear functions. For example, it is a class of polynomial functions up to a given degree. Notice that if $\mathcal{H}$ is restricted to be a class of all linear functions, $\rho_g(X, Y)$ is the absolute value of Pearson correlation $\rho(X, Y)$. Therefore, $\rho_g(X, Y)$ can be naturally considered as a generalization of the conventional Pearson correlation.

Assume that (X_1, Y_1), (X_2, Y_2), ... ,(X_n, Y_n) are independent and identically distributed observed pairs of two random variables X and Y. The generalized correlation $\rho_g(X, Y)$ between X and Y can be estimated as follows:

$$\hat{\rho}_g(X, Y) = \sup_{h \in \mathcal{H}} \frac{\sum_{i=1}^n \left\{h(X_i) - \overline{h}\right\}\left(Y_i - \overline{Y}\right)}{\sqrt{\sum_{i=1}^n \left\{h(X_i)^2 - \overline{h}^2\right\} \cdot \sum_{i=1}^n \left(Y_i - \overline{Y}\right)^2}}, \tag{77.32}$$

where $\overline{h} = n^{-1}\sum_{i=1}^n h(X_i)$ and $\overline{Y} = n^{-1}\sum_{i=1}^n Y_i$.

The proposed generalized correlation characterizes both linear and nonlinear relationships between two random variables. Therefore, the generalized correlation $\rho_g(X_j, Y)$ can be considered as a marginal utility to measure the influential effort of the predictor X_j on the response Y.

Hall and Miller (2009) suggest that in practice we rank the predictors based on the magnitude of estimated generalized correlation $\hat{\rho}_g(X_j, Y)$. As such, we order $\hat{\rho}_g\left(X_{\hat{j}_1}, Y\right) \geq \hat{\rho}_g\left(X_{\hat{j}_2}, Y\right) \geq \ldots \geq \hat{\rho}_g\left(X_{\hat{j}_p}, Y\right)$ and have

$$\hat{j}_1 \succeq \hat{j}_2 \succeq \cdots \succeq \hat{j}_p$$

denote the empirical ranking of the indices of all predictors. Intuitively, the higher ranking the predictor has, the more important it is on the response in terms of the generalized correlation. Therefore, given a suitable cutoff, one can select predictors with higher rankings and thus reduce the ultrahigh dimensionality to a relatively low scale.

Hall and Miller (2009) suggest a bootstrap procedure to choose a cutoff. To illustrate this procedure, let $S = \{(X_1, Y_1), (X_2, Y_2), \ldots, (X_n, Y_n)\}$ represent the original data set and $S^* = \{(X_1^*, Y_1^*), (X_2^*, Y_2^*), \ldots, (X_n^*, Y_n^*)\}$ be a resample drawn randomly from S with replacement. Denote by $r(j)$ the ranking of the jth predictor X_j such as $\hat{j}_{r(j)} = j$. Let $r^*(j)$ be the ranking of X_j using the bootstrapped resample S^*. Given a value α, such as 0.05, compute a nominal $(1 - \alpha)$-level two-sided prediction interval of the ranking, $[\hat{r}_-(j), \hat{r}_+(j)]$. Hall and Miller (2009) propose a criterion to regard the predictor X_j as influential if $\hat{r}_+(j) < \frac{1}{2}p$. In practice, the cutoff can also be replaced by some smaller fraction of p, such as $\frac{1}{4}p$. Therefore, the proposed generalized correlation ranking reduces the ultrahigh p down to the size of the selected model $\hat{\mathcal{M}}_k = \{j : \hat{r}_+(j) < kp\}$, where $0 < k < 1/2$ is a constant multiplier to control the size of the selected model $\hat{\mathcal{M}}_k$.

77.3.3 Sure Independence Screening for GLIM

The SIS procedure (Fan and Lv 2008) provides one possible method for dealing with ultrahigh-dimensional problems. However, the procedure applies to ordinary linear regression models only, and the theoretical properties of SIS rely heavily on the joint normality assumptions on the response and predictors. These constraints limit significantly the applicability of the SIS for categorical variables, even within the context of linear models.

To this end, Fan and Song (2010) propose a more general version of sure independence screening procedure for generalized linear models. They considered the maximum marginal likelihood estimator (the MMLE, for short) or the marginal likelihood ratio as a marginal utility to rank the importance of each predictor. The conditions under which the proposed MMLE possesses the sure screening property are also explored. Moreover, Fan and Song (2010) discuss how to choose the size of the selected model.

First, consider the generalized linear model (GLIM) with canonical link. That is, the response variable Y conditional on the predictors $\mathrm{x} = (X_1, \ldots, X_p)^{\mathrm{T}}$ is from an exponential family, whose probability density function takes the canonical form

$$f_{Y|\mathrm{x}}(y|\mathrm{x}) = \exp\{y\theta(\mathrm{x}) - b(\theta(\mathrm{x})) + c(y)\}, \tag{77.33}$$

for some known functions $b(\cdot)$, $c(\cdot)$, and $\theta(\mathrm{x}) = \mathrm{x}^{\mathrm{T}}\beta$. Without loss of generality, assume that the dispersion parameter $\phi = 1$ and each predictor are standardized with mean 0 and variance 1.

Therefore, the log-likelihood for the natural parameter θ of the GLIM is

$$\ell(\theta, y) = b(\theta) - y\theta. \tag{77.34}$$

Parallel to Fan and Lv (2008), let $\mathcal{M}_* = \{1 \le j \le p : \beta_j \ne 0\}$ be the true model with the model size $s = |\mathcal{M}_*|$. Fan and Song (2010) defined the maximum marginal likelihood estimator (MMLE) $\hat{\beta}_j^M$ of the jth predictor X_j as

$$\hat{\beta}_j^M = \left(\hat{\beta}_{j,0}^M, \hat{\beta}_j^M\right) = \arg\min_{\beta_0, \beta_1} \sum_{i=1}^{n} \ell\left(\beta_0 + \beta_1 X_{ij}, Y_i\right), \tag{77.35}$$

where Y_i is the ith observed response and X_{ij} is the ith observation of the jth predictor. Although the MMLE $\hat{\beta}_j^M$ is a wrong estimated coefficient for jth predictor X_j in the joint model, the $\hat{\beta}_j^M$ can preserve useful non-sparsity information of X_j in the joint model for variables screening under some mild conditions. Therefore, it is reasonable to consider the magnitude of $\hat{\beta}_j^M$ as a marginal utility to rank the importance of X_j and select a submodel, given a prespecified threshold γ_n:

$$\hat{\mathcal{M}}_{\gamma \mathrm{n}} = \left\{1 \le j \le p : \left|\hat{\beta}_j^M\right| \ge \gamma_n\right\}. \tag{77.36}$$

Theorem 4 of Fan and Song (2010) shows that the MMLEs are uniformly convergent to the population values and establishes the sure screening property of the MMLE screening procedure under some regularity conditions.

Fan and Song (2010) further discuss how large the selected model $\hat{\mathcal{M}}_{\gamma \mathrm{n}}$ should be. Under certain regularity assumptions, they show that with probability approaching one,

$$\left|\hat{\mathcal{M}}_{\gamma \mathrm{n}}\right| = O\left\{n^{2\kappa}\lambda_{\max}(\Sigma)\right\}, \tag{77.37}$$

where k determines how large the thresholding parameter γn is and $\lambda_{\max}(\Sigma)$ is the maximum eigenvalue of the covariance matrix Σ of predictors x, which controls how correlated the predictors are. If $\lambda_{\max}(\Sigma) = O(n^{\tau})$, the size of $\hat{\mathcal{M}}_{\gamma \mathrm{n}}$ has the order $O(n^{2\kappa+\tau})$, which can guide practitioners to choose the thresholding rule.

77.3.4 Model-Free Feature Screening

Zhu et al. (2011) propose a model-free feature screening, called sure independent ranking and screening (SIRS), for ultrahigh-dimensional data. Compared with the SIS and other model-based sure independence screening approaches, the SIRS works for a very general multi-index model framework including many commonly used parametric and semiparametric models. Therefore, the proposed SIRS is more robust to possible model misspecification and can be considered as model-free.

Let Ψ_y be the support of the response Y and denote the conditional distribution function of Y given x as $F(y|\mathrm{x}) = P(Y \leq y|\mathrm{x})$. Define the index sets of active predictors and inactive predictors, respectively, by

$$\mathcal{A} = \left\{k : F(y|\mathrm{x}) \quad \text{functionally depends on } X_k \text{ for some } \quad y \in \Psi_y\right\},$$

$$\mathcal{I} = \left\{k : F(y|\mathrm{x}) \text{ does not functionally depend on } X_k \text{ for any } y \in \Psi_y\right\}.$$

X_k for $k \in \mathcal{A}$ is called an active predictor, whereas X_k for $k \in \mathcal{I}$ is called an inactive predictor.

Considering a general model framework, we assume that $F(y|\mathrm{x})$ depends on x only through $\beta^{\mathrm{T}}\mathrm{x}_{\mathcal{A}}$ for some $p_1 \times K$ constant matrix β. That is,

$$F(y|\mathrm{x}) = F_0\left(y|\beta^{\mathrm{T}}\mathrm{x}_{\mathcal{A}}\right), \tag{77.38}$$

where $F_0(\cdot|\cdot)$ is an unknown function.

Without loss of generality, assume that $E(X_k) = 0$ and $\mathrm{var}(X_k) = 1$ for $k = 1, \ldots, p$. Define

$$\Omega(y) = E\{\mathrm{x}F(y|\mathrm{x})\} = E\{\mathrm{x}E[1(Y \leq y)|\mathrm{x}]\} = \mathrm{cov}\{\mathrm{x}, 1(Y \leq y)\}.$$

Then define a new marginal utility ω_k at the population level by

$$\omega_k = E\left\{\Omega_k^2(Y)\right\}, \quad k = 1, \ldots, p,$$

where $\Omega_k(y)$ is the kth element of $\Omega(y)$. Intuitively, if X_k and Y are independent, then X_k and $1(\mathrm{Y} \leq \mathrm{y})$ for any $y \in \Psi_y$ are independent resulting in that $\omega_k = 0$. On the other hand, if X_k and Y are correlated, then X_k and $1(Y \leq y)$ for some $y \in \Psi_y$ are correlated and thus $\omega_k > 0$.

For a random sample $\{(X_{i1}, \ldots, X_{ip}, Y_i), i = 1, \ldots, n\}$ from $\{\mathrm{x}, Y\}$, the sample moment estimator of ω_k is derived by

$$\hat{\omega}_k = \frac{1}{n}\sum_{j=1}^{n}\hat{\Omega}_k^2(Y_j) = \frac{1}{n}\sum_{j=1}^{n}\left\{\frac{1}{n}\sum_{i=1}^{n}X_{ik}1(Y_i \leq Y_j)\right\}^2, \quad k = 1, \ldots, p,$$

Zhu et al. (2011) suggest to employ the sample estimate $\hat{\omega}_k$ to rank all the candidate predictors and select the top ones as the estimate of the active predictors. Further, Theorem 2 of Zhu et al. (2011) theoretically shows that under some regularity conditions, for $\delta = \min_{k\in\mathcal{A}} \omega_k - \max_{k\in\mathcal{I}} \omega_k$, there exists a sufficiently small constant $s_\delta > 0$ such that

$$\mathbf{P}\left(\max_{k\in\mathcal{I}}\hat{\omega}_k < \min_{k\in\mathcal{A}}\hat{\omega}_k\right) \leq 2p\exp\{n\log(1-\delta s_\delta/4)/3\}. \tag{77.39}$$

This theorem demonstrates the ranking consistency property of the SIRS — namely, the SIRS screening method using $\hat{\omega}_k$ always ranks an active predictor ahead of an inactive one with the probability tending to one. This property provides a clear separation between the active and inactive predictors. Thus, the SIRS is asymptotically consistent in selection for the ultrahigh-dimensional problems.

For the independence screening, Fan and Lv (2008) suggest a hard threshold rule to choose the top variables in the order of $O(n/\log\, n)$, while Zhu et al. (2011) recommend a soft thresholding rule based on adding artificial auxiliary variables to the data. First, randomly generate q auxiliary variables $\{Z_1, \ldots, Z_q\}$ which are independent of both $\mathbf{x}$ and $\mathbf{Y}$. Then, consider the (p + q) dimensional vector $(X1, \ldots, X_p, Z_1, \ldots, Z_q)$ as the predictors and apply the independence screening method to pick top variables. In details, denote ω_k as the marginal utility for kth predictor for $k = 1, \ldots, p + q$. Because $\{Z_1, \ldots, Z_q\}$ are truly inactive, it can be shown that under some mild conditions, $\max_{l=1,\ldots,q}\hat{\omega}_{p+l} < \min_{k\in\mathcal{A}}\hat{\omega}_k$ holds with probability tending to one. Then one can select the predictor subset $\hat{\mathcal{M}}_s = \left\{k : \hat{\omega}_k > \max_{l=1,\ldots,q}\hat{\omega}_{p+l}\right\}$. Zhu et al. (2011) suggest to choose $q = p$ empirically and used numerical studies to show that the soft thresholding rule with this choice can work quite well.

77.3.5 Extensions of Independence Screening

77.3.5.1 Iterative Version of Independence Screening

Fan and Lv (2008) have shown that the SIS can perform very well when the assumed conditions are satisfied. However, when these restrictive conditions fail, the SIS procedure may be problematic. For example, when a variable is jointly correlated, but marginally uncorrelated with the response, the SIS is unlikely to select this important variable, resulting in high false-negative rate. On the other hand, when a variable is jointly uncorrelated but highly marginally correlated with the response, the SIS is likely to select this unimportant variable, resulting in high false-positive rate. To overcome this problem, Fan and Lv (2008) provide an important methodological extension of the SIS, called the iterative sure independence screening (ISIS).

The steps of the ISIS procedure are provided as follows:

Step 1: Apply the SIS to the full data set and select an index set $\hat{\mathcal{A}}_1$ of size $d = [n/\log n]$. Then implement the variable selection approaches, such as penalized least squares with SCAD penalty, on the index set $\hat{\mathcal{A}}_1$ to select a submodel $\hat{\mathcal{M}}_1$. Let $\hat{\mathcal{M}} = \hat{\mathcal{M}}_1$.

Step 2: Compute the residuals from regressing the response Y over $\left\{X_j : j \in \hat{\mathcal{M}}\right\}$. Then treat these residuals as the new responses and apply the same procedure in Step 1 to the remaining variables with indices $\{1,\ldots,p\}\backslash\hat{\mathcal{M}}$ to obtain another submodel1 $\hat{\mathcal{M}}_2$. Let $\hat{\mathcal{M}} = \hat{\mathcal{M}}_1 \cup \hat{\mathcal{M}}_2$.

Step 3: Iterate the process until $|\hat{\mathcal{M}}| \leq d'$, where d' is the prescribed number and $d' < n$. The index set $\hat{\mathcal{M}}$ is the final selected submodel by the ISIS.

Fan and Lv (2008) demonstrate empirically that the ISIS procedure can outperform the ordinary SIS.

Also, Fan et al. (2009) extend this version of ISIS by using the marginal likelihood to rank the importance of variables. Fan et al. (2011a) provide iterative nonparametric independence screening for the sparse ultrahigh-dimensional additive models. Finally, Zhu et al. (2011) develop an iterative version of the model-free independence screening with iteratively transforming the space of predictors.

77.3.5.2 Reduction of False-Positive Rate

The independence screening procedures are commonly used for feature selection, but they are usually conservative and result in many false-positive variables. Fan et al. (2009) propose a simple resampling technique that can help reduce the false-positive rate.

Let $\mathcal{A}$ be the set of active indices. We partition the samples randomly into two parts with the same sample size and then apply one independence screening, such as the SIS and the ISIS, to these two subsamples. Denote $\hat{\mathcal{A}}_1$ and $\hat{\mathcal{A}}_2$ as the selected submodel based on the first half and the second half of the samples, respectively. Under some conditions, both $\hat{\mathcal{A}}_1$ and $\hat{\mathcal{A}}_2$ possess the sure screening property. That is, both $\hat{\mathcal{A}}_1$ and $\hat{\mathcal{A}}_2$ can contain all active indices (i.e., $\mathcal{A}$) with the probability tending to one, i.e.,

$$P\left(\mathcal{A} \subseteq \hat{\mathcal{A}}_1\right) \to 1, \;\; P\left(\mathcal{A} \subseteq \hat{\mathcal{A}}_2\right) \to 1, \text{ as } n \to \infty.$$

Then define $\hat{\mathcal{A}} = \hat{\mathcal{A}}_1 \cap \hat{\mathcal{A}}_2$ as a new estimate of the active set $\mathcal{A}$. Therefore, the estimate $\hat{\mathcal{A}}$ also satisfies the sure screening property:

$$P\left(\mathcal{A} \subseteq \hat{\mathcal{A}}\right) \to 1, \text{as } \; n \to \infty.$$

Intuitively, the probability that one unimportant variable has to be selected twice into both $\hat{\mathcal{A}}_1$ and $\hat{\mathcal{A}}_2$ is very small, so $\hat{\mathcal{A}}$ can be expected to contain much fewer unimportant variables which may be falsely selected into $\hat{\mathcal{A}}_1$ or $\hat{\mathcal{A}}_2$. In the result, this simple resampling approach reduces the false-positive rate efficiently.

77.4 Variable Selection Versus Sufficient Dimension Reduction

Of course, methods for analyzing (ultra)high-dimensional data are by no means limited to those aforementioned. In particular, dimension reduction techniques have constituted an appealing alternative to feature selection in dealing with the curse of dimensionality (Bellman 1961). While both variable selection and dimension reduction are regularized approaches aiming to alleviate the high variance and overfitting problems in (ultra) high-dimensional analysis, it might be useful to draw a broad comparison between them.

There has been a growing interest in dimension reduction methods for regression analysis since the early 1990s, perhaps due to the introduction of sliced inverse regressions (Li 1991) and sliced average variance estimation (Cook and Weisberg 1991). Notably, the proposition of sufficient dimension reduction theory (SDR) by Cook (1998) offers a unified paradigm that can help reduce the dimension of the predictor vector X without loss of information about the response y. The underlying idea is to replace X with a minimal set of their linear combinations $\eta' X$, which concentrates the relevant information in X. More formally, SDR seeks a subspace $\mathcal{S}$ based on η such that $Y \perp (X \backslash P_{\mathcal{S}} X)$, where $\perp$ stands for independence and $P_{\mathcal{S}}$ denotes the orthogonal projection onto $\mathcal{S}$.

As can be seen from their definitions, the fundamental difference between variable selection and SDR lies in their assumptions about the model structure. While variable selection approaches are underpinned by the sparsity principle that assumes that only a small number of original predictors X contribute to the response y, SDR grounds on the existence of latent variables $\eta' X$, the common sources of systematic variation in y. As to the analysis of high-dimensional data in finance, there is no obvious conclusion which framework is preferable. Consider the empirical model where y denotes the aggregate return/volatility of a particular asset class and X consists of hundreds of macroeconomic and financial indicators. On the one hand, the classic asset pricing theory postulates that variations in y depend on a few key risk factors, e.g., aggregate consumption surplus, endowment volatility, and investor sentiment, which typically cannot be directly measured. Thus, the “latent factor” interpretation seems more relevant in this situation. On the other hand, most existing SDR methods suffer because the estimated linear reductions usually involve all of candidate predictors X. As a consequence, the results can be hard to interpret, the return-driving force may be difficult to identify, and the efficiency gain may be less than that possible with variable selection.

To sum, there are several methods available for conducting an ultrahigh-dimensional analysis. Depending on the type of application, some methods rely on dimension reduction, some others rely on variable selection, and a few employ both tactics. How to maximize dimension reduction and simultaneously improve model interpretability poses a significant challenge to financial econometricians. Recent developments in sparse sufficient dimension reduction theory (Li 2007; Zhou and He 2008; Chen et al. 2010) show promise in providing a unified method that can screen out irrelevant and redundant predictors and at the same time lead to a few linear combinations of active predictors.

77.5 Conclusion

One important development in the financial markets over the past decade or so is the explosion of financial data of unprecedented size and complexity, such as intraday transaction data on both exchange-traded securities (e.g., TAQ) and OTC market securities (e.g., TRACE). As such, high- and ultrahigh-dimensional data analysis has attracted a lot of attention from both researchers and practitioners. In this chapter, we have reviewed several existing variable selection methods and independence screening procedures developed recently in the statistics literature that are used to estimate a sparse model and select significant variables simultaneously.

While shrinkage-based variable selection methods have been successfully applied in many high-dimensional analyses, their direct applications to ultrahigh-dimensional statistical learning problems may raise issues associated with computational expediency and algorithm stability. In these situations, feature screening seems to be an essential step, as it is only concerned with the marginal response of the dependent variable to individual candidate predictors and thus enjoys great computational efficiency.

Appendix 1: One-Step Sparse Estimates

This appendix outlines the local linear approximation (LLA) algorithm, discussed in Sect. 77.2.3, for finding a solution of penalized least squares for a broad class of penalty functions. For illustration, we focus on the nondifferentiable, nonconvex $L_{0.5}$ and SCAD penalties introduced in Sect. 77.2.2, as both are ideally suited to the LLA algorithm (Zou and Li 2008). Recall that these two penalty functions are given by Eqs. (77.18) and (77.23), respectively. We note that $L_{0.5}$ and SCAD apparently lead to concave objective functions that are singular at the origin. We illustrate later that in linear regressions this optimization problem can be reduced into solving penalized least squares with an L_1 penalty.

We begin with the following one-step LLA estimator:

$$\hat{\beta} = \arg\min_{\beta} \left\{ \frac{1}{2} \|\mathrm{y} - \mathbf{X}\beta\|^2 + n \sum_{j=1}^{p} p'_{\lambda}\left(\left| \widetilde{\beta}_j \right| \right) |\beta_j| \right\}, \tag{77.40}$$

where the initial estimate $\widetilde{\beta}$ is usually represented by the ordinary least squares estimator in practice if $n > p$. Theoretically, $\widetilde{\beta}$ could be any root-n-consistent estimator for β. In the case where $n < p$, the plain vanilla L_1 estimator (LASSO) would be employed instead.

Following Zou and Li (2008), we discuss separately algorithms for the objective function $Q(\beta)$ with two different types of penalty functions, which correspond to $L_{0.5}$ and SCAD, respectively. Nonetheless, the idea underlying the two algorithms is the same – namely, using a data transform to simplify the penalized least squares to a standard L_1 regularization problem, for which we can take advantage of the related efficient algorithms. In particular, the LARS algorithm (Efron et al. 2004) has been widely applied to obtain the entire solution path of LASSO and forward stage-wise regressions. The relevant R package can be downloaded from http://www.stanford.edu/~hastie/swData.htm#SvmP.

Consider first type one of penalty functions.

Type 1. The tuning parameter could be disentangled from the penalty function, i.e., $p_\lambda(\theta)$ satisfies that $p_\lambda(\theta) = \lambda p_\lambda(\theta)$ and $p'_\lambda(\theta) > 0$. For example, bridge penalties $p_\lambda(|\theta|) = \lambda|\theta|^q$ for $0 < q < 1$ and the logarithm penalty $p_\lambda(|\theta|) = \lambda \log|\theta|$.

Algorithm 1

Step 1. *Create working data by* $x_j^* = x_j / p'_\lambda\left(\left|\widetilde{\beta}_j\right|\right)$, *for* $j = 1, \ldots, p$.

Step 2. *Apply the LARS algorithm to solve the penalized* $L1$ *regression*:

$$\hat{\beta}^* = arg\min_{\beta}\left\{\frac{1}{2}\|y - \mathbf{X}^*\beta\|^2 + n\lambda\sum_{j=1}^{p}\left|\beta_j\right|\right\}.$$

Step 3. *Let the final one-step estimator be* $\hat{\beta}_j = \hat{\beta}_j^* / p'_\lambda\left(\left|\widetilde{\beta}_j\right|\right)$ *for* $j = 1, \ldots, p$.

As we can see, through the one-step sparse estimator, LLA for the $L_{0\cdot 5}$ penalty is equivalent to the adaptive LASSO, with $\hat{w}_j = \left|\widetilde{\beta}_j\right|^{-0.5}$ (see Eq. 4 in Zou 2006). In computing the LARS estimates, tuning parameter λ can be chosen using the *cv.lars* routine embedded in the "lars" package. For example, the following command

$$cv.lars(xstar, y, K = 5, plot.it = TRUE, se = TRUE, type = \text{"}lasso\text{"})$$

instructs the R to plot fivefold cross-validated mean squared prediction error (MSE) for different values of λ. Then we are able to pick the λ with the smallest MSE and find out the step in LARS corresponding to that λ value. The main function of the package

$$lars(xstar, y, type = \text{"}lasso\text{"})$$

provides the entire sequence of coefficients and fits, starting from zero to the least squares fit.

Next, consider type 2 of penalty functions.

Type 2. $p_\lambda(\theta)$ satisfies that the derivative $p'_\lambda(\theta)$ is zero for some values. In addition, the regularization parameter λ cannot be separated from $p_\lambda(\theta)$. For example, the SCAD penalty with the first derivative

$$p'_\lambda(\theta) = \lambda\left\{I(\theta \le \lambda) + \frac{(a\lambda - \theta)_+}{(a-1)\lambda} I(\theta > \lambda)\right\}, \tag{77.41}$$

where $\theta > 0$, $p_\lambda(0) = 0$, and $a > 2$.

We define $U = \{j : p'_\lambda(\theta) = 0\}$ and $V = \{j : p'_\lambda(\theta) > 0\}$. Accordingly, we write $\mathbf{X} = [\mathbf{X}_U, \mathbf{X}_V]$ and $\hat{\beta} = \left(\hat{\beta}_U^T, \hat{\beta}_V^T\right)^T$.

Algorithm 2

Step 1. *Create working data by* $x_j^* = x_j \lambda / p'_\lambda\left(\left|\widetilde{\beta}_j\right|\right)$, *for* $j \in V$. *Let* H_U *be the projection matrix in the space of* $\{x_j^*, j \in U\}$. *Compute* $\widetilde{y} = (I - H_U)y$ *and* $\widetilde{X}_V^* = (I - H)X_V^*$.

Step 2. *Apply the LARS algorithm to solve the penalized* L_1 *regression*:

$$\hat{\beta}_V^* = arg\ \min_\beta \left\{ \frac{1}{2} \left\| \widetilde{\mathrm{y}} - \widetilde{\mathrm{X}}_V^* \beta \right\|^2 + n\lambda \sum_{j=1}^{p} |\beta_j| \right\}.$$

Step 3. *Compute* $\hat{\beta}_U^* = \left(\mathbf{X}_U^{*T}\mathbf{X}_U^*\right)^{-1}\mathbf{X}_U^{*T}\left(\widetilde{\mathrm{y}} - \mathbf{X}_V^*\hat{\beta}_V^*\right)$. *Then, the final one-step estimator* $\hat{\beta}$ *is obtained by*

$$\hat{\beta}_U = \hat{\beta}_U^* \ \ and\ \ \hat{\beta}_j = \hat{\beta}_j^* \lambda / p_\lambda' \left(\left| \widetilde{\beta}_j \right| \right) \ \ for\ \ j \in V.$$

We note that cross-validation routine is not applicable in this situation, as the sets U and V could change as λ varies. In other words, different values of the tuning parameter lead to different transforms of observations. As a result, the SCAD-type penalty requires reexecuting the cross-validation and solving the one-step estimator for each fixed λ. Alternatively, we can determine λ by a non-data-driven approach such as the BIC-based tuning parameter selector. Indeed, Wang et al. (2007) show that the commonly used generalized cross-validation tends to overshoot the correct number of nonzero coefficients. Instead, BIC can be used to consistently identify the true model. For practitioners, it would be more convenient to directly call the procedure *getfinalSCADcoef* included in the "SIS" package (Fan et al. 2010). The option *tune.method* = *c* ("*AIC*", "*BIC*") is used to specify the selection criterion.

Finally, we use a simple numerical example to illustrate the performance of one-step sparse estimates. In this example, simulation data is generated by executing the following command in the R 2.15.0 program:

$$\begin{aligned}
&set.seed(0)\\
&b <- \ c(4,4,4,\ -6 * sqrt(2))\\
&n = 150\\
&p = 300\\
&x = matrix(rnorm(n * p, mean = 0, sd = 1), n, p)\\
&y <- x[, 1:4]\% * \%b + rnorm(150)
\end{aligned}$$

The one-step sparse estimates with the $L_{0 \cdot 5}$ and SCAD penalty are summarized as follows. Both selection methods include all four significant variables, and the $L_{0.5}$ penalty falsely selects one noise variable. The estimated nonzero coefficients under these two methods are reported in the following:

$$\begin{aligned}
b_{scad} &= [3.986, 3.937, 3.944,\ -8.369],\\
b_{L05} &= [3.942, 3.964, 3.997,\ -8.388, 0.237].
\end{aligned}$$

Note that this example alone does not mean that the SCAD outperforms the $L_{0.5}$ penalty, because a valid Monte Carlo simulation study usually requires at least 1,000 simulated data sets.

References

Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In B. N. Petrov & F. Csaki (Eds.), *Second international symposium on information theory* (Vol. 1, pp. 267–281). Budapest: Akademiai Kiado.

Allen, D. (1974). The relationship between variable selection and data augmentation and a method for prediction. *Technometrics, 16*, 125–127.

Anderson, H. M., & Vahid, F. (2007). Forecasting the volatility of Australian stock returns. *Journal of Business and Economic Statistics, 25*, 76–90.

Antoniadis, A., & Fan, J. (2001). Regularization of wavelets approximations (with discussion). *Journal of the American Statistical Association, 96*, 939–967.

Bai, J., & Ng, S. (2008). Forecasting economic time series using targeted predictors. *Journal of Econometrics, 146*, 304–317.

Bellman, R. (1961). Adaptive control processes: A guided tour. *Pattern Recognition Letters, 29*, 1351–1357.

Breiman, L. (1996). Heuristics of instability and stabilization in model selection. *Annals of Statistics, 24*, 2350–2383.

Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (1997). *The econometrics of financial markets*. Princeton: Princeton University Press.

Candes, E., & Tao, T. (2007). The Dantzig selector: Statistical estimation when p is much larger than n (with discussion). *Annals of Statistics, 35*, 2313–2404.

Chen, X., Zou, C., & Cook, R. D. (2010). Coordinate-independent sparse sufficient dimension reduction and variable selection. *Annals of Statistics, 38*, 3696–3723.

Comon, P. (1994). Independent component analysis, a new concept? *Signal Processing, 36*, 287–314.

Cook, R. D. (1998). *Regression graphics: Ideas for studying regressions through graphics*. New York: John Wiley & Sons.

Cook, R. D., & Weisberg, S. (1991). Sliced inverse regression for dimension reduction: Comment. *Journal of the American Statistical Association, 86*, 328–332.

Craven, P., & Wahba, G. (1979). Smoothing noisy data with spline functions: Estimating the correct degree of smoothing by the method of generalized cross-validation. *Numerische Mathematik, 31*, 377–403.

Demartines, P., & Hérault, J. (1997). Curvilinear component analysis: A self-organizing neural network for nonlinear mapping of data sets. *Neural Networks, 8*, 148–154.

Efron, B., Hastie, T., Johnstone, I., & Tibshirani, R. (2004). Least angle regression (with discussion). *Annals of Statistics, 32*, 409–499.

Fan, J., & Li, R. (2001). Variable selection via nonconcave penalized likelihood and it oracle properties. *Journal of the American Statistical Association, 96*, 1348–1360.

Fan, J., & Li, R. (2006). Statistical challenges with high dimensionality: Feature selection in knowledge discovery. *Proceedings of the International Congress of Mathematicians, European Mathematical Society, Zurich, III*, 595–622.

Fan, J., & Lv, J. (2008). Sure independence screening for ultrahigh dimensional feature space (with discussion). *Journal of the Royal Statistical Society, Series B, 70*, 849–911.

Fan, J., & Song, R. (2010). Sure independence screening in generalized linear models with NP-dimensionality. *The Annals of Statistics, 38*, 3567–3604.

Fan, J., Samworth, R., & Wu, Y. (2009). Ultrahigh dimensional feature selection: Beyond the linear model. *Journal of Machine Learning Research, 10*, 1829–1853.

Fan, J., Feng, Y., Samworth, R., & Wu, Y. (2010). *SIS: Sure Independence Screening*. R package version 0.6. http://CRAN.R-project.org/package=SIS

Fan, J., Feng, Y., & Song, R. (2011a). Nonparametric independence screening in sparse ultra-high dimensional additive models. *Journal of the American Statistical Association, 106*, 544–557.

Fan, J., Lv, J., & Qi, L. (2011b). Sparse high-dimensional models in economics. *Annual Review of Economics, 3*, 291–317.
Frank, I. E., & Friedman, J. H. (1993). A statistical view of some chemometrics regression tools. *Technometrics, 35*, 109–148.
Friedman, J. (2008). Fast sparse regression and classification. *International Journal of Forecasting, 28*, 722–738.
Goto, S., & Xu, Y. (2010). *On mean variance portfolio optimization: Improving performance through better use of hedging relations* (Working paper), University of Rhode Island.
Hall, P., & Miller, H. (2009). Using generalized correlation to effect variable selection in very high dimensional problems. *Journal of Computational and Graphical Statistics, 18*, 533–550.
Hastie, T. (1994). *Principal curves and surfaces*. DTIC Document.
Hastie, T., Tibshirani, R., & Friedman, J. (2009). *The elements of statistical learning: Data mining, inference, and prediction* (2nd ed.). New York: Springer.
Hoerl, A., & Kennard, R. (1970). Ridge regression: Biased estimation for nonorthogonal problems. *Technometrics, 12*, 55–67.
Huang, J.-Z., & Shi, Z. (2010). *Determinants of bond risk premia*. (American Finance Association 2011 Denver meetings paper). Penn State University. http://ssrn.com/abstract=1573186.
Ji, P., & Jin, J. (2012). UPS delivers optimal phase diagram in high dimensional variable selection. *Annals of Statistics, 40*, 73–103.
Li, K. C. (1991). Sliced inverse regression for dimension reduction. *Journal of the American Statistical Association, 86*, 285–299.
Li, L. (2007). Sparse sufficient dimension reduction. *Biometrika, 94*, 603–613.
Li, R., Zhong, W., Zhu, L. (2012). Feature screening via distance correlation learning. *Journal of the American Statistical Association, 107*, 1129–1139.
Mallows, C. L. (1973). Some comments on Cp. *Technometrics, 15*, 661–675.
Mardia, K. V., Kent, J. T., & Bibby, J. M. (1980). *Multivariate analysis*. London: Academic.
Miller, A. J. (2002). *Subset selection in regression* (2nd ed.). New York: Chapman & HALL/CRC.
Schwartz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*, 461–464.
Shao, J. (1997). An asymptotic theory for linear model selection. *Statistica Sinica, 7*, 221–264.
Stock, J. H., & Watson, M. W. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association, 97*, 1167–1179.
Tibshirani, R. (1996). Regression shrinkage and selection via LASSO. *Journal of the Royal Statistical Society, Series B, 58*, 267–288.
Wang, H., Li, R., & Tsai, C. L. (2007). Tuning parameter selectors for the smoothly clipped absolute deviation method. *Biometrika, 94*, 553–568.
Zhang, C. (2010). Nearly unbiased variable selection under minimax concave penalty. *Annals of Statistics, 38*, 894–942.
Zhou, J., & He, X. (2008). Dimension reduction based on constrained canonical correlation and variable filtering. *Annals of Statistics, 36*, 1649–1668.
Zhu, L., Li, L., Li, R., & Zhu, L. (2011). Model-free feature screening for ultrahigh dimensional data. *Journal of the American Statistical Association, 106*, 1464–1475.
Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American Statistical Association, 101*, 1418–1429.
Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society, Series B, 67*, 301–320.
Zou, H., & Li, R. (2008). One-step sparse estimates in nonconcave penalized likelihood models. *Annals of Statistics, 36*, 1509–1533.
Zou, H., & Zhang, H. H. (2009). On the adaptive elastic-net with a diverging number of parameters. *Annals of Statistics, 37*, 1733–1751.

Hedonic Regression Models

78

Ben J. Sopranzetti

Contents

Abstract

This provides a basic overview of the nature and variety of hedonic empirical pricing models that are employed in the economics literature. It explores the history of hedonic modeling and summarizes the field's utility-theory-based, microeconomic foundations. It also provides a discussion of and potential solutions for common problems associated with hedonic modeling.

B.J. Sopranzetti
Rutgers, The State University of New Jersey, Newark, NJ, USA
e-mail: sopranze@business.rutgers.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_78,

This paper examines three specific, different hedonic specifications, the linear, semilog, and Box–Cox transformed hedonic models, and applies them to real estate data. It also discusses recent innovations related to hedonic models and how these models are being used in contemporary studies.

Keywords
Hedonic models • Regression • Real estate • Box–Cox • Pricing • Price indexes • Semilog • Least squares • Housing • Property

78.1 Introduction

Hedonic modeling first originated as a method for valuing the demand and the price of farm land.[1] Although Court (1939) is widely considered to be the father of hedonic modeling, his paper had nothing to do with real estate; instead, Court created a hedonic pricing index for automobiles. Regardless of how they are used, hedonic regressions deconstruct the price of an asset into the asset's component parts and then use some form of ordinary least squares regression analysis to examine how each individual piece uniquely contributes to the item's overall value.

The consumer price index is probably the most famous example of the use of hedonic regressions as a control mechanism for differences in the quality of products over time. The consumer price index measures the change over time in the price of a bundle of goods. But, if the quality of the goods in the bundle changes over time, then one can imagine that obtaining a high quality prediction of the value of the index at some future point in time can be problematical. For example, imagine trying to predict the price of an automobile today, based upon pricing information from the 1963. A Chevrolet Corvette costs substantially more money today that it did back in 1963. Some of the increase in the car's price is due to inflation, but another part of the price increase is because Corvettes are far safer, faster, and lighter today than they were back then. Corvettes nowadays enjoy a vast technical superiority over their vintage forbearers. The technical improvements clearly add value. As we will see below, they add "hedonic utility." So when trying to predict the price of a new Corvette with a turbo-charged engine, air conditioning, 6-speed transmission, airbags, and a sport-tuned suspension, one cannot examine the simple inflationary price increase of a Corvette, but in addition one must examine the price increases of the additional individual improvements.

Although the consumer price index might arguably be the most famous use of hedonic modeling techniques, hedonic pricing models are also widely utilized to price other items, such as electronics, clothing, and, in particular, real estate

[1] Hass (1922) and Wallace (1926) both use hedonic-style models to value farmland in the Midwestern United States.

(the focus of this paper), where they are most often utilized to correct for the heterogeneity among properties and houses. Since each house has idiosyncratic characteristics that make it unique, estimating demand or prices from real estate data can be challenging. Rather than pricing a given house or property directly, a researcher can deconstruct the house and property into their value-adding components, such as lot size, square feet, the number of bathrooms, the number of bedrooms, and neighborhood quality. A well-specified hedonic model will estimate the contribution to the total price of each of these features separately. If it is the price that is estimated, then the hedonic model is called an "additive" model. If, instead, the elasticity is estimated, then it is called a "log" model. This short primer will explore in some detail the nature and variety of hedonic pricing models and should provide a solid foundation for any researcher interested in employing this widely utilized empirical technique.

78.2 The Theoretical Foundation

Although many empirical papers using hedonic modeling techniques were published in the years that followed Court's work, Lancaster's (1966) seminal paper is the first attempt to create a theoretical foundation for hedonic modeling. To this end, Lancaster presented a groundbreaking theory of hedonic utility. Lancaster argues that it is not necessarily a good itself that creates utility, but instead the individual "characteristics" of a good that create utility. Specifically, an item's utility is simply the aggregated utility of the individual utility of each of its characteristics. For example, the utility that comes from owning an expensive car comes not so much from the car itself, but from the fact that it provides not only transportation but also fast acceleration, enhanced safety, attractive styling, increased prestige, etc. Furthermore, he argues that items can be arranged into groups based on the characteristics they contain. Consumers make their purchasing decisions within a group based on the number of characteristics a good possesses per unit cost. For example, people make their home purchase decisions based upon the number of bedroom, number of bathrooms, etc.

Although Lancaster is the first to discuss hedonic utility, he says nothing about pricing or pricing models. Rosen (1974) is the first to present a theory of hedonic pricing. Rosen argues that an item can be valued as the sum of its utility generating characteristics; that is, an item's total price should be the sum of the individual prices of its characteristics. This implies that an item's price can be regressed upon the characteristics to determine the way in which each characteristic uniquely contributes to the price. Although Rosen did not formally present a functional form for the hedonic pricing function, his model clearly implies a nonlinear pricing structure.

For a more thorough review of the extant literature, please see the excellent surveys by Follain and Jimenez (1985) and Sheppard (1999).

78.3 The Data

The data are 7,088 observations of real estate transactions of properties that are located in the Klein School District, of suburban Houston, between January 1, 1992 and December 31, 1995. The data set provides information on property, home, and transaction characteristics. Property characteristics include the lot size in square feet, distance to the central business district, and information on neighborhood submarkets within the overall Klein area. Home characteristics include the size of the house in square feet, year built, number of bathrooms, number of bedrooms, a dummy variable for the existence of a pool, and whether the house has any known defects. Transaction characteristics include the list price, the transaction price, the list date, the number of days that the property remained on the market prior to being sold, whether the property is sold "as is," whether the home was sold by a financial institution that had previously foreclosed on the property, and the identity of the listing and selling real estate brokers.

Out of the initial 7,088 observations, 146 observations are deleted due to missing or obviously incorrect data. In addition, several screens are employed to increase the homogeneity of the properties in our sample and to eliminate observations in which data may not have been entered correctly. Properties are omitted if their:

1. Age exceeds 30 years old (24 observations).
2. Lot size is smaller than 5,000 or larger than 50,000 ft^2 (165 observations).
3. Living space exceeds 4,000 ft^2 (362 observations).
4. Number of bathrooms (full and half) exceeds six (six observations).

The final sample includes 6,385 observations. Table 78.1 includes the summary statistics for the dependent variables.

Table 78.1 Summary statistics

Name	Mean	St. dev	Minimum	Maximum
SALE	98,114	44,438	18,000	385,000
FT	2,250	668	819	3,994
AGE	12.60	5.90	0	28
LOT	9,563	3,972	5,000	49,933
BED	3.63	0.60	1	6
BATH	2.22	0.44	1	5.1
POOL	0.14	0.35	0	1
FC	0.05	0.21	0	1
ASIS	0.01	0.12	0	1
D1DUMMY	0.03	0.16	0	1
CBD	20.26	3.50	10.61	28.86
NEW	0.02	0.13	0	1
SUMMER	0.41	0.49	0	1

78.4 The Linear Model

The basic additive hedonic equation is one where the value of an asset is regressed against the characteristics that determine its value. This linear model is appropriate when there are heterogeneous products and heterogeneous buyers, and the heterogeneous items to be valued can be easily replaced/restocked once they have been purchased, i.e., there is a continuous, uninterrupted supply of the item to be priced. In this case, the pricing model of the item is very simply the sum of the prices of its component parts, where

$$\mathrm{Value} = \mathrm{f}(\mathrm{S}, \mathrm{N}, \mathrm{L}, \mathrm{C}, \mathrm{T})$$

where

S represents the structural characteristics of the home and property, e.g., square footage, property size, and number of bedrooms

N represents the neighborhood characteristics

L represents the location within a giving market

C represents the contract conditions, e.g., is the property sold as is and is there a condo fee?

T represents the date or time that the transaction price is observed.

For ease of notation, we allow the matrix X to represent the combination of the individual vectors S, N, L, C, and T

$$VALUE = X\beta + \varepsilon \tag{78.1}$$

Thus an item's expected price is the characteristics X times β, where β represents a vector of marginal prices.

78.5 Empirical Specification

78.5.1 The Dependent Variable

In real estate modeling, it is common to use the most recent transaction price as a dependent variable. There are also studies that use rent as the dependent variable, but rents are problematical since different apartments may have different terms in the rental agreement; for example, some might include heat and hot water or parking. One way of dealing with the "rent" problem is to obtain the cost of utilities (or parking) for properties where the utilities are not included in the rental agreement and add these costs to the base rental price to get an adjusted rental price. Another possibility is to use the actual rental price as a dependent variable and then add a dummy-independent variable that equals one if the unit includes utilities or parking and zero otherwise. Using actual transaction prices circumvents these problems but exposes the researcher to another problem that current transactions may not be representative of the total housing stock: a selection bias.

78.5.2 Independent Variables

One of the principal criticisms of the hedonic modeling of real estate prices is the severity of the omitted variable problem: the coefficient estimates are often not robust to changes in the model's specification. This implies that researchers must be very careful when interpreting the coefficients of a hedonic regression. A sampling of recent hedonic real estate models yields the following common dependent variables:

Structural characteristics of the home and property: square footage of the unit, square footage of the property, total number of rooms, total number of bedrooms, total number of bathrooms, the existence of a pool, any known defects, structural type (single family, duplex, condominium, etc.), age, air conditioning, finished basement, fireplaces, garages, etc.

Neighborhood characteristics: quality of the school system, quality of the neighborhood, median salary.

Location within a giving market: distance from the central business district, proximity to a train station, distance to supermarket, distance to schools, flooding area.

Contract conditions: was the property sold as is? Was it a foreclosure?

For a more complete discussion of the potential explanatory variables, please see Hocking (1976), Leamer (1978), and Amemiya (1980).

78.5.3 Example Using the Linear Model

Table 78.2 presents the results of a linear hedonic pricing model where the dependent variable is the property's transaction price. The independent variables include the square footage of the home and its square, the age of the home and its square, the size of the lot and its square, the number of bedrooms and its square, dummy variables representing the number of bathrooms, the proximity in miles to the central business district, a dummy variable for a pool, foreclosure, sold as is, if there are defects, if the home is brand new, and if the property was listed in the summer.

The results of the hedonic regression will likely not be surprising to anyone that has ever shopped for a home. The hedonic model has parceled out the value of the home/property into its component parts and has succeeded in explaining 84.3 % of the value of transaction price. The relationship between real estate transaction prices and the square footage of the home is concave; for small homes small increases in size have a larger marginal impact than for larger homes. The same is true for the square footage of the property and the bedrooms. The relationship between transaction price and age is convex; implying that the market price of young houses depreciates at a larger rate than that of older houses. It is not surprising that given the hot summers in Texas, a pool adds substantial value (in this case $11,447 of value), nor is it surprising that foreclosed upon properties and properties that are sold as is tend to sell for less money than other properties.

Table 78.2 OLS regression of transaction price on the independent variables

OLS sale		
	Coefficient	T-stat
FT	0.94591	10.33
SQFT	−8.65E-04	−15.35
AGE	−3,035.3	−20.71
SQAGE	59.356	10.25
LOT	1.7881	10.01
SQLOT	−1.62E-05	−3.59
BED	42,839	11.40
SQBED	−6,207.4	−12.35
BATH 2	6,305.7	2.11
BATH 3	4,199.3	5.82
BATH 4	15,936	17.26
BATH 5	16,196	6.26
POOL	11,447	16.85
FORECLOSE	−10,025	−9.03
ASIS	−11,932	−6.21
DEFECT	−383.98	−0.28
CBD	267.98	4.12
NEW	3,004.3	1.60
SUMMER	1,260.8	2.81
CONSTANT	−17,355	−2.78
R^2	84.3 %	

Properties farther from the central business district also tend to sell for higher prices. Interestingly, properties listed in the summer tend to sell for on average \$1,260 higher prices than those not listed in the summer.

The principal use of a linear hedonic model is to help researchers construct a property's transaction price from its component parts. So, using the above model, a 3-year-old, 1,800 ft^2 home, with a 200 ft by 200 ft lot, three bedrooms and two baths, and a pool, sold at a foreclosure sale, with no known defects, which is located 7 miles from the central business district and listed in the summertime, would have a predicted transaction price of \$102,092.

78.6 The Semilog Model

When the item to be priced cannot be easily restocked, for example, a home, then nonlinearities arise in the hedonic pricing structure. To deal with this, it is common for researchers to employ the following semilogarithmic functional form for the hedonic model:

$$VALUE = e^{x\beta\varepsilon} \tag{78.2}$$

thus

$$\ln VALUE = X\beta + \varepsilon \tag{78.3}$$

In this case, the log of an item's expected price is the sum of its characteristics $\boldsymbol{X}$ times $\boldsymbol{\beta}$, and the marginal price of each individual attribute $\boldsymbol{x}$ is

$$PRICE(x) = e^{xb} \tag{78.4}$$

where $\boldsymbol{x}$ is the current level of the characteristic and $\boldsymbol{b}$ is the regression coefficient. Notice that the semilog form implies that the price of a given characteristic varies with its level, i.e., the prices are nonlinear.

The semilog structural hedonic pricing model has several advantages over its linear counterpart. The principal advantage is that it permits the value of a given characteristic (e.g., the number of bathrooms) to vary proportionately with the value of other characteristics (the number of bedrooms). This is not the case with a linear model, where a second bathroom adds the same value to a house that has one bedroom as it does to one that has five bedrooms.

78.6.1 Example Using the Semilog Model

Table 78.3 presents the results of a linear hedonic pricing model where the dependent variable is the log of the property's transaction price. Again, the results are not surprising. An advantage of the semilog form is that the model's coefficients are easily interpreted. The percentage change in the value of the house for a unit change in the dependent variable can be represented as $e^b - 1$, where $\boldsymbol{b}$ is the regression coefficient.[2] For example, if the coefficient on the variable that represents a pool equals 0.104, then adding a pool to a house would increase its value by $e^{0.104} - 1 = 10.96\%$.

78.7 The Box–Cox Model

Although not as widely employed due to the difficulty of interpreting the coefficients, a more general form of the hedonic pricing model was first presented by Halvorsen and Pollakowski (1981), which applies the seminal work of Box and Cox (1964) to hedonic modeling. The basic Box–Cox transformation is

[2]See Halvorsen and Palmquist (1980).

Table 78.3 OLS regression of LOGSALE on the dependent variables

OLS LOGSALE		
Coefficient	T-stat	
FT	6.76E-04	28.68
SQFT	−5.57E-08	−11.89
AGE	−2.50E-02	−20.48
SQAGE	4.33E-04	8.98
LOT	2.02E-05	13.62
SQLOT	−2.70E-10	−7.17
BED	0.3214	10.28
SQBED	−4.46E-02	−10.66
BATH 2	1.29E-02	0.52
BATH 3	5.32E-02	8.86
BATH 4	0.11094	14.44
BATH 5	9.35E-02	4.35
POOL	0.10416	18.43
FORECLOSE	−0.13635	−14.77
ASIS	−0.18429	−11.52
DEFECT	−9.26E-03	−0.81
CBD	4.17E-03	7.71
NEW	−2.02E-02	−1.29
SUMMER	1.38E-02	3.69
CONSTANT	9.5387	183.40
R^2	87.5 %	

$$X^{(\lambda)} = \frac{X^{\lambda} - 1}{\lambda}, \quad \text{if } \lambda \neq 0, X > 0 \tag{78.5}$$

$$= \ln X, \quad \text{if } \lambda = 0$$

So the basic form of the Box–Cox model is given by

$$VALUE = X^{(\lambda)}\beta + \varepsilon \tag{78.6}$$

Please see the Appendix for a discussion of how the Box–Cox coefficients are estimated. Notice that when λ equals one, the Box–Cox structural form reduces to the linear form. A more general form of the Box–Cox model can be expressed by

$$VALUE^{(\theta)} = \beta_0 + \sum_j \beta_j X_j^{\lambda} + \frac{1}{2}\sum_j \sum_k \gamma_{jk} X_j^{\lambda} X_k^{\lambda} \tag{78.7}$$

In this case, when θ and λ are both equal to one and the cross products γ_{jk} are all zero, then Eq. 78.7 reduces down to a simple linear regression model. When θ and λ are both equal to zero and the cross products γ_{jk} are all also equal to zero, then the model reduces down to a straightforward log-log functional form.

Table 78.4 Box–Cox regression of the transaction price on the dependent variables

Box sale		
Coefficient	T-stat	
FT	2.35E-04	31.29
SQFT	−2.19E-08	−14.67
AGE	−7.82E-03	−20.13
SQAGE	1.34E-04	8.71
LOT	6.49E-06	13.69
SQLOT	−8.79E-11	−7.33
BED	0.10082	10.12
SQBED	−1.39E-02	−10.43
BATH 2	8.94E-03	1.13
BATH 3	1.70E-02	8.87
BATH 4	3.37E-02	13.76
BATH 5	2.86E-02	4.17
POOL	3.28E-02	18.21
FORECLOSE	−4.49E-02	−15.26
ASIS	−6.17E-02	−12.10
DEFECT	−3.08E-03	−0.85
CBD	1.39E-03	8.05
NEW	−7.83E-03	−1.57
SUMMER	4.40E-03	3.69
CONSTANT	6.1796	372.70
	R^2	87.5 %
	λ	−0.100

The Box–Cox model was first introduced into the mainstream finance literature by C. F. Lee in his seminal 1976 paper, which examines the Functional Form and the Dividend Effect of the Electric Utility Industry. Other excellent examples of the application of the Box–Cox model to finance include Lee and Kau (1976), and (Lee et al. 1980).

One of the most innovative uses of Box–Cox hedonic models in real estate has been to back out property depreciation rates for federal tax reasons. Hulton and Wykcoff (1981) present evidence of a constant geometric rate derived from a hedonic model with a Box–Cox transformation of the value industrial and commercial buildings. The more flexible Box–Cox approach permits relationships to emerge instead of forcing a predetermined, and perhaps ad hoc, functional form.

78.7.1 Example Using the Box–Cox Model

Table 78.4 presents the results of a linear hedonic pricing model where the dependent variable is the Box–Cox transformation of the property's transaction price. Notice that lambda emerges from the model and is not exogenously imposed.

78.8 Problems with Hedonic Modeling

78.8.1 The Identification Problem

There is an inherent identification problem that occurs when one attempts to model an item's demand function when the prices that one has to work with come from the interaction of the item's supply and demand functions: it is difficult to separate out the supply and demand impact on price. There is a second problem with hedonic modeling that comes from the nonlinear nature of the pricing structure. In a simple demand model, the price of an item is taken as given, and consumers make their purchase decision (quantity) based upon the exogenous price, i.e., consumers are price takers. But, as seen above, nonlinear hedonic models imply that the price of a characteristic is correlated with quantity; so consequently, buyers will select not only the quantity of a characteristic but, by design, also its price. Several authors, Blomquist and Worley (1982) or Diamond and Smith (1985), have attempted to solve this problem through the use of instrumental variables.

78.8.2 The Equilibrium Pricing Problem

A key aspect of demand modeling is that observed prices are assumed to be equilibrium prices. Unfortunately, in markets such as real estate where adjustment costs can be large, the notion of observed prices being equilibrium ones becomes problematical. There are several papers that attempt to deal with the disequilibrium character of real estate transaction prices. See Maclennan (1977, 1982) for a good overview of the disequilibrium problem.

There are several ways in which researches have attempted to deal with the disequilibrium problem. Bowden (1978) addresses the problem by utilizing only those observations that are either at or near equilibrium. He employs a switching regression technique. See Anas and Eum (1984) for an application of this "disequilibrium" hedonic modeling technique. Although switching regression models can be effective at mitigating the disequilibrium problem, they are not without their challenges. The major problem is how to differentiate between equilibrium and disequilibrium prices. Another problem is that modelers are often more interested in predicting actual future transaction prices rather than so-called equilibrium prices.

Another way to deal with the disequilibrium problem is to find a way to adjust prices back to their equilibrium levels. There are several papers that use this technique.[3] The technique involves estimating a time series index of prices and then finding a way to determine which prices are equilibrium ones (e.g., prices may be deemed to be near their equilibrium values if there was little or no price change from period to period). Take this set of equilibrium prices and estimate their determining

[3]See Abraham and Hendershott (1996), Malpezzi (1998), and Drieman and Follain (2000).

characteristics, so that for each period there is an actual price, and estimated index price, and an estimated equilibrium price. From these prices it is possible to determine the extent to which the price is out of equilibrium. The last step is to find the determinants of the disequilibrium and adjust the initial prices accordingly.

78.9 Recent Developments

There have been some interesting recent developments in the area of hedonic modeling. Below is a brief synopsis of a few recent papers and issues that are on the cutting edge of the literature. Costanigro et al. (2007) argue that when researchers disregard heterogeneity across assets, they introduce an aggregation bias into their estimated prices. They suggest that a model that estimates hedonic functions that are specific to price ranges yields more accurate predictions. Collins et al. (2007) examine a similar uniqueness problem but with respect to art rather than wine.

Goetzmann and Peng (2006) analyze and present a model that adjusts for the bias that occurs because of a house seller's reservation price in transaction-based hedonic price indexes. They present a hedonic model where the ratio of sellers' reservation prices to the actual market value has an impact on trading volume and can lead to a bias in the observed transaction prices. They find that there is an upward bias to index returns when trading volume decreases.

Diewert (2002) and Feenstra and Knittel (2004) examine the problem of quality adjustments in a hedonic model. These papers examine whether the output price index for a durable good can also be used as (partial) input price index. Although these papers focus on the producer price index, the quality adjustments can be applied more widely. The authors find that the relevant input price is the value of the characteristic bundle of the underlying asset (i.e., the market price) divided by the quantity of the individual characteristics associated with the bundle. For example, if one were to buy a cleaning service, one would take the price of the cleaning service divided by the total number of benefits provided by the service (e.g., the time savings, the convenience, the trustworthiness, the quality of the work).

Bajari and Benkard (2005) reexamine hedonic models of demand for differentiated products. They nicely generalize Rosen's original hedonic model to allow for unobservable product characteristics and for the hedonic pricing function to have a nonseparable form. They use a semi-parametric approach to demonstrate that if there are only a few products, one can construct bounds on an individual's utility parameters, and in addition other important considerations, for example, aggregate demand and consumer surplus.[4] Heckman et al. (2010) examine nonadditive

[4]Please see Bontemps et al. (2008), Jensen and Webster (2007), Ferreira and McMillan (2007), Hahn and Mathews (2007), and Epple et al. (2006) for additional work on this interesting theme.

hedonic models. In specific, they examine the issue of nonparametric identification and estimation of these models.

Recently, hedonic models that have been employed examine issues of substantial political importance. Sander and Polasky (2009), Poudyal et al. (2009), Hoshino and Kuriyama (2010), Cutter et al. (2011), (Brander and Koeste 2011), and Nordman and Wagner (2012) all provide evidence on the high value of wide-open spaces in the United States. Jiao and Liu (2010) examines this same issue in Wuhan, China. Donovan and Butry (2010) examine a related issue: the value of trees that line city streets. Gopalakrishnan et al. (2011) examine the value of disappearing beaches due to beach erosion. Kim et al. (2010) and Bayer et al. (2009) utilize hedonic models to measure the value of air quality. Lastly, Bishop and Murphy (2011) have an innovative paper that utilized a dynamic hedonic model to estimate the willingness to pay in order to avoid violent crime.

Hedonic regressions are being increasingly used to better understand the drivers of prices for consumer products. Costanigro et al. (2007, 2009), and Panzone (2011) have utilized hedonic models to examine fluctuations in wine prices. Thrane (2009) uses a hedonic model to determine whether sensory or objective attributes drive wine prices. Benfratello et al. (2009) examine the issue of taste versus reputation. Kassie et al. (2011) utilize a hedonic model to examine the prices of cattle.

78.10 Summary

This short primer provides an overview of the literature and the microeconomic theory that underpins modern hedonic pricing models. At their most basic, hedonic pricing models deconstruct an asset's price into the price of the asset's individual component parts and then use some form of ordinary least squares regression analysis, using either a linear, semilog, or Box–Cox structural form, to examine how each individual component part uniquely contributes to the item's overall value. This paper explores each of these aforementioned structural forms and their associated problems and in addition offers some guidance on common treatments of the dependent and independent variables.

Appendix

This appendix describes the maximum likelihood technique used for the estimation of the nonlinear parameters of the Box–Cox transformation. Given the functional form specified in Eqs. 78.5–78.7 and employing the assumption that there is some λ for which the error term in Eq. 78.6 has an approximate normal distribution with mean of zero and a variance of σ^2, then for the nth observation, the density function can be represented as

$$f(z|\beta,\Sigma) = \frac{e^{\left[-\frac{1}{2}(z-\mu)'\sum^{-1}(z-\mu)\right]}}{(2\pi)^{\frac{n}{2}}|\Sigma|^{\frac{1}{2}}}, \text{where } z = y^{(\lambda)} \tag{78.8}$$

If $E(y)$ is linear, that is, $\mu = X\beta$ and $\sum = \sigma^2 I$, then the density function transforms into something more tractable:

$$f(z|\beta,\Sigma) = \frac{e^{\left[-\frac{1}{2\sigma^2}(z-X\beta)'(z-X\beta)\right]}}{(2\pi\sigma^2)^{\frac{n}{2}}} \tag{78.9}$$

Next, in order to get the density function for y rather than z, we must multiply the density function for z by its Jacobian. If we do so, then we obtain

$$f(y|\beta,\Sigma,\lambda) = \frac{e^{\left[-\frac{1}{2\sigma^2}(z-X\beta)'(z-X\beta)\right]}}{(2\pi\sigma^2)^{\frac{n}{2}}}\prod_{i=1}^{n} y_i^{\lambda-1} \tag{78.10}$$

the corresponding log-likelihood function of which is delineated as

$$\begin{aligned}\ln L\left(\beta,\sigma^2,\lambda|y\right) = -\frac{1}{2\sigma^2}(z-X\beta)'(z-X\beta) - \frac{n}{2}\ln\left(2\pi\sigma^2\right)\\ + (\lambda-1)\sum_{i-1}^{n}\ln y_i\end{aligned} \tag{78.11}$$

Now that we have the log of the likelihood function, the values of the parameters β, λ, and σ are determined by maximum likelihood estimation.

References

Abraham, J. M., & Hendershott, P. H. (1996). Bubbles. *Journal of Housing Research, 7*(2), 191–208.

Amemiya, T. (1980). Selection of regressors. *International Economic Review, 21*, 331–354.

Anas, A., & Eum, S. J. (1984). Hedonic analysis of a housing market in disequilibrium. *Journal of Urban Economics, 15*, 107–123.

Bajari, P., & Lanier Benkard, C. (2005). Demand estimation with heterogeneous consumers and unobserved product characteristics: A hedonic approach. *Journal of Political Economy, 113*(6), 1239–1276.

Bayer, P., Keohane, N., & Timmins, C. (2009). Migration and hedonic valuation: The case of air quality. *Journal of Environmental Economics and Management, 58*(1), 1–14.

Benfratello, L., Piacenza, M., & Sacchetto, S. (2009). Taste or reputation: What drives market prices in the wine industry? Estimation of a hedonic model for Italian premium wines. *Applied Economics, 41*(17), 2197–2209.

Bishop, K., & Murphy, A. (2011). Estimating the willingness to pay to avoid violent crime: A dynamic approach. *The American Economic Review, 101*(3), 625–629.

Blomquist, G., & Worley, L. (1982). Specifying the demand for housing characteristics: The exogeneity issue. In D. B. Diamond & G. Tolley (Eds.), *The economics of urban amenities*. New York: Academic Press.

Bontemps, C., Simioni, M., & Surry, Y. (2008). Semiparametric hedonic price models: Assessing the effects of agricultural nonpoint source pollution. *Journal of Applied Econometrics, 23*(6), 825–842.

Bowden, R. J. (1978). *The econometrics of disequilibrium*. Amsterdam: North Holland.

Box, G. E. P., & Cox, D. (1964). An analysis of transformations. *Journal of the American Statistical Association Society Series B, 26*, 211–252.

Brander, L., & Koeste, M. (2011). The value of urban open space: Meta-analyses of contingent valuation and hedonic pricing results. *Journal of Environmental Management, 92*(10), 2763–2773.

Collins, A., Scorcu, A. E., & Zanola, R. (2007). *Sample selection bias and time instability of hedonic art price indexes* Working Papers 610, Dipartimento Scienze Economiche, Universita' di Bologna.

Costanigro, M., McCluskey, J. J., & Mittelhammer, R. C. (2007). Segmenting the wine market based on price: Hedonic regression when different prices mean different products. *Journal of Agricultural Economics, 58*(3), 454–466.

Costanigro, M., Mittelhammer, R., & McCluskey, J. (2009). Estimating class-specific parametric models under class uncertainty: Local polynomial regression clustering in an hedonic analysis of wine markets. *Journal of Applied Econometrics, 24*(7), 1117–1135.

Court, A. T. (1939). Hedonic Price Indexes with automotive examples. In C. F. Roos (ed.), *The dynamics of automobile demand* (pp. 99–177). New York: The General Motors Corporation.

Cutter, B., Fernandez, L., Sharma, R., & Scott, T. (2011). *Dynamic analysis of open space value using a repeat sales/hedonic approach* (Unpublished Working Paper). University of Oregon.

Diamond, D. B., Jr., & Smith, B. (1985). Simultaneity in the market for housing characteristics. *Journal of Urban Economics, 17*, 280–292.

Diewert, W. E. (2002). Hedonic producer price indexes and quality adjustment (Discussion Paper No. 02-14). Vancouver: University of British Columbia

Donovan, G., & Butry, D. (2010). Trees in the city: Valuing street trees in Portland, Oregon. *Landscape and Urban Planning, 94*(2), 77–83.

Dreiman, M., & Follain. J. R. (2000). Drawing inferences about housing supply elasticity from house price responses to income shocks, Freddie Mac, Processed.

Epple, D., Romano, R., & Sieg, H. (2006). Admission, tuition, and financial aid policies in the market for higher education. *Econometrica, 74*(4), 885–928.

Feenstra, R. C., & Knittel, C. R. (2004). Re-assessing the U.S. quality adjustment to computer prices: The role of durability and changing software, NBER chapters. In *Price index concepts and measurement* (pp. 129–160). Cambridge, MA: National Bureau of Economic Research, Inc.

Follain, J. R., & Jimenez, E. (1985). Estimating the demand for housing characteristics: A survey and critique. *Regional Science and Urban Economics, 15*(1), 77–107.

Goetzmann, W., & Peng, L. (2006). Estimating house price indexes in the presence of seller reservation prices. *Review of Economics and Statistics, 88*(1), 100–112.

Gopalakrishnan, S., Smith, M., Slott, J., & Murray, A. (2011). The value of disappearing beaches: A hedonic pricing model with endogenous beach width. *Journal of Environmental Economics and Management, 61*(3), 297–310.

Haas, G. C. (1922). *Sales prices as a basis for farm land appraisal*. (Technical Bulletin 9) St. Paul: The University of Minnesota Agricultural Experiment Station.

Hahn, W. F., & Mathews, K. H. (2007). Characteristics and hedonic pricing of differentiated beef demands. *Agricultural Economics, 36*(3), 377–393.

Halverson, R., & Pollakowski, H. O. (1981). Choice of functional form equations. *Journal of Urban Economics, 10*(1), 37–49.

Halvorsen, R., & Palmquist, R. (1980). The interpretation of dummy variables in Semilogarithmic regressions. *American Economic Review, 70*, 474–475.

Heckman, J., Matzkin, R., & Nesheim, L. (2010). Nonparametric identification and estimation of nonadditive hedonic models. *Econometrica, 78*(5), 1569–1591.
Hocking, R. R. (1976). The analysis and selection of variables in linear regression. *Biometrics, 32*, 1–49.
Hoshino, T., & Kuriyama, K. (2010). Measuring the benefits of neighbourhood park amenities: Application and comparison of spatial hedonic approaches. *Environmental and Resource Economics, 45*(3), 429–444.
Hulton, C., & Wycoff, F. (1981). The measurement of economic depreciation. In C. Hulten (Ed.), *Depreciation, inflation, and the taxation of income from capital*. Washington, DC: Urban Institute Book.
Jensen, P. H., & Webster, E. (2007). Labelling characteristics and demand for retail grocery products in Australia. *Australian Economic Papers, 47*(2), 129–140.
Jiao, L., & Liu, Y. (2010). Geographic field model based hedonic valuation of urban open spaces in Wuhan, China. *Landscape and Urban Planning, 98*(1), 47–55.
Kassie, G., Abdulai, A., & Wollny, C. (2011). Heteroscedastic hedonic price model for cattle in the rural markets of central Ethiopia. *Applied Economics, 43*(24), 3459–3464.
Kim, S. G., Cho, S. H., Lambert, D., & Roberts, R. (2010). Measuring the value of air quality: Application of the spatial hedonic model. *Air Quality, Atmosphere & Health, 3*(1), 41–51.
Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy, 74*, 57–132.
Leamer, E. E. (1978). *Specification searches: Ad Hoc inference with nonexperimental data*. New York: Wiley.
Lee, C. F. (1976). Functional form and the dividend effect of the electric utility industry. *Journal of Finance, 31*(5), 1481–1486.
Lee, C. F., & Kau, J. B. (1976). The functional form in estimating the density gradient: An empirical investigation. *Journal of American Statistical Association, 71*, 326–327.
Lee, C. F., Fabozzi, F. J., & Francis, J. C. (1980). Generalized functional form for mutual fund returns. *Journal of Financial and Quantitative Analysis, 15*, 1107–1120.
Maclennan, D. (1977). Some thoughts on the nature and purpose of hedonic price functions. *Urban Studies, 14*, 59–71.
Maclennan, D. (1982). *Housing economics*. London: Longman.
Malpezzi, S. (1998). *A simple error correction model of house prices* (Wisconsin-Madison CULER Working Papers. 98–11). Madison: University of Wisconsin, Center for Urban Land Economic Research.
Nordman, E., & Wagner, J. (2012). Public purchases and private preferences: Challenges for analyzing public open space acquisitions. *Urban Forestry & Urban Greening, 11*(2), 179–186.
Panzone, L. (2011). The lost scent of Eastern European wines in Western Europe: A hedonic model applied to the UK market. *British Food Journal, 113*(8), 1060–1078.
Poudyal, N., Hodges, D., & Merrett, C. (2009). A hedonic analysis of the demand for and benefits of urban recreation parks. *Land Use Policy, 26*(4), 975–983.
Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy, 82*(1), 34–55.
Sander, H., & Polasky, S. (2009). The value of views and open space: Estimates from a hedonic pricing model for Ramsey County, Minnesota, USA. *Land Use Policy, 26*(3), 837–845.
Sheppard, S. (1999). Hedonic analysis of housing markets. In P. C. Cheshire & E. S. Mills (Ed.), *Handbook of regional and urban economics*, (vol 3). Amsterdam: Elsevier.
Thrane, C. (2009). Explaining variation in wine prices: The battle between objective and sensory attributes revisited. *Applied Economics Letters, 16*(13), 1383–1386.
Wallace, H. A. (1926). Comparative farmland values in Iowa. *Journal of Land and Public Utility Economics, 2*, 385–392.

Optimal Payout Ratio Under Uncertainty and the Flexibility Hypothesis: Theory and Empirical Evidence

79

Cheng-Few Lee, Manak C. Gupta, Hong-Yi Chen, and Alice C. Lee

Contents

Disclaimer: Views and opinions presented in this publication are solely those of the authors and do not necessarily represent those of State Street Corporation, which is not associated in any way with this publication and accepts no liability for its contents.

C.-F. Lee (✉)
Department of Finance and Economics, Rutgers Business School, Rutgers, The State University of New Jersey, Piscataway, NJ, USA

Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
e-mail: lee@business.rutgers.edu; cflee@business.rutgers.edu

M.C. Gupta
Temple University, Philadelphia, PA, USA
e-mail: mcgupta@temple.edu

H.-Y. Chen
Department of Finance, National Central University, Taoyuan, Taiwan
e-mail: fnhchen@ncu.edu.tw

A.C. Lee
State Street Corp., USA
e-mail: alice.finance@gmail.com

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_79,

Abstract

We theoretically extend the proposition of DeAngelo and DeAngelo's (*Journal of Financial Economics* 79, 293–315, 2006) optimal payout policy in terms of the flexibility dividend hypothesis. We also introduce growth rate, systematic risk, and total risk variables into the theoretical model. Our empirical findings show that based on flexibility considerations, a company will reduce its payout when the growth rate increases. In addition, a nonlinear relationship exists between the payout ratio and the risk. In other words, the relationship between the payout ratio and risk is negative (or positive) when the growth rate is higher (or lower) than the rate of return on total assets.

We use a panel data collected in the USA from 1969 to 2009 to empirically investigate the impact of growth rate, systematic risk, and total risk on the optimal payout ratio in terms of the fixed-effects model. Furthermore, we implement the moving estimates process to find the empirical breakpoint of the structural change for the relationship between the payout ratio and risks and confirm that the empirical breakpoint is not different from our theoretical breakpoint. Our theoretical model and empirical results can therefore be used to identify whether flexibility or the free cash flow hypothesis should be used to determine the dividend policy.

Keywords

Dividends • Payout policy • Optimal payout ratio • Flexibility hypothesis • Free cash flow hypothesis • Signaling hypothesis • Fixed effect • Clustering effect • Structural change model • Moving estimates processes • Systematic risk • Total risk • Market perfection

79.1 Introduction

Corporate dividend policy has long engaged the attention of financial economists, dating back to the irrelevance theorem of Miller and Modigliani (1961; M&M hereafter), in which they state that a rational and perfect economic environment is free of illusion. Since then, their rather controversial findings have been challenged and tested by weakening the assumptions or introducing imperfections into the analysis. For example, the signaling models developed by Bhattacharya (1979), Miller and Rock (1985), and John and Williams (1985) and the free cash flow hypothesis proposed by Easterbrook (1984) and Jensen (1986) are the two well-known propositions challenging M&M's dividend irrelevance theorem; however, empirical studies examining signaling and free cash flow hypotheses have yielded mixed results.

DeAngelo and DeAngelo (2006) reexamine the irrelevance of the M&M dividend irrelevance theorem by allowing partial dividend payout. They argue that the original M&M (1961) irrelevance result is due to the fact that they consider either paying out all earnings or not paying any of them at all. Therefore, payout policy might be relevant if partial payout is allowed. In other words, DeAngelo and DeAngelo (2006) use dividend flexibility hypothesis to show their dividend relevance results. Underlying the flexibility hypothesis, the current paper (1) develops a theoretical model to support the proposition of DeAngelo and DeAngelo's (2006) optimal payout policy when the partial payout is allowed, (2) reconciles the dispute between free cash flow hypothesis and flexibility hypothesis in dividend policy literature, and (3) performs empirical tests in terms of theoretical results derived in this paper.

First of all, following DeAngelo and DeAngelo (2006), we develop a dynamic model allowing firms to hold some amount of cash into a positive NPV project for the reason of financial flexibility. Under the assumption of stochastic rate of return and the dividend flexibility hypothesis, we carry out the optimization procedure to maximize firm value, and the final expression of the optimal dividend policy of the firm is thus derived. The model is comprehensive and allows structural analysis of different variables that could be relevant for corporate dividend policy. For example, the model incorporates the rate of return on assets, growth rate, and risk (systematic, firm-specific, and total risk) to name a few variables that could affect corporate dividend decisions. Comparative statics provide insights into the effect of each of these parameters on corporate dividend policy, and this is followed by an analysis of the interaction effects of these variables on corporate dividend policy.

Second, the implications of the optimization results are explained. Our results show that the relationship between the optimal payout ratio and the growth rate is negative in general. We investigate the separate and then the combined effects of market-dependent and market-independent components of risk on the optimal dividend policy. We perform comparative static analyses of the relationships between the payout ratio and (1) change in total risk, (2) change in systematic risk, (3) simultaneous changes in both total risk and systematic risk, and (4) no change in risk. We examine in detail the effects of variations in the profitability rate, its distribution parameters, and their dynamic behavior on the optimal dividend policy of the firm. The theoretical relationship between the payout ratio and the growth ratio implies that high growth firms need to reduce the payout ratio and retain more earnings to build up "precautionary reserves" for flexibility considerations, but low growth firms are likely to be more mature and already build up their reserves. More importantly, the relationship between the payout ratio and the risks reflects different dividend policies at high growth firms and low growth firms. With higher risk, the costs of external funds increase; therefore, to optimize shareholders' wealth, high growth firms tend to reduce their payouts and keep more relatively low-cost funds to sustain their high growth, whereas low growth firms tend to pay more dividends and reduce the risk for shareholders.

Third, our theoretical model and its implications lead us to three testable hypotheses. Although a large and growing body of empirical research on the optimal dividend payout policy has emerged, none of it has a solid theoretical model to support findings or introduces a nonlinear structured model on payout policy.[1] We here try to empirically examine three hypotheses derived from our theoretical optimal payout model. Using data collected in the USA from 1969 to 2009, we analyze 28,333 dividend-paying firm-years. Our empirical results show that firms' payout ratios are negatively related to firms' growth, and a negative (or positive) relationship exists between firms' risks and the dividend payout ratios among firms with higher (or lower) growth rates relative to their rate of return on assets. Our empirical results are consistent with our theoretical model under the dividend flexibility hypothesis. We also find that growth and risk interact in explaining the payout ratio, indicating that the payout ratio is not linearly related to the growth rate or to the risk of the firm. Furthermore, we implement the moving estimates process to find the empirical breakpoint of the structural change for the relationship between the payout ratio and risks and confirm that the empirical breakpoint is not different from our theoretical breakpoint.

The primary contributions of this paper are our theoretical derivation of an optimal payout ratio under the dividend flexibility hypothesis and our demonstration of a negative but nonlinear relationship between the payout ratio and the growth rate. More importantly, we theoretically and empirically locate a structural change point for the relationship between the payout ratio and the growth rate. Rozeff (1982), Aivazian et al. (2003), Blau and Fuller (2008), and others conclude that a firm's risk and optimal payout are negatively related. Contrary to these conclusions and generally held beliefs, the dynamic optimization model developed here shows that the optimal payout is not necessarily negatively related to risk but implies that payout policies differ in high growth firms and low growth firms. High growth firms pay out their dividends for flexibility considerations, but low growth firms pay out their dividends to reduce their free cash flow problem. In addition, this paper also shows that the relationship between firms' payouts and their growth rates (or risks) can be affected by their risks (or growth rates). Contrary to earlier studies, the theoretical model developed here shows that the optimal payout ratio is not linearly related to the growth rate or to the risk of the firm, and whether the rate of return on assets is higher or lower than the growth rate has significant effect on the relationships between these variables. The results are borne out by rather extensive empirical research done in this paper based on 40 years of data collected in the USA from 1969 to 2009.

The stochastic dynamic optimization model developed here challenges many commonly held beliefs and results obtained from earlier studies. It provides more meaningful relationships among total risk and its separate components (systematic and firm-specific), growth rate, rate of return on assets, and optimal

[1]For example, Rozeff (1982), Jagannathan et al. (2000), Grullon et al. (2002), Aivazian et al. (2003), and Blau and Fuller (2008).

dividend policy. Our results represent an advance in corporate finance literature because all the important variables included in this study were never simultaneously included in those earlier studies, and more importantly, the interaction effects among these variables were neither recognized nor fully appreciated. We believe that this study is the first of its kind to provide a theoretical basis for relating dividend policies with firm risks, growth rates, and return on assets (ROAs) in an interrelated fashion.

The remainder of this paper is organized as follows: Section 79.2 contains a review of the literature on dividend-relevant theories and the findings of empirical work on dividend policy. In Sect. 79.3 we lay out a dynamic model used in subsequent sections to examine the existence, or nonexistence, of an optimal dividend policy. Section 79.4 provides the final expression of the optimal dividend policy of the firm derived from the optimization procedure to maximize firm value. In Sect. 79.5 we present both a detailed form and an approximated form of the relationship between the optimal dividend payout ratio and the growth rate. Section 79.6 includes a discussion of the effects of -market-dependent and market-independent components of risk on the optimal dividend policy. Section 79.7 includes empirical evidence supporting the model and implications in previous sections, and the conclusion appears in Sect. 79.8.

79.2 Review of the Literature

Corporate dividend policy has puzzled financial economists, dating back to the dividend irrelevance theorem proposed by M&M (1961). Since then, their controversial findings have been challenged and tested by weakening the assumptions or introducing imperfections into the analysis. The signaling models developed by Bhattacharya (1979), Miller and Rock (1985), and John and Williams (1985) suggest that because of the asymmetric information between managers and shareholders, managers use dividends as a signal to release private information to the market; however, empirical studies examining the signaling hypothesis have yielded mixed results. Nissim and Ziv (2001), Brook et al. (1998), Bernheim and Wantz (1995), Kao and Wu (1994), and Healy and Palepu (1988) support the signaling (asymmetric information) hypothesis by finding a positive association between dividend increases and future profitability. Kalay and Lowenstein (1986) and Asquith and Mullins (1983) find that dividend changes are positively associated with stock returns in the days surrounding dividend announcements, and Sasson and Kolodny (1976) identify that there is a positive association between the payout ratio and average rates of return. Studies by Benartzi et al. (1997), DeAngelo et al. (1996), and Grullon et al. (2002), however, reveal no support for the hypothesized relationship between dividend changes and future profitability.

Using agency cost theory, Easterbrook (1984) and Jensen (1986) propose the free cash flow hypothesis, arguing that because managers cannot credibly precommit to shareholders, they will not invest excess cash in negative-NPV projects. Dividend changes may convey information about how the firm will use future cash flow. Again, the results of empirical studies have been mixed at best. Several researchers, including Agrawal and Jayaraman (1994), Jensen et al. (1992), and Lang and Litzenberger (1989), find positive supports for the agency cost hypothesis, but others find no support for this hypothesis, for example, Howe et al. (1992), Denis et al. (1994), and Yoon and Starks (1995).

DeAngelo and DeAngelo (2006) argue that M&M's (1961) dividend irrelevance occurs because they consider either paying out all earnings or not paying any of them at all; therefore, payout policy might have impact on firm value if partial payout is allowed. In other words, DeAngelo and DeAngelo (2006) use the dividend flexibility hypothesis to show their dividend relevance results. Blau and Fuller (2008) have theoretically and empirically shown that the dividend flexibility hypothesis is indeed a reasonable dividend policy. In addition, several empirical studies show evidence of firms preferring financial flexibility, for example, Lie (2005), DeAngelo et al. (2006), Denis and Osobov (2008), and Gabudean (2007).

Besides focusing on the relevance of dividend policy, a growing body of literature deals with the determinants of optimal dividend payout policy. For example, Rozeff (1982) shows that the optimal dividend payout is related to the fraction of insider holdings, the growth of the firm, and the firm's beta coefficient. He also finds evidence that the optimal dividend payout is negatively correlated to beta risk, and supporting that beta risk reflects the leverage level of a firm. Jagannathan et al. (2000) empirically show that operation risk is negatively related to the propensity to increase payouts. Grullon et al. (2002) show that dividend changes are related to the change in the growth rate and the change in the rate of return on assets. They also find that dividend increases should be associated with subsequent declines in profitability and risk. Aivazian et al. (2003) examine eight emerging markets and show that, similar to US firms, dividend policies in emerging markets can also be explained by profitability, debt, and the market-to-book ratio. None of the foregoing scholars, however, has a solid theoretical model to support their finding.

We use these authors' works as a springboard to deal with the flexibility hypothesis and develop the theoretical underpinnings and a stochastic dynamic optimization model determining the conditions for an optimal dividend policy. We believe that ours is the first attempt ever made to include growth rate, risks, and more importantly, the rate of return and its distribution parameters in a single model deriving the conditions for an optimal dividend policy; furthermore, this paper empirically tests the findings and conclusions of the theoretical model developed here, using an extensive data set involving 28,333 dividend-paying firm-years from 1969 to 2009. The results of the empirical tests confirm and validate the findings of the theoretical model and provide new insights in the determination of optimal dividend policy, the influence of each explanatory variable listed above, and more importantly, their interaction effects on the corporate dividend policy. Notably, some of these results run counter to categorically stated conclusions reached in earlier studies.

79.3 The Model

We develop the dividend policy model under the assumptions that the capital markets represent the closest approximation to the economists' ideal of a perfect market – zero transaction costs, rational behavior on the part of investors, and the absence of tax differentials between dividends and capital gains. It is assumed that the firm is not restricted to financing its growth only by retained earnings and that its rate of return, $\widetilde{r}(t)$, is a nonstationary random variable, normally distributed with mean, μ, and variance, $\sigma(t)^2$.

Let $A(0)$ represent the initial assets of the firm and h be the growth rate. Then the earnings of this firm are given by Eq. 79.1, which is

$$\widetilde{x}(t) = \widetilde{r}(t)A(0)e^{ht} \tag{79.1}$$

where $\widetilde{x}(t)$ represents the earnings of the firm, and the tilde ($\sim$) denotes its random character.

Based upon DeAngelo and DeAngelo's (2006) assumption, we allow that the firm partially pay out its earnings and retain a certain amount of earnings in support of its growth.[2] The retained earnings of the firm, $y(t)$, can therefore be expressed as follows:

$$y(t) = \widetilde{x}(t) - m(t)\widetilde{d}(t), \tag{79.2}$$

where $\widetilde{d}(t)$ is the dividends per share and $m(t)$ is the total number of shares outstanding at time t.

Equation 79.2 further indicates that the focus of the firm's decision making is on retained earnings, which implies that dividend $\widetilde{d}(t)$ also becomes a random variable. The growth of a firm can be financed by retained earnings or by issuing new equity.

The new equity raised by the firm at time t can be defined as follows:

$$e(t) = \delta p(t)\dot{m}(t), \tag{79.3}$$

Where

$p(t)$ = price per share;

$\dot{m}(t) = dm(t)/dt;$

δ = degree of market perfection, $0 < \delta \leq 1$.

The value of δ equal to one indicates that new shares can be sold by the firm at current market prices.

[2]DeAngelo and DeAngelo (2006) carefully explain why partial payout is important to obtain an optimal ratio under perfect markets. In addition, they also argue that partial payout is important to avoid the suboptimal solution for optimal dividend policy.

From Eqs. 79.1, 79.2 and 79.3, investment in period t is the sum of retained earnings and funds raised by new equity, so the investment in period t can be written as follows:

$$hA(0)e^{ht} = \tilde{x}(t) - m(t)\tilde{d}(t) + \delta\dot{m}(t)p(t). \tag{79.4}$$

This implies that

$$\tilde{d}(t) = \left\{[\tilde{r}(t) - h]A(0)e^{ht} + \delta\dot{m}(t)p(t)\right\}/m(t), \tag{79.5}$$

and the mean and variance of the dividends per share can be expressed as follows:

$$E\left[\tilde{d}(t)\right] = \left\{[\mu - h]A(0)e^{ht} + \delta\dot{m}(t)p(t)\right\}/m(t) Var\left[\tilde{d}(t)\right] = A(0)^2\sigma(t)^2e^{2th}/m^2(t). \tag{79.6}$$

Also, let us postulate an exponential utility function of the following form[3]:

$$U\left[\tilde{d}(t)\right] = -e^{\alpha\tilde{d}(t)}, \text{ where } \alpha > 0. \tag{79.7}$$

Following the moment generating function, we have

$$E\left(-e^{-\alpha\tilde{d}(t)}\right) = -e^{-\alpha E\left[\tilde{d}(t)\right] + \frac{\alpha^2}{2}Var\left[\tilde{d}(t)\right]}, \tag{79.8}$$

where $\overline{d}(t)$ is the certainty equivalent value of $\tilde{d}(t)$.[4]

From Eqs. 79.6 and 79.8, the certainty equivalent dividend stream can be written as

$$\overline{d}(t) = \frac{(\mu - h)A(0)e^{th} + \delta\dot{m}(t)p(t)}{m(t)} - \frac{\alpha' A(0)^2\sigma(t)^2e^{2th}}{m(t)^2}, \tag{79.9}$$

where $\alpha' = \alpha/2$. Therefore, taking advantage of exponential utility, we can obtain a risk adjusted dividend stream. Furthermore, $\overline{d}(t)$ will reduce to the certainty case if we assume $\sigma(t)^2 = 0$.

In accordance with the capital asset pricing theory developed by Sharpe (1964), Lintner (1963), and Mossin (1966), the total risk can be decomposed into systematic risk and unsystematic risk; that is, $\tilde{r}(t)$ can be defined as follows:

$$\tilde{r}(t) = a + b\tilde{I}(t) + \tilde{\varepsilon}(t), \tag{79.10}$$

[3]Pratt (1964) provides a detailed analysis of the various utility functions. Exponential, hyperbolic, and quadratic forms have been variously used in the literature, but the first two seem to have preference over the quadratic form because the latter has the undesirable property that it ultimately turns downwards.

[4]From the moment generating function discussed in Hogg and Craig (2004), we know that $E(-e^{ty}) = -e^{tE(y) + \frac{1}{2}t^2Var(y)}$. Let $t = -\alpha$, then the right-hand side of (8) is easily obtained.

where $\widetilde{I}(t)$ is the market index; $\widetilde{\varepsilon}(t) \sim N\left(0, \sigma_{\varepsilon}^{2}\right)$; a and b are regression parameters; and $Var\left(b\widetilde{I}(t)\right)$ and $Var(\widetilde{\varepsilon}(t))$ represent the systematic and unsystematic risk, respectively.

Following Eqs. 79.10 and 79.6 can be rewritten as

$$\begin{aligned} E\left[\widetilde{d}(t)\right] &= \left[\left(a+b\bar{I}-h\right)A(o)e^{ht} + \delta\dot{m}(t)p(t)\right]/m(t)Var\left[\widetilde{d}(t)\right] \\ &= A(0)^{2}\left[b^{2}Var\left(\widetilde{I}(t)\right) + Var(\varepsilon(t))\right]e^{2th}/m(t)^{2} \\ &= A(0)^{2}\left[\rho(t)^{2}\sigma(t)^{2} + \left(1-\rho(t)^{2}\right)\sigma(t)^{2}\right]e^{2th}/m(t)^{2}, \end{aligned} \tag{79.11}$$

where
$\rho(t)$ = the correlation coefficient between $\widetilde{r}(t)$ and $\widetilde{I}$;
a = market-independent component of the firm's rate of return;
$b\bar{I}$ = market-dependent component of the firm's rate of return;
$\rho(t)^{2}\sigma(t)^{2}$ = nondiversifiable risk; and
$(1-\rho(t)^{2})\sigma(t)^{2}$ = diversifiable risk.

The unsystematic risk usually can be diversified away by the investors,[5] so the certainty equivalent value in Eq. 79.9 should be revised as

$$\hat{d}'(t) = \frac{\left(a+b\bar{I}-h\right)A(o)e^{th} + \delta\dot{m}(t)p(t)}{m(t)} - \frac{\alpha' A(o)^{2}\rho(t)^{2}\sigma(t)^{2}e^{2th}}{m(t)^{2}}. \tag{79.12}$$

Following Lintner (1962), we observe that the stock price should equal the present value of this certainty equivalent dividend stream discounted at a riskless rate of return. Therefore,

$$p(0) = \int_{0}^{T} \hat{d}'(t)e^{-kt}dt, \tag{79.13}$$

where
$p(0)$ = the stock price at $t = 0$;
k = the risk free rate of return;
T = the planning horizon.

This model will be used in subsequent sections to find the functional form of $m(t)$ and optimize the payout ratio. The formulation of our model is different from that of M&M (1961), Gordon (1962), Lerner and Carleton (1966a), and Lintner (1964). For example, in contrast to our model, M&M consider neither the nonstationarity of the firm's rate of return in their model nor explicitly incorporated uncertainty in their valuation model. Their models are also essentially static and would not permit an extensive analysis of the dynamic process of moving from one equilibrium state to another; furthermore, the formulation of our model is different from those who propose to capitalize the market-dependent and market-independent components of

[5] See Lintner (1965), Mossin (1966), and Sharpe (1964).

the uncertain stream of earnings at the risky and riskless rates, respectively.[6] Instead, we view the market value of a firm as the present value of certainty equivalents of random future receipts. In the next section, we carry out the optimization of Eq. 79.13 and derive the final expression for the optimal payout ratio.[7]

79.4 Optimum Dividend Policy

Based upon the evaluation model developed in the previous section, in this section we will derive an optimal dividend payout ratio.

Substituting Eq. 79.12 into Eq. 79.13, we obtain

$$p(0) = \int_0^T \left[\frac{\left(A + b\bar{I} - h\right)A(0)e^{th} + \delta \dot{m}(t)p(t)}{m(t)} - \frac{\alpha' A(0)^2 \rho(t)^2 \sigma(t)^2 e^{2th}}{m(t)^2} \right] e^{-kt} dt. \tag{79.14}$$

To maximize Eq. 79.14, we observe that

$$p(t) = \int_t^T \hat{d}'(s) e^{-k(s-t)} ds = e^{kt} \int_t^T \hat{d}'(s) e^{-ks} ds, \tag{79.15}$$

where $s =$ the proxy of time in the integration.

From Eq. 79.15, we can formulate a differential equation as

$$\frac{dp(t)}{d(t)} = \dot{p}(t) = kp(t) - \hat{d}'(t). \tag{79.16}$$

Substituting Eq. 79.12 into Eq. 79.16, we obtain the differential equation

$$\dot{p}(t) + \left[\delta \frac{\dot{m}(t)}{m(t)} - k \right] p(t) = -G(t), \tag{79.17}$$

Where

$$G(t) = \frac{\left(a + b\bar{I} - h\right)A(0)e^{th}}{m(t)} - \frac{\alpha' A(o)^2 \rho(t)^2 \sigma(t)^2 e^{2th}}{m(t)^2}. \tag{79.18}$$

Following Kreyszig (2010) and Lee and Shi (2010), we solve the differential Eq. 79.17 and obtain the solution as indicated in Eq. 79.19.

$$p(t) = \frac{e^{kt}}{m(t)^\delta} \int_t^T G(s) m(s)^\delta e^{-ks} ds. \tag{79.19}$$

[6]See Brennan (1973).

[7]For further explanation of the optimization of the deterministic and stochastic control models and their applications to economic problems, please see Aoki (1967), Bellman (1990 and 2003), and Intriligator (2002).

Then, Eq. 79.20 can be obtained from Eqs. 79.18 and 79.19, implying that the initial value of a stock can be expressed as the summation of present values of its earnings stream adjusted by the risk taken by the firm.

$$p(0) = \frac{1}{m(0)^{\delta}} \int_0^T \left\{ \left(a + b\overline{I} - h\right) A(0) e^{th} m(t)^{\delta-1} - \alpha' A(0)^2 \rho(t)^2 \sigma(t)^2 e^{2th} m(t)^{\delta-2} \right\} e^{-kt} dt. \tag{79.20}$$

To maximize firm value, the Euler-Lagrange condition for the optimization of $p(o)$ is given by Eq. 79.21,[8]

$$(\delta - 1)\left(a + b\overline{I} - h\right) A(0) e^{th} m(t)^{\delta-2} - \alpha' A(0)^2 \rho(t)^2 \sigma(t)^2 e^{2th} m(t)^{\delta-3} (\delta - 2) = 0. \tag{79.21}$$

Therefore, the optimal shares outstanding at time t can be derived.

$$m(t) = \frac{(2-\delta)\alpha' A(0) e^{th} \rho(t)^2 \sigma(t)^2}{(1-\delta)\left(a + b\overline{I} - h\right)}. \tag{79.22}$$

From Eqs. 79.18, 79.19 and 79.22, we can obtain the maximized stock value

$$p(t) = \frac{\left(a + b\overline{I} - h\right)^2 (1-\delta) e^{kt - th\delta} \int_t^T e^{\delta hs - ks} \left(\rho(s)\sigma(s)\right)^{2\delta-2} ds}{\alpha' (2-\delta)^2 \rho(t)^{2\delta} \sigma(t)^{2\delta}}. \tag{79.23}$$

From Eq. 79.22, we also obtain the optimal number of shares of new equity issued at time t

$$\dot{m}(t) = \frac{\left\{ h(2-\delta)\alpha' A(0) e^{th} \rho(t)^2 \sigma(t)^2 + (2-\delta)\alpha' A(0) e^{th} [\rho(t)^2 \dot{\sigma}(t)^2 + \sigma(t)^2 \dot{\rho}(t)^2 \right\}}{(1-\delta)\left(a + b\overline{I} - h\right)}. \tag{79.24}$$

From Eqs. 79.23 and 79.24, we have the amount generated from issuing new equity

$$\dot{m}(t)p(t) = \frac{\left(a + b\overline{I} - h\right) e^{kt-(\delta-1)th} A(0) \left(h\rho(t)^2 \sigma(t)^2 + \rho(t)^2 \dot{\sigma}(t)^2 + \sigma(t)^2 \dot{\rho}(t)^2\right) \int_t^T e^{s(\delta h - k)} \left(\rho(s)\sigma(s)\right)^{2\delta-2} ds}{(2-\delta)\rho(t)^{2\delta}\sigma(t)^{2\delta}}. \tag{79.25}$$

From Eqs. 79.5 and 79.25, we can obtain $\overline{D}(t) = m(t)\overline{d}(t)$. From Eqs. 79.1 and 79.10, we can obtain $\overline{x}(t) = \left(a + b\overline{I}\right) A(0) e^{ht}$. When δ approaches unity, we can derive the optimal payout ratio as

---

[8]For the derivation of Eq. 79.21, please refer to the Appendix 2.

$$\frac{\overline{D}(t)}{\overline{x}(t)} = \frac{\left(a + b\overline{I} - h\right)}{\left(a + b\overline{I}\right)} \left\{ 1 + \frac{\left[e^{(h-k)(T-t)} - 1\right]}{(h-k)} \left(h + \frac{\dot{\sigma}(t)^2}{\sigma(t)^2} + \frac{\dot{\rho}(t)^2}{\rho(t)^2} \right) \right\}. \quad (79.26)$$

Equation 79.26 implies an optimal payout ratio when we use an exponential utility function to derive the stochastic dynamic dividend policy model. This result does not necessarily imply that the dividend policy results derived by M&M (1961) are false because we allowed free cash flow to be paid out partially as assumed by DeAngelo and DeAngelo (2006) instead of paying out all free cash flows as assumed by M&M (1961).

In the following section, we use Eq. 79.26 to explore the implications of the stochasticity, the stationarity (in the strict sense), and the nonstationarity of the firm's rate of return for its dividend policy.[9] We also investigate in detail the differential effects of variations in the systematic and unsystematic risk components of the firm's stream of earnings on the dynamics of its dividend policy.

79.5 Relationship Between the Optimal Payout Ratio and the Growth Rate

In this section, we investigate the relationship between the optimal payout ratio and the growth rate in terms of both exact and approximate approaches. Taking the partial derivative of Eq. 79.26 with respective to the growth rate, we obtain

$$\begin{aligned}\frac{\partial\left[\overline{D}(t)/\overline{x}(t)\right]}{\partial h} &= \left(-\frac{1}{a + b\overline{I}}\right)\left[\frac{-k + he^{(h-k)(T-t)}}{h-k}\right] \\ &\quad + \left(1 - \frac{h}{a + b\overline{I}}\right)\left[\frac{[(-k) + h(h-k)(T-t)]e^{(h-k)(T-t)} + k}{(h-k)^2}\right].\end{aligned} \quad (79.27)$$

The sign of Eq. 79.27 is affected not only by the growth rate (h) but also by the expected rate of return on assets ($a + b\overline{I}$), the duration of future dividend payments ($T-t$), and the cost of capital (k).

Since the sign of Eq. 79.27 cannot be analytically determined, we use a sensitivity analysis approach to investigate the sign of Eq. 79.27. Table 79.1 shows the sign of partial derivatives of Eq. 79.27 under different values of the growth rate and the rate of return on assets as well as duration ($T-t$). We find that the relationship between the optimal payout ratio and the growth rate is always negative when the growth rate is higher than the rate of return on assets. If the growth rate is lower than the rate of return on asset, the direction of relationship essentially depends on the duration of the dividend payment ($T-t$). We find that the sign of Eq. 79.27 is

[9] See Hamilton (1994), pp. 45–46.

negative if the duration ($T - t$) is small and the growth rate and rate of return on assets are within a reasonable range. In addition, the curved lines in Table 79.1 also indicate a nonlinear relationship between the growth rate and the optimal payout ratio. We can therefore conclude that the relationship between the optimal payout ratio and the growth rate is nonlinear and generally negative.

Based upon Eq. 79.27, Fig. 79.1 plots the change in the optimal payout ratio with respect to the growth rate in different durations of dividend payments and costs of capital. We find a negative relationship between the optimal payout ratio and the growth rate, indicating that a firm with a higher rate of return on assets tends to pay out less when its growth opportunities increase. Moreover, a firm with a lower growth rate and higher expected rate of return will not decrease its payout when its growth opportunities increase, but a firm with a lower growth and a higher expected rate of return on asset is not a general case in the real world. We also find that the duration of future dividend payments is an important determinant of the dividend payout decision, but the effect on the cost of capital is relatively minor.

In the finite growth case, if $(h–k)(T–t) < 1$, then following the Maclaurin expansion, the optimal payout ratio under no change in risk defined in Eq. 79.26 can be written as

$$\left[\overline{D}(t)/\overline{x}(t)\right] \approx \left(1 - \frac{h}{a + b\overline{I}}\right)(1 + h(T - t)). \tag{79.28}$$

The partial derivative of Eq. 79.28 with respective to the growth rate is

$$\frac{\partial\left[\overline{D}(t)/\overline{x}(t)\right]}{\partial h} \approx \left(\frac{\left[\left(a + b\overline{I}\right) - h\right](T - t) - h(T - t) - 1}{a + b\overline{I}}\right). \tag{79.29}$$

Equation 79.29 indicates that the relationship between the optimal dividend payout and the growth rate depends on firm's level of growth, the rate of return on assets, and the duration of future dividend payment.[10] Equation 79.29 is negative when the rate of return on assets is lower than the growth rate. This implies that the firm will reduce its payout when its growth rate increases. The higher growth rate (h) and the lower rate of return on assets ($a + b\overline{I}$) will lead to a more negative relationship between dividend payout ratio and growth rate. More specifically, the condition of Eq. 79.29 will lead to a negative relationship between the optimal payout ratio and the growth rate.

$$h > \frac{1}{2}\left[\left(a + b\overline{I}\right) - \frac{1}{(T - t)}\right]. \tag{79.30}$$

Consistent with the sensitivity analysis of Eq. 79.27, when a firm with a high growth rate or a low rate of return on assets faces a growth opportunity, it will decrease its dividend payout to generate more cash to meet such a new investment.

A possible explanation is that high growth firms need more retained earnings to meet their future growth opportunities because the growth rate is the main

[10]Please see the Appendix 3 for the derivation of Eq. 79.28 and 79.29.

Table 79.1 Sensitivity analysis of the relationship between the optimal payout and the growth rate

Growth (%)	ROA=5 %	ROA=10 %	ROA=15 %	ROA=20 %	ROA=25 %	ROA=30 %	ROA=35 %	ROA=40 %
Panel A. $T-t=3$, Cost of capital = 7 %								
1	−18.32	−7.77	−4.25	−2.49	−1.44	−0.73	−0.23	0.15
2	−19.39	−8.26	−4.55	−2.70	−1.58	−0.84	−0.31	0.08
3	−20.52	−8.78	−4.87	−2.92	−1.74	−0.96	−0.40	**0.02**
4	−21.69	−9.32	−5.20	−3.14	−1.91	−1.08	−0.49	**−0.05**
5	−22.91	−9.89	−5.55	−3.38	−2.08	−1.21	−0.59	−0.13
6	−24.19	−10.49	−5.92	−3.63	−2.26	−1.35	−0.70	−0.21
7	−25.56	−11.12	−6.31	−3.90	−2.46	−1.50	−0.81	−0.29
8	−26.92	−11.75	−6.70	−4.17	−2.65	−1.64	−0.92	−0.38
9	−28.38	−12.43	−7.12	−4.46	−2.86	−1.80	−1.04	−0.47
10	−29.90	−13.14	−7.55	−4.76	−3.09	−1.97	−1.17	−0.57
11	−31.48	−13.88	−8.01	−5.08	−3.32	−2.14	−1.31	−0.68
12	−33.14	−14.65	−8.49	−5.41	−3.56	−2.33	−1.45	−0.79
13	−34.86	−15.46	−8.99	−5.76	−3.81	−2.52	−1.60	−0.90
14	−36.66	−16.30	−9.51	−6.12	−4.08	−2.72	−1.75	−1.03
15	−38.54	−17.18	−10.06	−6.50	−4.36	−2.94	−1.92	−1.16
16	−40.50	−18.10	−10.63	−6.89	−4.65	−3.16	−2.09	−1.29
17	−42.54	−19.05	−11.22	−7.31	−4.96	−3.39	−2.28	−1.44
18	−44.67	−20.05	−11.85	−7.74	−5.28	−3.64	−2.47	−1.59
19	−46.89	−21.09	−12.49	−8.20	−5.62	−3.90	−2.67	−1.75
20	−49.20	−22.18	−13.17	−8.67	−5.97	−4.17	−2.88	−1.91
Panel B of Table 79.1 $T-t=5$, Cost of capital = 7 %								
1	−17.33	−6.45	−2.83	−1.01	**0.07**	0.80	1.32	1.70
2	−18.99	−7.18	−3.24	−1.27	**−0.09**	0.70	1.26	1.68
3	−20.77	−7.96	−3.68	−1.55	−0.27	0.59	1.20	1.66

4	−22.70	−8.80	−4.17	−1.85	−0.46	0.46	1.13	1.62
5	−24.76	−9.71	−4.69	−2.18	−0.68	0.33	1.04	1.58
6	−26.97	−10.69	−5.26	−2.54	−0.91	0.17	0.95	1.53
7	−29.44	−11.78	−5.89	−2.95	−1.18	< **0.01**	0.84	1.47
8	−31.90	−12.87	−6.53	−3.35	−1.45	**−0.18**	0.72	1.40
9	−34.63	−14.09	−7.24	−3.81	−1.76	−0.39	0.59	1.32
10	−37.57	−15.39	−8.00	−4.31	−2.09	−0.61	0.44	1.23
11	−40.71	−16.80	−8.83	−4.85	−2.45	−0.86	0.28	1.13
12	−44.07	−18.31	−9.72	−5.43	−2.85	−1.13	**0.09**	1.01
13	−47.67	−19.93	−10.68	−6.05	−3.28	−1.43	**−0.11**	0.88
14	−51.53	−21.66	−11.71	−6.73	−3.74	−1.75	−0.33	0.74
15	−55.66	−23.52	−12.81	−7.46	−4.25	−2.10	−0.57	0.57
16	−60.07	−25.52	−14.00	−8.25	−4.79	−2.49	−0.84	0.39
17	−64.79	−27.66	−15.28	−9.09	−5.38	−2.90	−1.14	**0.19**
18	−69.84	−29.95	−16.65	−10.01	−6.02	−3.36	−1.46	**−0.03**
19	−75.23	−32.40	−18.13	−10.99	−6.70	−3.85	−1.81	−0.28
20	−81.00	−35.03	−19.70	−12.04	−7.45	−4.38	−2.19	−0.55
Panel C of Table 79.1 $T-t = 15$, Cost of capital = 7 %								
1	−13.56	−1.52	2.49	4.50	5.71	6.51	7.08	7.51
2	−17.06	−2.56	2.27	4.69	6.14	7.11	7.80	8.31
3	−21.34	−3.89	1.93	4.84	6.58	7.75	8.58	9.20
4	−26.58	−5.57	1.43	4.93	7.03	8.43	9.43	10.18
5	−32.96	−7.69	**0.73**	4.94	7.47	9.15	10.36	11.26
6	−40.72	−10.34	**−0.21**	4.85	7.89	9.91	11.36	12.45
7	−51.15	−14.03	−1.66	4.53	8.24	10.72	12.48	13.81
8	−61.58	−17.72	−3.10	4.21	8.59	11.52	13.61	15.17
9	−75.40	−22.75	−5.21	3.57	8.83	12.34	14.85	16.73

(*continued*)

Table 79.1 (continued)

Growth (%)	ROA=5 %	ROA=10 %	ROA=15 %	ROA=20 %	ROA=25 %	ROA=30 %	ROA=35 %	ROA=40 %
10	−92.10	−28.94	−7.89	2.63	8.95	13.16	16.17	18.42
11	−112.25	−36.53	−11.29	**1.33**	8.91	13.95	17.56	20.26
12	−136.53	−45.80	−15.55	**−0.43**	8.64	14.69	19.01	22.25
13	−165.74	−57.09	−20.88	−2.77	8.10	15.34	20.51	24.39
14	−200.85	−70.83	−27.49	−5.82	7.18	15.85	22.04	26.69
15	−243.01	−87.50	−35.67	−9.75	5.80	16.17	23.57	29.13
16	−293.57	−107.70	−45.74	−14.77	3.82	16.21	25.06	31.70
17	−354.16	−132.13	−58.11	−21.11	**1.10**	15.90	26.47	34.40
18	−426.70	−161.62	−73.25	−29.07	**−2.57**	15.11	27.73	37.20
19	−513.48	−197.17	−91.74	−39.02	−7.39	13.70	28.76	40.06
20	−617.19	−239.98	−114.24	−51.37	−13.65	11.49	29.46	42.93

This table shows the sensitivity analysis of the relationship between the optimal payout and the growth rate. Panel A presents, when the duration of dividend payments (T–t) is equal to 3 years and the cost of capital is equal to 7 %, the values of Eq. 79.32 under different settings of the rate of return on assets and the growth rate. Similar to Panel A, Panel B and Panel C present the values of Eq. 79.32 when the duration of dividend payments is equal to 3 years and 5 years and the cost of capital is equal to 7 %

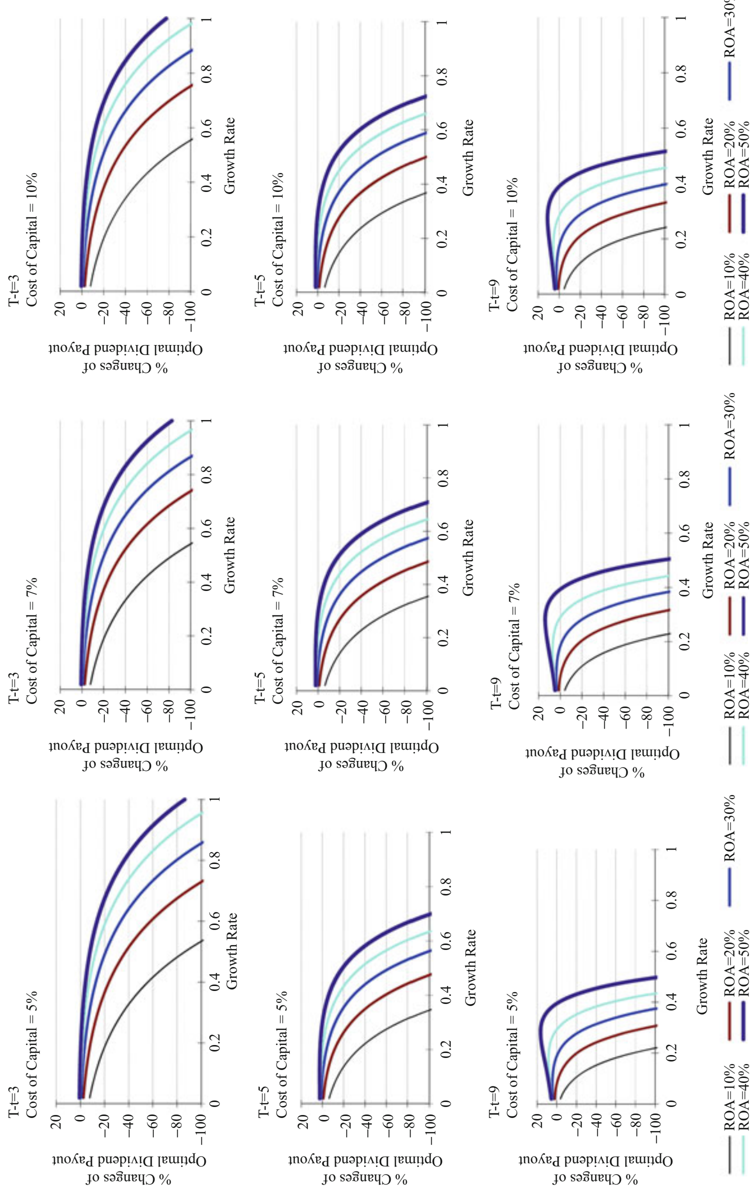


Fig. 79.1 **Sensitivity analysis of the relationship between the optimal payout and the growth rate**. The figures show the sensitivity analysis of the relationship between the optimal payout and the growth rate. In each figure, each *line* shows the percentage change of the optimal dividend payout with a 1 % change in the growth rate. The *various lines* represent different levels of the rate of return on assets. The figures present different durations of dividend payments ($T-t$) and costs of capital

determinant of value in the case of such companies, but low growth firms do not need more earnings to maintain their low growth perspective and can afford to increase their payouts. Based on the flexibility concerns, the relationship between firms' payout ratios and their growth rates is therefore negative.

79.6 Relationship Between Optimal Payout Ratio and Risks

Equation 79.26 implies that the optimal payout ratio is a function of the expected profitability rate ($a + b\bar{I}$), growth rate (h), cost of capital (k), age (T–t), total risk ($\sigma(t)^2$), and the correlation coefficient between profit and market rate of return ($\rho(t)^2$). In addition, Eq. 79.26 is also a function of two dynamic variables – the relative time rate of change in the total risk of the firm, $\left[\dot{\sigma}(t)^2/\sigma(t)^2\right]$, and the relative time rate of change in the covariability of the firm's earnings with the market, $\left[\dot{\rho}(t)^2/\rho(t)^2\right]$. This theoretical dynamic relationship between the optimal payout ratio and other determinants can be used to do empirical studies to determine dividend policy. The dynamic effects of variations in $\left[\dot{\sigma}(t)^2/\sigma(t)^2\right]$ and $\left[\dot{\rho}(t)^2/\rho(t)^2\right]$ on the time path of optimal payout ratio can be investigated under the following four cases: (1) changes in total risk, (2) changes in correlation between profit and the market rate of return (i.e., systematic risk), (3) changes in total risk and systematic risk, and (4) no changes in risk.

79.6.1 Case 1: Total Risk

First, we examine the effect of $\left[\dot{\sigma}(t)^2/\sigma(t)^2\right]$ on the optimal payout ratio. By differentiating Eq. 79.26 with respect to $\left[\dot{\sigma}(t)^2/\sigma(t)^2\right]$, we obtain

$$\frac{\partial\left[\overline{D}(t)/\overline{x}(t)\right]}{\partial\left[\dot{\sigma}(t)^2/\sigma(t)^2\right]} = \left(1 - \frac{h}{a + b\bar{I}}\right)\left[\frac{e^{(h-k)(T-t)} - 1}{h - k}\right]. \tag{79.31}$$

In Eq. 79.31 the cost of capital, k, can be either larger or smaller than the growth rate h. We can show that $\frac{e^{(h-k)(T-t)}-1}{h-k}$ is always larger than 0, regardless of whether k is larger or smaller than h.[11] Thus, the sign of Eq. 79.31 depends on the sign of $\left(1 - \frac{h}{a+b\bar{I}}\right)$, which depends on the growth rate h relative to $\left(a + b\bar{I}\right)$.

[11]If $h > k$, then $e^{(h-k)(T-t)} > 1$, and both the numerator and denominator are greater than zero, resulting in a positive value; if $h < k$, then $e^{(h-k)(T-t)} < 1$, and both the numerator and denominator are less than zero, resulting in a positive value. Thus, $\frac{e^{(h-k)(T-t)}-1}{h-k}$ is always larger than 0, regardless of whether k is larger or smaller than h.

If the growth rate h is equal to $\left(a + b\bar{I}\right)$, then $\left(1 - \frac{h}{a+b\bar{I}}\right)$ is equal to zero. Equation 79.31 is thus zero, and the change in total risk will not affect the payout ratio because the first derivative of the optimized payout ratio, Eq. 79.26, with respect to $\left[\dot{\sigma}(t)^2/\sigma(t)^2\right]$ is always zero.

If growth rate h is larger than $\left(a + b\bar{I}\right)$, then the entire first derivative of Eq. 79.26 with respect to $\left[\dot{\sigma}(t)^2/\sigma(t)^2\right]$ is negative (i.e., Eq. 79.31 is negative); furthermore, $h > \left(a + b\bar{I}\right)$ implies that the growth rate of a firm is larger than its expected profitability rate. An alternative case is $h < \left(a + b\bar{I}\right)$, which implies that the growth rate of a firm is less than its expected profitability rate. This situation can occur when a company is either in a low growth, no growth, or negative growth stage. Under this situation, a company will increase its payout ratio as shown in Eq. 79.31. If $h < \left(a + b\bar{I}\right)$, then Eq. 79.31 is positive, indicating that a relative increase in the risk of the firm would increase its optimal payout ratio. This implies that a relative increase in the total risk of the firm would decrease its optimal payout ratio. Lintner (1965) and Blau and Fuller (2008) have found this kind of relationship, yet they did not theoretically show how it can be derived.

Jagannathan et al. (2000) empirically show that operation risk is negatively related to the propensity to increase payouts in general and dividends in particular. Our theoretical analysis in terms of Eq. 79.31 shows that the change in total risk is negatively or positively related to the payout ratio, conditional on the higher growth rate relative to the expected profitability rate. We find negative relationships between payout and the change in total risk for high growth firms $\left(h > \left(a + b\bar{I}\right)\right)$. A possible explanation is that in the case of high growth firms, a firm must reduce the payout ratio and retain more earnings to build up "precautionary reserves," which become all the more important for a firm with volatile earnings over time. High growth firms thus tend to retain more earnings when they face higher risk. By contrast, in the case of established low growth firms $\left(h < \left(a + b\bar{I}\right)\right)$, low growth firms are likely to be more mature and have most likely already built such reserves over time. They probably do not need more earnings to maintain their low growth perspective and can afford to increase the payout; consequently, when facing higher risk on their earnings, low growth firms can reduce their shareholders' risk by paying more dividends to their shareholders.

The age of the firm (T–t), which is one of the variables in Eq. 79.31, becomes an important factor because the very high growth firms are also the newer firms with very little builtup precautionary reserves.

Under more dynamic conditions, we provide further evidence of the validity of Lintner's (1965) observations that, ceteris paribus, optimal dividend payout ratios vary directly with the variance of the firm's profitability rates. The rationale for such relationships, even when the systematic risk concept is incorporated into the analysis, is obvious, that is, holding $\rho(t)^2$ constant and letting the $\sigma(t)^2$ increase imply that the covariance of the firm's earnings with the market does not change though its relative proportion to the total risk increases.

79.6.2 Case 2: Systematic Risk

To examine the effect of a relative change in $\left[\dot{\rho}(t)^2/\rho(t)^2\right]$ (i.e., systematic risk) on the dynamic behavior of the optimal payout ratio, we differentiate Eq. 79.26 to obtain

$$\frac{\partial\left[\overline{D}(t)/\overline{x}(t)\right]}{\partial\left[\dot{\rho}^2(t)/\rho^2(t)\right]} = \left(1 - \frac{h}{a + b\overline{I}}\right)\left[\frac{e^{(h-k)(T-t)} - 1}{h-k}\right]. \tag{79.32}$$

The sign of Eq. 79.32 can be analyzed as with Eq. 79.31, so the conclusions of Eq. 79.32 are similar to those of Eq. 79.31. A relative change in $\rho(t)^2$ can either decrease or increase the optimal payout ratio, all things being equal. The effect of nonstationarity in the firm's nondiversifiable risk would tend to be obliterated should both the systematic and the unsystematic components of total risk not be clearly identified in the expression for optimal payout ratio. Although the total risk of the firm is stationary (i.e., $\left[\dot{\sigma}(t)^2/\sigma(t)^2\right]$ is equal to zero), a change in the total risk complexion of the firm could still conceivably occur because of an increase or decrease in the covariability of its earnings with the market. Equations 79.26 and 79.32 clearly identify the effect of such a change in the risk complexion of the firm on its optimal payout ratio.

An examination of Eq. 79.26 indicates that only when the firm's earnings are perfectly correlated with the market (i.e., $\rho^2 = 1$), whether the management arrives at its optimal payout ratio using the total variance concept of risk or the market concept of risk does not matter. For every other case, the optimal payout ratio followed by management using the total variance concept of risk would be an overestimate of the true optimal payout ratio for the firm based on the market concept of risk underlying the capital asset pricing theory.

Management may decide not to use the truly dynamic model and instead substitute an average of the long run systematic risk of the firm, but for $\dot{\rho}^2(t) > 0$, because the initial average is higher than the true $\rho^2(t)$, the management would pay out less or more in the form of dividends than is optimal. In other words, the payout ratio followed in the initial part of the planning horizon would be an overestimate or an underestimate of the optimal payout under truly dynamic specifications.

Rozeff (1982) empirically shows a negative relationship between the β coefficient (systematic risk) and the payout level. The theoretical analysis in terms of Eq. 79.32 provides a more detailed analytical interpretation of his findings. The explanations of these results are similar to those discussed for findings in previous sections.

79.6.3 Case 3: Total Risk and Systematic Risk

In our third case, we attempt to investigate the compounded effect of a simultaneous change in the total risk of the firm and also a change in its decomposition into the market- dependent and market-independent components. Taking the total differential of Eq. 79.26 with respect to $\left[\dot{\sigma}(t)^2/\sigma(t)^2\right]$ and $\left[\dot{\rho}(t)^2/\rho(t)^2\right]$, we obtain

$$d\left[\overline{D}(t)/\overline{x}(t)\right] = \gamma d\left(\frac{\dot{\sigma}(t)^2}{\sigma(t)^2}\right) + \gamma d\left(\frac{\dot{\rho}(t)^2}{\rho(t)^2}\right), \tag{79.33}$$

where $\gamma = \left(1 - \frac{h}{a+b\overline{I}}\right)\left[\frac{e^{(h-k)(T-t)} - 1}{h-k}\right]$. Also, γ can be either negative or positive, as shown above.

Now from Eq. 79.33, the greatest decrease or increase in the optimal payout ratio would obviously occur when both $\dot{\sigma}(t)^2$ and $\dot{\rho}(t)^2$ are positive. This implies that the total risk of the firm increases, and in addition, its relative decomposition into systematic and unsystematic components also changes, making the firm's earnings still more correlated with the market. Under this circumstance, the decrease or increase in the optimal payout would now represent the compounded effect of both these changes; however, it is conceivable that although $\dot{\sigma}(t)^2$ is positive, $\dot{\rho}(t)^2$ is negative, tending to offset the decrease or increase in the optimal payout ratio resulting from the former. Alternatively, $\dot{\sigma}(t)^2$ could be negative, indicating a reduction in the total risk of the firm and may offset the increase in the optimal payout ratio resulting from a positive $\dot{\rho}(t)^2$.

To what extent the inverse variations in the total risk and the risk complexion of the firm will offset each other's effects on the optimal payout ratio for the firm would, of course, be dependent upon the relative magnitudes of $\dot{\rho}(t)^2$ and $\dot{\sigma}(t)^2$. To see the precise trade-off between the two dynamic effects of $\left[\dot{\sigma}(t)^2/\sigma(t)^2\right]$ and $\left[\dot{\rho}(t)^2/\rho(t)^2\right]$ on the optimal payout ratio, let the total differential of Eq. 79.26, given in Eq. 79.33, be set equal to zero, yielding

$$d\left[\dot{\sigma}(t)^2/\sigma(t)^2\right] = -d\left[\dot{\rho}(t)^2/\rho(t)^2\right]. \tag{79.34}$$

Equation 79.34 implies that the relative increase (or decrease) in $\sigma(t)^2$ has a one-to-one correspondence with the relative decrease (or increase) in $p(t)^2$, so in Eq. 79.34 conditions are established for relative changes in $p(t)^2$ and $\sigma(t)^2$, which lead to a null effect on the optimal dividend payout ratio.

79.6.4 Case 4: No Change in Risk

Now we consider the least dynamic situation, in which no changes exist in total risk or systematic risk, assuming $\dot{\sigma}(t)^2 = 0$ and $\dot{\rho}(t)^2 = 0$. Under this circumstance, Eq. 79.26 reduces to

$$\left[\overline{D}(t)/\overline{x}(t)\right] = \left(1 - \frac{h}{a+b\overline{I}}\right)\left[\frac{-k + he^{(h-k)(T-t)}}{h-k}\right]. \tag{79.35}$$

Thus, when the firm's total risk and covariability of its earnings with the market are assumed stationary, Eq. 79.35 indicates that a firm's optimal payout ratio is independent of its risk. Notice that neither $\sigma(t)^2$ nor $\rho(t)^2$ now appears in the expression for the optimal payout ratio given in Eq. 79.35. These conclusions, like those of Wallingford (1972a, b), for example, run counter to the intuitively appealing and well-accepted theory of finance emphasizing the relevance of risk for financial decision making.[12] Our model clearly shows that the explanation for such unacceptable implications of the firm's total risk and its market-dependent and market-independent components for the firm's optimal payout policy lies, of course, in the totally unrealistic assumptions of stationarity underlying the derivation of such results as illustrated in Eq. 79.35.

79.7 Empirical Evidence

A growing body of literature focuses on the determinants of optimal dividend payout policy. Rozeff (1982), Jagannathan et al. (2000), Grullon et al. (2002), Aivazian et al. (2003), Blau and Fuller (2008), and others empirically investigate the determination of dividend policy, but none of them has a solid theoretical model to support their findings. Based upon our theoretical model and its implications discussed in the foregoing sections, we develop the three testable hypotheses that follow:

Hypothesis 1 Firms generally reduce their dividend payouts when their growth rates increase.

The negative relationship between the payout ratio and the growth ratio in our theoretical model implies that high growth firms need to reduce the payout ratio and retain more earnings to build up "precautionary reserves," but low growth firms are likely to be more mature and already build up their reserves for flexibility considerations. Rozeff (1982), Fama and French (2001), Blau and Fuller (2008), and others argue that high growth firms will have higher investment opportunities and tend to pay out less in dividends. Based upon flexibility concerns, we predict that high growth firms pay higher dividends. This result is obtained when risk factor is not explicitly considered.

Following Eqs. 79.31 and 79.32, we theoretically find that the relationship between the payout ratio and the risk can be either negative or positive, depending upon whether the growth rate is higher or lower than the rate of return on total assets. Based upon this finding, we develop two other hypotheses.

Hypothesis 2 The relationship between the firms' dividend payouts and their risks is negative when their growth rates are higher than their rates of return on asset.

High growth firms need to reduce the payout ratio and retain more earnings to build up "precautionary reserves," which become more important for a firm with

[12]For example, Lintner (1963).

volatile earnings over time. For flexibility considerations, high growth firms tend to retain more earnings when they face higher risk. This theoretical result is consistent with the flexibility hypothesis.

Hypothesis 3 The relationship between the firms' dividend payouts and their risks is positive when their growth rates are lower than their rates of return on asset.

Low growth firms are likely to be more mature and have most likely already built such reserves over time, and they probably do not need more earnings to maintain their low growth perspective and can afford to increase the payout (see Grullon et al. 2002). Because the higher risk may involve higher cost of capital and make the free cash flow problem worse, for free cash flow considerations, low growth firms tend to pay more dividends when they face higher risk. This theoretical result is consistent with the free cash flow hypothesis.

79.7.1 Sample Description

We collect the firm information, including total asset, sales, net income, and dividends payout, from Compustat. Stock price, stock returns, share codes, and exchange codes are retrieved from the Center for Research in Security Prices (CRSP) files. The sample period is from 1969 to 2009. Only common stocks (SHRCD = 10, 11) and firms listed on NYSE, AMEX, or NASDAQ (EXCE = 1, 2, 3, 31, 32, 33) are included in our sample. We exclude utility services (SICH = 4900–4999) and financial institutions (SICH = 6000–6999).[13] The sample includes those firm-years with at least 5 years of data available to compute average payout ratios, growth rate, return on assets, beta, total risk, size, and book-to-market ratios. The payout ratio is measured as the ratio of the dividend payout to the net income. The growth rate is the sustainable growth rate proposed by Higgins (1977). The beta coefficient and total risk are estimated by the market model over the previous 60 months. For the purpose of estimating their betas, firm-years in our sample should have at least 60 consecutive previous monthly returns. To examine the optimal payout policy, only firm-years with five consecutive dividend payouts are included in our sample.[14] Considering the fact that firm-years with no dividend payout one 1 year before (or after) might not start (or stop)

[13]We filter out those financial institutions and utility firms based on the historical Standard Industrial Code (SIC) available from COMPUSTAT. When a firm's historical SIC is unavailable for a particular year, the next available historical SIC is applied instead. When a firm's historical SIC is unavailable for a particular year and all the years after, we use the current SIC from COMPUSTAT as a substitute.

[14]To avoid creating a large difference in dividend policy, on one hand managers partially adjust firms' payout by several years to reduce the sudden impacts of the changes in dividend policy. On the other hand, they use not only 1-year firm conditions but also multiyear firm conditions to decide how much they will pay out. In examining the optimal payout policy, we use the 5-year rolling averages for all variables.

their dividend payouts in the first (fourth) quarter of the year, we exclude from our sample firm-years with no dividend payouts 1 year before or after to ensure the dividend payout policy reflects the firm's full-year condition.

Table 79.2 shows the summary statistics for 2,645 sample firms during the period from 1969 to 2009. Panel A of Table 79.2 lists the number of firm-year observations for all sample high growth firms and low growth firms, respectively. High growth firm-years are those firm-years that have 5-year average sustainable growth rates higher than their 5-year average rate of return on assets. Low growth firm-years are those firms with 5-year average sustainable growth lower than their

Table 79.2 Summary statistics of sample firm characteristics

Panel A. Sample size

	Number of firm-years				Number of firm-years		
Year	All	Growth > ROA	Growth < ROA	Year	All	Growth > ROA	Growth < ROA
1969	345	161	184	*1990*	690	522	168
1970	360	175	185	*1991*	668	511	157
1971	404	201	203	*1992*	653	494	159
1972	513	269	244	*1993*	642	460	182
1973	535	308	227	*1994*	655	479	176
1974	572	371	201	*1995*	651	483	168
1975	609	432	177	*1996*	693	530	163
1976	650	486	164	*1997*	725	582	143
1977	678	530	148	*1998*	743	620	123
1978	711	553	158	*1999*	725	612	113
1979	779	620	159	*2000*	709	607	102
1980	764	636	128	*2001*	659	569	90
1981	929	785	144	*2002*	599	503	96
1982	1,203	1,003	200	*2003*	571	475	96
1983	1,151	933	218	*2004*	525	433	92
1984	1,067	832	235	*2005*	481	391	90
1985	1,010	744	266	*2006*	510	430	80
1986	958	669	289	*2007*	542	451	91
1987	897	645	252	*2008*	579	470	109
1988	847	615	232	*2009*	610	484	126
1989	721	531	190	*All years*	28,333	21,065	6,728

Panel B of Table 79.2 Descriptive statistics of characteristics of sample

	Payout ratio	Growth rate	ROA	Beta	Total risk	Size ($MM)	M/B
All sample (N = 28,333)							
Mean	0.3793	0.1039	0.0723	1.0301	0.0106	3,072	1.7940
Median	0.3540	0.0886	0.0648	1.0251	0.0089	291	1.3539
Stdev	0.1995	0.7444	0.0389	0.4272	0.0078	14,855	1.9479
High growth firms (N = 21,065)							
Mean	0.3180	0.1233	0.0698	1.0624	0.0112	3,267	1.7951
Median	0.2996	0.1002	0.0638	1.0581	0.0095	314	1.3757

(*continued*)

Table 79.2 (continued)

Panel B of Table 79.2 Descriptive statistics of characteristics of sample							
	Payout ratio	Growth rate	ROA	Beta	Total risk	Size ($MM)	M/B
Stdev	0.1658	0.8060	0.0355	0.4352	0.0070	15,806	1.6496
Low growth firms (N = 6,728)							
Mean	0.5762	0.0413	0.0800	0.9265	0.0087	2,447	1.7904
Median	0.5542	0.0524	0.0692	0.9375	0.0071	229	1.3007
Stdev	0.1690	0.4918	0.0476	0.3822	0.0099	11,250	2.6909

This table presents the descriptive statistics for those major characteristics of our sample firms. Sample includes those firms listed on NYSE, AMEX, and NASDAQ with at least 5 years of data available to compute average payout ratios, growth rate, return on assets, beta, total risk, size, and book-to-market ratios. All financial service operations and utility companies are excluded. Panel A lists the numbers of firm-years observations for all sample firms, high growth firms, and low growth firms, respectively, during the period between year 1969 and year 2009. High growth firm-years are defined as firm-years with sustainable growth rates higher than their rates of return on assets. Low growth firm-years are defined as firm-years with sustainable growth rates lower than their rate of return on assets. Panel B lists the mean, median, and standard deviation values of the 5-year average of the payout ratio, growth rate, rate of return on assets, beta risk, total risk, size, and book-to-market ratio. The payout ratio is measured as the ratio of the dividend payout to the earnings. Growth rate is the sustainable growth rate proposed by Higgins (1977). The beta coefficient and total risk are estimated by the market model over the previous 60 months. Size is defined as market capitalization calculated by the closing price of the last trading day of June of that year times the outstanding shares at the end of June of that year

5-year average rate of return on assets. The sample size increases from 345 firms in 1969 to 1,203 firms in 1982, while declining to 610 firms by 2009. A total of 28,333 dividend-paying firm-years are included in the sample. When classifying high growth firms and low growth firms relative to their return on assets, the proportion of high growth firms increases over time. The proportion of firm-years with a growth rate higher than return on assets increases from less than 50 % during the late 1960s and early 1970s to 80 % in 2008. Panel B of Table 79.2 shows the 5-year moving averages of mean, median, and standard deviation values for the measures of payout ratio, growth rate, rate of return, beta coefficient, total risk, market capitalization, and market-to-book ratio across all firm-years in the sample. Among high growth firms, the average growth rate is 12.33 %, and the average payout ratio is 31.80 %; but for low growth firms, the average growth rate is 4.13 %, and the average payout ratio is 57.62 %. High growth firms undertake more beta risk and total risk, indicating that high growth firms undertake both more systematic risk and unsystematic risk to pursue a higher rate of return.

79.7.2 Univariate Analysis

To examine Hypothesis 1, we divide our sample into five groups by growth rate. As Table 79.3 Panel A indicates, the average (median) payout ratio is 48.64 % (48.35 %) for the lowest growth group and is 22.13 % (19.67 %) for the highest growth group. The argument of "precautionary reserves" for flexibility concerns

Table 79.3 Payout ratios partitioned by growth rate and risks

Panel A. Payout ratios partitioned by growth rate

		Low growth				High growth
		1	2	3	4	5
All sample	*Mean*	0.4864	0.3835	0.3272	0.2696	0.2213
	Median	0.4835	0.3795	0.3196	0.2556	0.1967

Panel B. Payout ratios partitioned by beta risk

		Low beta				High beta
		1	2	3	4	5
All sample	*Mean*	0.3912	0.3940	0.3963	0.3609	0.3371
	Median	0.3716	0.3765	0.3737	0.3328	0.3009
High growth firms	*Mean*	0.3173	0.3229	0.3204	0.3025	0.2886
	Median	0.3027	0.3081	0.3028	0.2874	0.2588
Low growth firms	*Mean*	0.6115	0.6064	0.6182	0.6189	0.6292
	Median	0.5864	0.5748	0.5947	0.6042	0.6090

Panel C. Payout ratios partitioned by total risk

		Low total risk				High total risk
		1	2	3	4	5
All sample	*Mean*	0.4274	0.4058	0.3618	0.3285	0.3041
	Median	0.4099	0.3808	0.3333	0.2899	0.2588
High growth firms	*Mean*	0.3499	0.3286	0.2957	0.2698	0.2531
	Median	0.3381	0.3138	0.2745	0.2431	0.2190
Low growth firms	*Mean*	0.6078	0.6202	0.6269	0.6091	0.6115
	Median	0.5732	0.5942	0.6062	0.5863	0.6025

This table presents the average and the median payout ratios in different groups partitioned by growth rate, beta risk, or total risk during the sample period from 1969 to 2009. Panel A reports the average and the median payout ratios by one-way sort on the growth rate. Panel B reports the average and the median payout ratios by independent two-way sort on the growth rate and the beta. Panel C reports the average and the median payout ratios by independent two-way sort on the growth rate and the total risk. High growth firm-years are defined as firm-years with sustainable growth rates higher than their rates of return on assets. Low growth firm-years are defined as firm-years with sustainable growth rates lower than their rate of return on assets

and Hypothesis 1 can therefore be confirmed. We further divide our sample in to five groups by beta risk and total risk indicated in Panel B and Panel C of Table 79.3. We can observe a monotonic decrease in payout ratios when the beta risk or the total risk increases, which is consistent with the findings of Rozeff (1982), Fenn and Liang (2001), Grullon et al. (2002), Aivazian et al. (2003), and Blau and Fuller (2008), but they do not consider the relationship between the payout ratio, and the risk may alter by the growth. To examine the Hypothesis 2 and Hypothesis 3, we further divide our sample by 2-way sorts on the growth rate and risks. High growth groups are firm-years with sustainable growth rates higher than their rate of return on total assets. Low growth groups are firm-years with sustainable growth rates lower than their rate of return on total assets. In Table 79.3 Panel B, among high growth firms, the average (median) payout ratio decreases from 31.73 % (30.10 %)

to 28.86 % (25.88 %) when the beta risk increases. Among low growth firms, however, the average (median) payout ratio increases from 61.15 % (58.64 %) to 62.92 % (60.90 %) when the beta risk increases. Similar to Panel B, Panel C also shows a negative relationship between the payout ratio and the total risk among high growth firms and a positive relationship between the payout ratio and the total risk among low growth firms. Above all, the static analysis results of Panel B and Panel C support Hypotheses 2 and 3 that the relationship between the payout ratio and the risk depends upon the growth rate of a firm.

79.7.3 Multivariate Analysis

To examine the relationship between the payout ratio and other financial variables, we propose fixed-effects models of the payout ratio as follows[15]:

$$
\begin{aligned}
ln\left(\frac{payout\ ratio_{i,t}}{1-\left(payout\ ratio_{i,t}\right)}\right) = {} & \alpha + \beta_1 Risk_{i,t} + \beta_2 D_{i,t}\left(g_{i,t} < c \cdot ROA_{i,t}\right) \cdot Risk_i \\
& + \beta_3 Growth_{i,t} + \beta_4 Risk_{i,t} \times Growth_{i,t} \\
& + \beta_5 \ln(Size)_{i,t} + \beta_6 ROA_{i,t} + e_{i,t}.
\end{aligned} \tag{79.36}
$$

In the regression, the dependent variable is the logistic transformation of the payout ratio. Independent variables include risk measure (beta coefficient or total risk), the interaction of dummy variable and risk measure, growth rate, the interaction of risk measure and growth rate, log of size, and rate of return on total assets.[16] Based upon the theoretical model and its implications from Sect. 79.6, we assume that c is equal to 1. The dummy variable (D_i) is equal to 1 if a firm's 5-year average growth rate is less than its 5-year average rate of return on assets and 0 otherwise. Such structure allows us to analyze the relationship between payout ratio and growth rate and the relationship between payout ratio and risk under different growth rate levels.

Thompson (2010), Peterson (2009), and Cameron et al. (2006) have pointed out that standard errors of 2-way fixed-effects estimates can be biased if 2-dimensional clustering (clustering in the cross-sectional errors and clustering in time-series errors) is not controlled for. Thompson (2010) and Boehmer et al. (2010) have empirically found that these clustering effects are not important for large samples.

[15]The dummy variable $D_{i,t}(g_{i,t} < c\ .\ ROA_{i,t})$ used in Eq. 79.36 implies that the relationship between the payout ratio and risks is nonlinear (piecewise regression). In other words, the breakpoint of the structural change is at $g_{i,t} = c\ .\ ROA_{i,t}$. Based upon our theoretical model, we assume that c is equal to 1 in our empirical work.

[16]Besides merely adding an interaction dummy as indicated in Eq. 79.36, we include an intercept dummy to take care of the individual effect of two groups. We also run regressions for high growth firms and low growth firms separately. Results from both models are qualitatively the same as those from Eq. 79.36 and also support Hypotheses 1–3.

The robust standard errors are very similar to standard errors for ordinary least square (OLS), suggesting that the fixed effects and control variables are removing most of the correlation that is present across observations. In addition, our dataset cannot be meaningfully applied to the clustering effect model[17]; therefore, statistical inferences in this paper are conducted using fixed-effects standard errors.

Table 79.4 provides the results of fixed-effects regressions for 2,645 firms during the period 1969 to 2009. *Model* (*1*) and *Model* (*2*) show that the estimated coefficients of the growth rate are −0.03 with a *t*-statistics of −4.85 and −0.03 with a *t*-statistics of −4.85, respectively. Such significantly negative coefficients confirm Hypothesis 1, which states that high growth firms will pay less in dividends for the consideration of flexibility. We also include an interaction term of risk and growth rate into *Model* (*3*) and *Model* (*4*). The results in *Model* (*3*) and *Model* (*4*) also support Hypothesis 1.

Models (*1–4*) show that the relationship between the payout ratio and the risk is significantly negative. The results are similar to the findings of Rozeff (1982), Jagannathan et al. (2000), and Grullon et al. (2002), indicating that dividend payouts are negatively correlated to firm risks; but our theoretical model shows that if firms follow their optimal dividend payout policy, the relationship between dividend payouts and firm risks depends on their growth rates relative to their rate of return on total assets as our theoretical analysis presented in Sect. 79.6. In Table 79.2, we find the number of firms with a higher growth rate with respect to their rate of return on assets greater than the number of firms with a lower growth rate with respect to their rates of return on assets. When pooling high growth firms and low growth firms together, the negative risk effect of high growth firms will dominate the positive risk effect of low growth firms due to the larger proportion of high growth firms in the observations. The results of the negative relationship between the payout ratio and the risk shown in *Models* (*1–4*) may therefore result from the greater proportion of high growth firms. Based on our subsequent analysis, the effect of growth rates on dividend payout policies can be more accurately found when firms are separated into high growth firms and low growth firms relative to their rates of return on total assets of return on assets.

To test Hypotheses 2 and 3, we introduce an interaction term of the dummy variable and the risk. In *Model* (*5*) and *Model* (*6*), the estimated coefficients of risk are −0.23 with a *t*-statistics of −13.88 and −0.22 with a *t*-statistics of −12.69, respectively. The significantly negative coefficients support the hypothesis 2 that, because of the consideration of flexibility, the payout ratio and the risk are negatively correlated for firms with a higher growth rate relative to their rate of return on assets. In addition, significant and positive coefficients of the interaction term of the dummy variable and the risk indicate that, when the risk changes, the dividend policy for low growth firms is different from that of high growth firms. By summing the coefficient of risk and the coefficient of interaction term, we can obtain coefficients

[17]Because our sample is an unbalanced panel data, the clustering computer program cannot meaningfully estimate the variance components, variance of firm $\left(\hat{V}_{firm}\right)$, variance of time $\left(\hat{V}_{time}\right)$, and heteroskedasticity-robust OLS variance $\left(\hat{V}_{white}\right)$.

Table 79.4 Fixed-effects regressions using theoretical structural change point

Model	Intercept	Beta	D*beta	Growth* beta	Total risk	D*total risk	Growth* total risk	Growth	ln(Size)	ROA	Adj-R2	F-test (p-value)
(1)	0.32***	−0.18***						−0.03***	0.04***	−13.24***	0.1945	
	(2.70)	(−9.82)						(−4.85)	(3.76)	(−63.14)		
(2)	0.45***				−17.47***			−0.03***	0.02**	−13.07***	0.1982	
	(3.76)				(−14.64)			(4.68)	(2.10)	(−62.42)		
(3)	0.31***	−0.14***		−0.29***				0.09***	0.04***	−12.96***	0.1970	
	(2.64)	(−7.76)		(−8.84)				(6.38)	(3.47)	(−61.17)		
(4)	0.45***				−17.14***		−2.90	<0.01	0.02**	−13.06***	0.1982	
	(3.77)				(−13.13)		(−0.63)	(0.01)	(2.06)	(−61.97)		
(5) $c = 1$	0.33***	−0.23***	0.71***					−0.02***	0.002	−11.54***	0.2867	<0.01
	(2.94)	(−13.88)	(57.57)					(−3.70)	(0.21)	(−57.83)		
(6) $c = 1$	0.52***				−27.81***	63.35***		−0.02***	−0.002	−11.40***	0.2753	<0.01
	(4.58)				(−24.15)	(52.23)		(−3.53)	(−0.27)	(−56.55)		
(7) $c = 1$	0.32***	−0.22***	0.70***	−0.13***				0.03***	−0.001	−11.43***	0.2871	<0.01
	(2.91)	(−12.69)	(56.96)	(−4.02)				(2.35)	(−0.10)	(−56.76)		
(8) $c = 1$	0.51***				−29.61***	63.68***	15.30***	−0.16***	−0.001	−11.47***	0.2756	<0.01
	(4.53)				(−23.44)	(52.35)	(3.46)	(−3.89)	(−0.06)	(−56.63)		

*, **, and *** indicate significance at the 10 %, 5 %, and 1 % levels, respectively

This table presents results from fixed-effects regressions of dividend payout ratios on firm characteristics for 2,645 firms during the period from 1969 to 2009. The regressions are as follows:

$$\ln\left(\frac{payout\ ratio_{i,t}}{1-\left(payout\ ratio_{i,t}\right)}\right) = \alpha + \beta_1 Risk_{i,t} + \beta_2 D\left(g_{i,t} < c \cdot ROA_{i,t}\right) \cdot Risk_i + \beta_3 Growth_{i,t} + \beta_4 Risk_{i,t} \times Growth_{i,t} + \beta_5 \ln(Size)_{i,t} + \beta_6 ROA_{i,t} + e_{i,t}$$

The dependent variable is the payout ratio with a logistic transformation. Following the theoretical model, we assume c is equal to 1. The dummy variable is equal to 1 if a firm's 5-year average growth rate is less than its 5-year average ROA and 0 otherwise. The independent variables are beta risk (total risk), dummy times beta risk (total risk), interaction between growth rate and beta risk (total risk), growth rate, log of size, and the rate of return on assets. t-statistics are presented in parentheses, and the last column presents p-values of the F-test on the null hypothesis that the restricted *Models* (*1–4*) are not different from the restricted *Models* (*5–8*) of 0.48 and 35.54 for beta and total risk, respectively, indicating the relationship between the payout ratio and the risk for low growth firms is positive. That is, when the risk increases, low growth firms will follow their optimal payout policies to increase their dividend payouts. Hypothesis 3 is thus confirmed in our empirical work. Blau and Fuller (2008) find that the flexibility hypothesis is more suitable than the free cash flow hypothesis to explain the dividend policy, but their method cannot separate the dividend policy decisions between high growth firms and low growth firms. Consequently, the model used in this paper can be regarded as a generalization of their results.

In *Model* (*7*) and *Model* (*8*) in Table 79.4, we further include the interaction term of the risk and the growth rate into the regressions. The results still support our hypotheses that the relationship between the payout ratio (Hypothesis 1) and the growth rate is negative and the relationship between the payout ratio and the risk depends on firm's growth rate with respect to its rate of return on total assets (Hypotheses 2 and 3). In addition, we find that the interaction terms of the risk and the growth rate are significantly different from zero. We also find that the adjusted R-squares for *Models* (*5–8*) are higher than those for models without dummies. F-tests also reject the null hypothesis that the regression with the interaction of the dummy variable and the risk is not different from the regression without the interaction of the dummy and the risk. We can thus conclude that the payout ratio is not linearly related to the growth rate or to the risk, and previous empirical studies on dividend policy using linear model may suffer from model misspecification.

79.7.4 Moving Estimates Process for Structural Change Model

In Eqs. 79.31 and 79.32 in Sect. 79.6, we theoretically show that the structural change breakpoint for the relationship between the payout ratio and risks is at $g_{i,t} = ROA_{i,t}$. In empirical works, we use dummy variable approach to separate the sample into a high growth ($g_{i,t} > c \,.\, ROA_{i,t}$) group and a low growth ($g_{i,t} < c \,.\, ROA_{i,t}$) group, assuming c is equal to 1, and empirically test the relationship between the payout ratio and risks for high growth firms and low growth firms. Based on the moving estimates process developed by Chow (1960), Hansen (1996, 1999, 2000), and Zeileis et al. (2002), we try to estimate the empirical breakpoint of the structural change and examine whether the empirical breakpoint is different from our theoretical breakpoint or not.

By using the moving estimates process, we can obtain an empirical estimate of $c = 0.93$ when using beta risk and $c = 0.97$ when using total risk.[18] That is, the breakpoint of the structural change is at $g_{i,t} = 0.93 \times ROA_{i,t}$ or $g_{i,t} = 0.93 \times ROA_{i,t}$. Then we redo *Model* (*5*) to *Model* (*8*) of Table 79.4 by using the empirical structural change point instead of the theoretical structural change point. In Table 79.5,

[18]Gujarati (2009) shows that this kind of problem can be regarded as piecewise regression by using moving estimates processes. Please see the Appendix 4 for the details of the moving estimates process.

Table 79.5 Fixed-effects regressions using empirical structural change point

Model	Intercept	Beta	D*beta	Growth* beta	Total risk	D*total risk	Growth* total risk	Growth	ln(Size)	ROA	Adj-R2	F-test (p-value)
***(1)* $c = 0.93$**	0.29***	−0.21***	0.77***					−0.02***	0.005	−11.63***	0.2891	<0.01
	(2.58)	(−12.57)	(58.41)					(−3.46)	(0.55)	(−58.46)		
***(2)* $c = 0.97$**	0.47***				−27.03***	66.22***		−0.02***	0.002	−11.50***	0.2758	<0.01
	(4.15)				(−23.53)	(52.43)		(−3.50)	(0.25)	(−57.13)		
***(3)* $c = 0.93$**	0.28***	−0.20***	0.77***	−0.13***				0.04***	0.004	−11.51***	0.2896	<0.01
	(2.56)	(−11.39)	(57.82)	(−4.19)				(2.60)	(0.44)	(−57.33)		
***(4)* $c = 0.97$**	0.46***				−28.67***	66.52***	13.00***	−0.15***	0.004	−11.56***	0.2761	<0.01
	(4.10)				(−22.76)	(52.53)	(3.17)	(−3.60)	(0.45)	(−57.17)		

*, **, and *** indicate significance at the 10 %, 5 %, and 1 % levels, respectively

This table presents results from fixed-effects regressions of dividend payout ratios on firm characteristics for 2,645 firms during the period from 1969 to 2009. The regressions are as follows:

$$\ln\left(\frac{payout\ ratio_{i,t}}{1-\left(payout\ ratio_{i,t}\right)}\right) = \alpha + \beta_1 Risk_{i,t} + \beta_2 D\left(g_{i,t} < c \cdot ROA_{i,t}\right) \cdot Risk_i + \beta_3 Growth_{i,t} + \beta_4 Risk_{i,t} \times Growth_{i,t} + \beta_5 \ln(Size)_{i,t} + \beta_6 ROA_{i,t} + e_{i,t}$$

The dependent variable is the payout ratio with a logistic transformation. Based upon results from the moving estimates process, c is equal to 0.93 when using beta risk and 0.97 when using total risk. The dummy variable is equal to 1 if a firm's 5-year average growth rate is less than its 5-year average ROA and 0 otherwise. The independent variables are beta risk (total risk), dummy times beta risk (total risk), interaction between growth rate and beta risk (total risk), growth rate, log of size, and the rate of return on assets. t-statistics are presented in parentheses, and the last column presents p-values of the F-test on the null hypothesis that the restricted *Models* (*1–4*) in Table 79.4 are not different from the restricted *Models* (*1–4*) in Table 79.5

we compare to the theoretical model and find that the estimates and significances using the empirical breakpoint of the structural change are almost the same as those using the theoretical breakpoint indicated in Table 79.4. Results obtained from the moving estimates process show the existence of an empirical breakpoint for the relationship between the payout ratio and risks and also confirm that the dummy variable used in Eq. 79.36 is both theoretically and empirically acceptable.

79.8 Summary and Concluding Remarks

In this paper, we extend the model developed in earlier studies and improve upon the results obtained in those studies on optimal dividend payout. By allowing partial dividend payout for the consideration of flexibility, we theoretically show the existence of the optimal dividend payout. The results obtained in this paper are different from those of M&M (1961) because, contrary to their model, the model developed here is dynamic, is under uncertainty, and allows the firm to retain some earnings for positive NPV projects, following DeAngelo and DeAngelo (2006) and Blau and Fuller (2008). The dynamic stochastic model developed and empirically tested here is comprehensive and for the very first time includes several variables simultaneously that have not been included in any of the earlier studies. More specifically, the optimization model developed here includes, among other things, a stochastic rate of return on assets, corporate growth rate, and total risk (broken down into its systematic and firm-specific components).

We further analyze the effects of different parameters on the optimal payout ratio. A sensitivity analysis and an approximation form show a negative but nonlinear relationship between the optimal dividend payout ratio and the growth rate. We also explicitly derive the theoretical relationship between the optimal payout ratio and risks. Results of the comparative static analyses cast a different light on the relationship between corporate dividend policy and risk. That the risk and dividend payout are negatively related is generally believed. On the contrary, the relationship between risk and dividend payout is positive when the growth rate is less than the company's rate of return on assets. Only when the growth rate exceeds the rate of return on assets does the relationship between dividend payout and risk turns negative. In sum, this paper shows for the first time that the relationship between growth and dividend payout is nonlinear. In addition, no research has ever shown that the relationship between risk rates and the dividend payout is nonlinear and influenced by other important variables, those of the company's rate of return on assets and growth rate.

Based upon our theoretical model, we develop three testable hypotheses and try to reconcile conflicts between the flexibility hypothesis and the free cash flow hypothesis in existing literature. The empirical research covers 40 years of US data from 1969 to 2009 for 28,333 dividend-paying firm-years. Our empirical results show that the optimal dividend payout ratio is negatively, but not linearly, related to the growth rate. In addition, the optimal dividend payout ratio is negatively (positively) related to both total risk and systematic risk when the growth rate is higher (lower) than the rate of return on assets. Such results also confirm the argument that high growth firms pay dividends due to flexibility considerations and low growth firms pay dividends due to the consideration of the free cash flow problem.

In sum, this paper develops a dynamic stochastic model deriving relationships between corporate dividend payout and several important financial variables. The relationships derived from the model are then tested by extensive empirical research. The relationships found between dividend payout and the growth rate and between dividend payout and risk run counter to those observed in earlier studies and also counter to generally held beliefs. Extensive empirical research validated the conclusions derived from the dynamic stochastic optimization model developed in this paper. For the first time we report here in the literature that these relationships are nonlinear and not solely negative as found in other studies. Thus, this paper extends the model developed in earlier studies and improves upon the results obtained in those studies on optimal dividend payout. We believe that this paper may contribute to an understanding dividend policy in the literature.

Appendix 1: Derivation of Eq. 79.19

This appendix presents a detailed derivation of the solution to the variable partial differential equation, Eq. 79.17, which is similar to Gould's (1968) Eq. 79.9 in investigation the adjustment cost. Following Gould's (1968) approach, we first derive a general solution for a standard variable partial differential equation. Then we apply this general equation to solve Eq. 79.17. The standard variable partial differential equation can be defined as

$$\dot{p}(t) + g(t)p(t) = q(t). \tag{79.37}$$

As a particular case of Eq. 79.37, the equation

$$\dot{p}(t) + g(t)p(t) = 0 \text{ or } \frac{\dot{p}(t)}{p(t)} = -g(t) \tag{79.38}$$

has a solution

$$p(t) = c \cdot \exp\left(-\int g(t)dt\right). \tag{79.39}$$

By substituting constant c with function $c(t)$, we have the potential solution to Eq. 79.37

$$p(t) = c(t) \cdot \exp\left(-\int g(t)dt\right). \tag{79.40}$$

Taking a differential with respect to t, we obtain

$$\begin{aligned}\dot{p}(t) &= \dot{c}(t)\cdot\exp\left(-\int g(t)dt\right) - c(t)\cdot\exp\left(-\int g(t)dt\right)g(t)\\ &= \dot{c}(t)\cdot\exp\left(-\int g(t)dt\right) - p(t)g(t).\end{aligned} \tag{79.41}$$

Therefore,

$$\dot{p}(t) + p(t)g(t) = \dot{c}(t)\cdot\exp\left(-\int g(t)dt\right). \tag{79.42}$$

From Eqs. 79.37 and 79.42, we have

$$\dot{c}(t)\cdot\exp\left(-\int g(t)dt\right) = q(t). \tag{79.43}$$

Equivalently,

$$\dot{c}(t) = q(t)\cdot\exp\left(\int g(t)dt\right). \tag{79.44}$$

Therefore,

$$c(t) = \int q(t)\cdot\exp\left(\int g(t)dt\right)dt. \tag{79.45}$$

Substituting Eq. 79.45 into Eq. 79.39, we have the general solution of Eq. 79.37

$$p(t) = \exp\left(-\int g(t)dt\right)\cdot\left[\int q(t)\exp\left(\int g(t)dt\right)dt\right]. \tag{79.46}$$

To solve Eq. 79.17, we will apply the above result. Let $g(t) = \delta\frac{\dot{m}(t)}{m(t)} - k$ and $q(t) = -G(t)$.

Because

$$\begin{aligned}&\exp\left(\int g(t)dt\right) = \exp\left(\int\left(\delta\frac{\dot{m}(t)}{m(t)} - k\right)dt\right) = \exp\left(\delta\int\frac{\dot{m}(t)}{m(t)}dt - kt\right)\\ &= \exp(\delta\mathrm{In}(m(t)) - kt + c) = c_1\cdot m(t)^{\delta}\exp(-kt), \text{where } c_1 > 0,\end{aligned} \tag{79.47}$$

then we have

$$P(t) = c_2\cdot m(t)^{-\delta}\exp(kt)\cdot\left[\int q(t)c_3\cdot m(t)^{\delta}\exp(-kt)dt\right], \tag{79.48}$$
$$\text{where } c_2 > 0, \text{and } c_3 > 0$$

or equivalently,

$$P(t) = c_4 \cdot \frac{e^{kt}}{m(t)^{\delta}} \cdot \left[\int -G(t) \frac{m(t)^{\delta}}{e^{kt}} dt \right], \tag{79.49}$$

where $c_4 > 0$.

Finally, we have

$$P(t) = c \frac{e^{kt}}{m(t)^{k}} \cdot \int G(t) m(t)^{\delta} e^{-kt} dt. \tag{79.50}$$

Changing from an indefinite integral to a definite integral, Eq. 79.49 can be shown as

$$p(t) = \frac{e^{kt}}{m(t)^{\delta}} \int_t^T G(s) m(s)^{\delta} e^{-ks} ds,$$

which is Eq. 79.19.

Appendix 2: Derivation of Eq. 79.21

This appendix presents a detailed derivation of Eq. 79.21. In Eq. 79.51, the initial value of the firm can be expressed as

$$p(0) = \frac{1}{m(0)^{\delta}} \int_0^T \left\{ \left(a + b\bar{I} - h\right) A(0) e^{th} m(t)^{\delta - 1} - \alpha' A(0)^2 \rho(t)^2 \sigma(t)^2 e^{2th} m(t)^{\delta - 2} \right\} e^{-kt} dt. \tag{79.51}$$

To maximize firm value, the number of shares outstanding at each point of time should be determined; therefore, the objective function can be written as follows:

$$\max_{\{m(t)\}_{t=0}^{T}} p(0). \tag{79.52}$$

Following the Euler-Lagrange condition (see Chiang 1984), we take first-order conditions on the objective function with respect to $m(t)$, where $t \in [0, T]$, and let such first-order conditions be equal to zero:

$$\begin{aligned} &\frac{1}{m(0)^{\delta}} \Big\{ (\delta - 1) \left(a + b\bar{I} - h\right) A(0) e^{th} m(t)^{\delta - 2} \\ &\quad - \alpha' A(0)^2 \rho(t)^2 \sigma(t)^2 e^{2th} m(t)^{\delta - 3} (\delta - 2) \Big\} e^{-kt} dt = 0, \end{aligned}$$

Where

$$t \in [0, T]. \tag{79.53}$$

To simplify Eq. 79.53,

$$(\delta-1)\left(a+b\bar{I}-h\right)A(0)e^{th}m(t)^{\delta-2}-\alpha' A(0)^2\rho(t)^2\sigma(t)^2e^{2th}m(t)^{\delta-3}(\delta-2)=0,$$

where

$$t\in[0,T], \tag{79.54}$$

which is Eq. 79.21.

Appendix 3: Derivation of Eqs. 79.28 and 79.29

This appendix presents a detailed derivation of both Eqs. 79.28 and 79.29. In Eq. 79.55, the optimal payout ratio with no changes in total risk or systematic risk is

$$\left[\overline{D}(t)/\overline{x}(t)\right]=\left(1-\frac{h}{a+b\bar{I}}\right)\left[\frac{-k+he^{(h-k)(T-t)}}{h-k}\right]. \tag{79.55}$$

Considering the finite growth case, if $(h-k)(T-t)<1$, then following Maclaurin expansion, the $e^{(h-k)(T-t)}$ can be expressed as

$$\begin{aligned}e^{(h-k)(T-t)}&=1+(h-k)(T-t)+\frac{(h-k)^2(T-t)^2}{2!}+\frac{(h-k)^3(T-t)^3}{3!}.\\&\approx 1+(h-k)(T-t)\end{aligned} \tag{79.56}$$

Therefore, Eq. 79.31 can be approximately written as

$$\left[\overline{D}(t)/\overline{x}(t)\right]\approx\left(1-\frac{h}{a+b\bar{I}}\right)(1+h(T-t)),$$

which is Eq. 79.38.

We further take the partial derivative of Eq. 79.33 with respect to the growth rate. Then the partial derivative of optimal payout ratio with respect to the growth rate can be approximately written as

$$\frac{\partial\left[\overline{D}(t)/\overline{x}(t)\right]}{\partial h}\approx\left(\frac{\left(a+b\bar{I}\right)(T-t)-2h(T-t)-1}{a+b\bar{I}}\right),$$

which is Eq. 79.29.

Table 79.6 Moving estimates process for testing the break point of structural change

	F-statistics			F-statistics	
c	Beta risk	Total risk	c	Beta risk	Total risk
0.20	120.38	92.78	1.01	658.58	533.26
0.25	147.77	114.94	1.02	646.43	524.02
0.30	205.83	153.89	1.03	636.99	516.94
0.35	252.56	179.52	1.04	621.57	506.64
0.40	333.30	244.85	1.05	619.94	507.41
0.45	395.33	298.05	1.06	610.91	503.46
0.50	421.89	320.42	1.07	616.34	505.71
0.55	445.75	340.36	1.08	610.81	502.08
0.60	505.14	213.02	1.09	608.96	505.16
0.65	532.57	205.06	1.10	609.88	496.72
0.70	572.82	411.85	1.15	610.61	499.96
0.75	600.57	452.04	1.20	575.99	473.31
0.80	633.15	485.64	1.25	543.25	445.85
0.85	674.48	524.66	1.30	532.27	431.02
0.90	675.53	542.20	1.35	505.57	408.10
0.91	682.40	547.18	1.40	487.91	388.90
0.92	678.08	542.58	1.45	446.97	364.58
0.93	**682.43**	549.74	1.50	408.25	326.46
0.94	679.55	547.81	1.55	383.32	304.93
0.95	679.52	546.92	1.60	345.32	266.34
0.96	675.88	**549.83**	1.65	327.34	249.22
0.97	668.42	543.73	1.70	309.13	231.42
0.98	669.35	547.23	1.75	296.67	223.43
0.99	664.33	548.16	1.80	278.69	212.60
1.00	622.77	545.64			

This shows the F-statistics of moving estimates processes. The nonstructural change regression and structural change regression are as follows:

$$\ln\left(\frac{payout\ ratio_{i,t}}{1-\left(payout\ ratio_{i,t}\right)}\right) = \alpha + \beta_1 Risk_{i,t} + \beta_2 Growth_{i,t} + \beta_3 \ln(Size)_{i,t} + \beta_4 ROA_{i,t} + e_{i,t}$$

$$\ln\left(\frac{payout\ ratio_{i,t}}{1-\left(payout\ ratio_{i,t}\right)}\right) = \alpha + \beta_1' Risk_{i,t} + \beta_2' D\left(g_{i,t} < c \cdot ROA_{i,t}\right) \cdot Risk_i + \beta_3' Growth_{i,t}$$
$$+ \beta_5' \ln(Size)_{i,t} + \beta_6' ROA_{i,t} + e_{i,t}$$

The dependent variable is the payout ratio with a logistic transformation. The breakpoint, c, is between 0.2 times and 1.8 times the rate of return on total assets. The dummy variable is equal to 1 if a firm's 5-year average growth rate is less than c times its 5-year average ROA and 0 otherwise. The independent variables are beta risk (total risk), dummy times beta risk (total risk), growth rate, log of size, and the rate of return on assets. F-statistics are under the null hypothesis that the relationship between the payout ratio and the risk does not depend on the growth rate of a firm

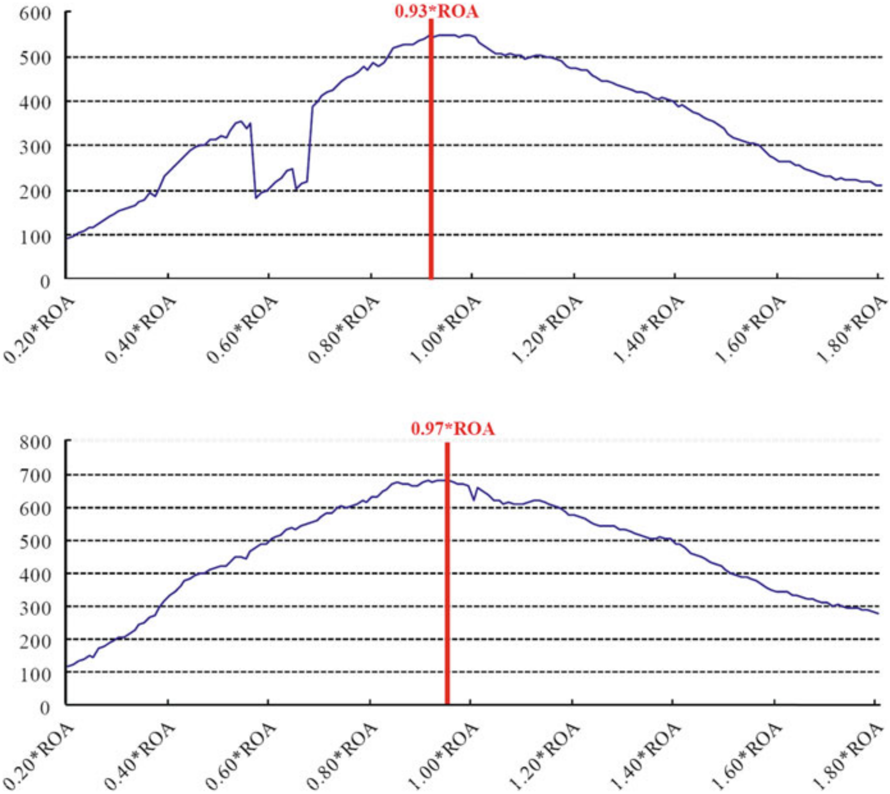


Fig. 79.2 Moving estimates process for testing the break point of structural change. The figures show the F-statistics of moving estimates processes. The nonstructural change regression and structural change regression are as follows:

$$\ln\left(\frac{payout\ ratio_{i,t}}{1-\left(payout\ ratio_{i,t}\right)}\right) = \alpha + \beta_1 Risk_{i,t} + \beta_2 Growth_{i,t} + \beta_3 \ln(Size)_{i,t} + \beta_4 ROA_{i,t} + e_{i,t}$$

$$\ln\left(\frac{payout\ ratio_{i,t}}{1-\left(payout\ ratio_{i,t}\right)}\right) = \alpha + \beta_1' Risk_{i,t} + \beta_2' D\left(g_{i,t} < c \cdot ROA_{i,t}\right) \cdot Risk_i + \beta_3' Growth_{i,t}$$
$$+ \beta_5' \ln(Size)_{i,t} + \beta_6' ROA_{i,t} + e_{i,t}.$$

The dependent variable is the payout ratio with a logistic transformation. The breakpoint, c, is between 0.2 times and 1.8 times rate of return on total assets. The dummy variable is equal to 1 if a firm's 5-year average growth rate is less than c times its 5-year average ROA and 0 otherwise. The independent variables are beta risk (total risk), dummy times beta risk (total risk), growth rate, log of size, and the rate of return on assets. F-statistics is under the null hypothesis that the relationship between the payout ratio and the risk does not depend on the growth rate of a firm. The risk used in Fig. 79. 2a is the beta, and the risk used in Fig. 79. 2b is the total risk

Appendix 4: Using Moving Estimates Process to Find the Structural Change Point in Eq. 79.36

To estimate the empirical breakpoint, we first assume the dummy variable associated with risk as $D(g_{i,t} < c\ .\ ROA_{i,t})$. We introduce a no structural change model, Eq. 79.57, and a structural change model, Eq. 79.58. In Eq. 79.58, c is a continuous constant variable, ranging from 0.2 to 1.8 ($c \in [0.2, 1.8]$). The breakpoint of structural change is at $g_{i,t} = c\ .\ ROA_{i,t}$.

$$\ln\left(\frac{payout\ ratio_{i,t}}{1-\left(payout\ ratio_{i,t}\right)}\right) = \alpha + \beta_1 Risk_{i,t} + \beta_2 Growth_{i,t} + \beta_3 \ln(\text{Size})_{i,t} + \beta_4 ROA_{i,t} + e_{i,t} \tag{79.57}$$

$$\ln\left(\frac{payout\ ratio_{i,t}}{1-\left(payout\ ratio_{i,t}\right)}\right) = \alpha + \beta_1^{'} Risk_{i,t} + \beta_2^{'} D\left(g_{i,t} < c \cdot ROA_{i,t}\right) \cdot Risk_i + \beta_3^{'} Growth_{i,t} + \beta_4^{'} \ln(Size)_{i,t} + \beta_5^{'} ROA_{i,t} + e_{i,t} \tag{79.58}$$

By using the moving estimates process, we calculate the F-statistics for all potential structural change points between $g_{i,t} = 0.2 \times ROA_{i,t}$ and $g_{i,t} = 1.8 \times ROA_{i,t}$. The F-test is under the null hypothesis that no structural change on the relationship between the payout ratio and the risk. That is, Eq. 79.57 is identical to Eq. 79.58. Finally, we can locate the breakpoint of the structural change at the point with highest value of F-statistics.

From Table 79.6, the process has a clear peak at $c = 0.93$ when using the beta risk as the independent variable and $c = 0.96$ when using the total risk as the independent variable. Figure 79.2 also graphically presents a peak F-statistics point at $c = 0.93$ (0.96) in terms of beta risk (total risk). Results from the moving estimates process indicate that a structural change on the relationship exists between the payout ratio and risks. The breakpoint of the structural change is at $g_{i,t} = 0.93 \times ROA_{i,t}$ or $g_{i,t} = 0.97 \times ROA_{i,t}$.

References

Agrawal, A., & Jayaraman, N. (1994). The dividend policies of all-equity firms: A direct test of the free cash flow theory. *Managerial & Decision Economics, 15*, 139–148.

Aivazian, V., Booth, L., & Cleary, S. (2003). Do emerging market firms follow different dividend policies from U.S. Firms? *Journal of Financial Research, 26*, 371–387.

Aoki, M. (1967). *Optimization of stochastic systems*. New York: Academic Press.

Asquith, P., & Mullins, D. W., Jr. (1983). The impact of initiating dividend payments on shareholders' wealth. *Journal of Business, 56*, 77–96.

Bar-Yosef, S., & Kolodny, R. (1976). Dividend policy and capital market theory. *Review of Economics & Statistics, 58*, 181–190.

Bellman, R. (1990). *Adaptive control process: A guided tour*. Princeton: Princeton University Press.

Benartzi, S., Michaely, R., & Thaler, R. (1997). Do changes in dividends signal the future or the past? *Journal of Finance, 52*, 1007–1034.

Bernheim, B. D., & Wantz, A. (1995). A tax-based test of the dividend signaling hypothesis. *American Economic Review, 85*, 532–551.

Bhattacharya, S. (1979). Imperfect information, dividend policy, and "the bird in the hand" fallacy. *Bell Journal of Economics, 10*, 259–270.

Blau, B. M., & Fuller, K. P. (2008). Flexibility and dividends. *Journal of Corporate Finance, 14*, 133–152.

Boehmer, E., Jones, C. M., & Zhang, X. (2010). *Shackling short sellers: The 2008 shorting ban*, Working paper.

Brennan, M. J. (1973). An approach to the valuation of uncertain income streams. *Journal of Finance, 28*, 661–674.

Brook, Y., Charlton, W. T., Jr., & Hendershott, R. J. (1998). Do firms use dividends to signal large future cash flow increases? *Financial Management, 27*, 46–57.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2006). *Robust inference with multiway Clustering*, Working paper.

Chiang, A. C. (1984). *Fundamental methods of mathematical economics* (3rd ed.). New York: McGraw-Hill.

Chow, G. C. (1960). Tests of equality between Sets of coefficients in two linear regressions. *Econometrica, 28*, 591–605.

DeAngelo, H., & DeAngelo, L. (2006). The irrelevance of the MM dividend irrelevance theorem. *Journal of Financial Economics, 79*, 293–315.

DeAngelo, H., DeAngelo, L., & Skinner, D. J. (1996). Reversal of fortune: Dividend signaling and the disappearance of sustained earnings growth. *Journal of Financial Economics, 40*, 341–371.

DeAngelo, H., DeAngelo, L., & Stulz, R. M. (2006). Dividend policy and the earned/contributed capital mix: A test of the life-cycle theory. *Journal of Financial Economics, 81*, 227–254.

Denis, D. J., Denis, D. K., & Sarin, A. (1994). The information content of dividend changes: Cash flow signaling¸overinvestment, and dividend clienteles. *Journal of Financial & Quantitative Analysis, 29*, 567–587.

Denis, D. J., & Osobov, I. (2008). Why do firms pay dividends? International evidence on the determinants of dividend policy. *Journal of Financial Economics, 89*, 62–82.

Easterbrook, F. H. (1984). Two agency-cost explanations of dividends. *American Economic Review, 74*, 650–659.

Fama, E. F., & French, K. R. (2001). Disappearing dividends: Changing firm characteristics or lower propensity to pay? *Journal of Financial Economics, 60*, 3–43.

Fenn, G. W., & Liang, N. (2001). Corporate payout policy and managerial stock incentives. *Journal of Financial Economics, 60*, 45–72.

Gabudean, R. C. (2007). *Strategic interaction and the co-determination of firms financial policies*, Working Paper.

Gordon, M. J. (1962). The savings investment and valuation of a corporation. *Review of Economics and Statistics, 44*, 37–51.

Gould, J. P. (1968). Adjustment costs in the theory of investment of the firm. *Review of Economic Studies, 35*, 47–55.

Grullon, G., Michaely, R., & Swaminathan, B. (2002). Are dividend changes a sign of firm maturity? *Journal of Business, 75*, 387–424.

Gujarati, D. N. (2009). *Basic econometrics* (5th ed.). New York: McGraw-Hill.

Hamilton, J. D. (1994). *Time series analysis* (1st ed.). Princeton: Princeton University Press.

Hansen, B. E. (1996). Inference when a nuisance parameter is not identified under the null hypothesis. *Econometrica, 64*, 413–430.

Hansen, B. E. (1999). Threshold effects in nondynamic panels: Estimation, testing, and inference. *Journal of Econometrics, 93*, 345–368.

Hansen, B. E. (2000). Sample splitting and threshold estimation. *Econometrica, 68*, 575–603.
Healy, P. M., & Palepu, K. G. (1988). Earnings information conveyed by dividend Initiations and omissions. *Journal of Financial Economics, 21*, 149–175.
Higgins, R. C. (1977). How much growth can a firm afford? *Financial Management, 6*, 7–16.
Hogg, R. V., & Craig, A. T. (2004). *Introduction to mathematical statistics* (6th ed.). Englewood Cliffs: Prentice Hall.
Howe, K. M., He, J., & Kao, G. W. (1992). One-time cash flow cash and free cash-flow theory: Share repurchases and special dividends. *Journal of Finance, 47*, 1963–1975.
Intriligator, M. D. (2002). *Mathematical optimization and economic theory*. Philadelphia: Society for Industrial and Applied Mathematics.
Jagannathan, M., Stephens, C. P., & Weisbach, M. S. (2000). Financial flexibility and the choice between dividends and stock repurchases. *Journal of Financial Economics, 57*, 355–384.
Jensen, G. R., Solberg, D. P., & Zorn, T. S. (1992). Simultaneous determination of insider ownership, debt, and dividend policies. *Journal of Financial & Quantitative Analysis, 27*, 247–263.
Jensen, M. C. (1986). Agency costs of free cash flow, corporate finance, and takeovers. *American Economic Review, 76*, 323–329.
John, K., & Williams, J. (1985). Dividends, dilution, and taxes: A signaling equilibrium. *Journal of Finance, 40*, 1053–1070.
Kalay, A., & Loewenstein, U. (1986). The informational content of the timing of dividend announcements. *Journal of Financial Economics, 16*, 373–388.
Kao, C., & Wu, C. (1994). Tests of dividend signaling using the March-Merton model: A generalized friction approach. *Journal of Business, 67*, 45–68.
Kreyszig, E. (2010). *Advanced engineering mathematics* (10th ed.). New York: Wiley.
Lang, L. H. P., & Litzenberger, R. H. (1989). Dividend announcements: Cash flow signaling vs. free cash flow hypothesis? *Journal of Financial Economics, 24*, 181–191.
Lee, C. F., & Shi, J. (2010). Application of alternative ODE in finance and econometric research. In: Lee, C. F., Lee A. C., & Lee, J. (Eds.), *Handbook of Quantitative Finance and Risk Management*. Springer, 1293-1300.
Lerner, E. M., & Carleton, W. T. (1966). Financing decisions of the firm. *Journal of Finance, 21*, 202–214.
Lie, E. (2000). Excess funds and agency problems: An empirical study of incremental cash disbursements. *Review of Financial Studies, 13*, 219–248.
Lie, E. (2005). Operating performance following open market share repurchase announcements. *Journal of Accounting and Economics, 39*, 411–436.
Lintner, J. (1962). Dividends, earnings, leverage, stock prices and the supply of capital to corporations. *Review of Economics and Statistics, 44*, 243–269.
Lintner, J. (1963). The cost of capital and optimal financing of corporate growth. *Journal of Finance, 18*, 292–310.
Lintner, J. (1964). Optimal dividends and corporate growth under uncertainty. *Quarterly Journal of Economics, 78*, 49–95.
Lintner, J. (1965). Security prices, risk, and maximal gains from diversification. *Journal of Finance, 20*, 587–615.
Miller, M. H., & Modigliani, F. (1961). Dividend policy, growth, and the valuation of shares. *Journal of Business, 34*, 411–433.
Miller, M. H., & Rock, K. (1985). Dividend policy under asymmetric information. *Journal of Finance, 40*, 1031–1051.
Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica, 34*, 768–783.
Nissim, D., & Ziv, A. (2001). Dividend changes and future profitability. *Journal of Finance, 56*, 2111–2133.
Petersen, M. A. (2009). Estimating standard errors in finance panel data sets: Comparing approaches. *Review of Financial Studies, 22*, 435–480.
Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica, 32*, 122–136.

Rozeff, M. S. (1982). Growth, beta and agency costs as determinants of dividend payout ratios. *Journal of Financial Research, 5*, 249–259.
Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance, 19*, 425–442.
Thompson, S. B. (2010). *Simple formulas for standard errors that cluster by both firm and time*, Working paper.
Wallingford, B. A., II. (1972a). An inter-temporal approach to the optimization of dividend policy with predetermined investments. *Journal of Finance, 27*, 627–635.
Wallingford, B. A., II. (1972b). A correction to "An inter-temporal approach to the optimization of dividend policy with predetermined investments". *Journal of Finance, 27*, 944–945.
Yoon, P. S., & Starks, L. T. (1995). Signaling, investment opportunities, and dividend announcements. *Review of Financial Studies, 8*, 995–1018.
Zeileis, A., Leisc, F., Hornik, K., & Kleiber, C. (2002). strucchange: An R package for testing for structural change in linear regression models. *Journal of Statistical Software, 7*, 1–38.

Modeling Asset Returns with Skewness, Kurtosis, and Outliers

80

Thomas C. Chiang and Jiandong Li

Contents

Abstract

This chapter uses an exponential generalized beta distribution of the second kind (EGB2) to model the returns on 30 Dow Jones industrial stocks. The model accounts for stock return characteristics, including fat tails, peakedness (leptokurtosis), skewness, clustered conditional variance, and leverage effect.

T.C. Chiang (✉)
Department of Finance, Drexel University, Philadelphia, PA, USA
e-mail: chiangtc@drexel.edu

J. Li
Chinese Academy of Finance and Development (CAFD) and Central University of Finance and Economics (CUFE), Beijing, China
e-mail: jiandongli@cufe.edu.cn

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_80,

The evidence suggests that the error assumption based on the EGB2 distribution is capable of taking care of skewness, kurtosis, and peakedness and therefore is also capable of making good predictions on extreme values. The goodness-of-fit statistic provides supporting evidence in favor of EGB2 distribution in modeling stock returns. This chapter also finds evidence that the leverage effect is diminished when higher moments are considered.

The EGB2 distribution used in this chapter is a four-parameter distribution. It has a closed-form density function and its higher-order moments are finite and explicitly expressed by its parameters. The EGB2 distribution nests many widely used distributions such as normal distribution, log-normal distribution, Weibull distribution, and standard logistic distribution.

Keywords

Expected stock return • Higher moments • EGB2 distribution • Risk management • Volatility • Conditional skewness • Risk premium

80.1 Introduction

Focusing on economic rationales, financial economists have identified a set of fundamental variables to predict stock returns over time, including market risk, change in interest rate, inflation rate, real activities, default risk, term premium, dividend yields, and earning yields, among other variables. In the cross-sectional analysis, Fama and French (1996) further emphasize size factor (SMB) and value factor (HML). Depending on the frequency of the data being studied, the Monday effect or the January effect is usually added to the model to highlight calendar anomalies. The empirical evidence of statistical significance to justify these variables is rather diverse. The mixed results have been attributed to variations in sample size, frequency, country, market, and/or model specification. As Avramov (2002) argues, the lack of consensus in choosing the "correct" variables may stem from model uncertainty, since the equilibrium asset pricing theories are not explicit about which variables should be included in the predictive regression.

To deal with this uncertainty, researchers occasionally resort to a missing variable. It becomes more apparent as GARCH-type models show that financial data demonstrate some sort of volatility clustering phenomenon. To incorporate the conditional variance into the mean equation is definitely helpful in tying stock returns to volatility (See French et al. 1987; Akgiray 1989; Baillie and DeGennaro 1990; and Bollerslev et al. 1992, among others). However, the GARCH-type specification based on a normal distribution is unsatisfactory for use with data that entail extreme values. Recent financial market developments show that significant daily loss occurs more frequently, and the volatility cannot reasonably be predicted from a normal distribution. The popularity of using a normal distribution assumption lies in the fact that the statistical analysis of stock returns can be simplified, allowing the analyst to focus on the first two moments. This simplification, however, misses the information contained in higher order moments.

To account for higher-order moments is important in modeling stock return series for the following reasons. First, from an econometric point of view, Hansen (1994) notes that empirical specifications of asset pricing models are incomplete unless the full conditional model is specified. Estimation and forecasting accuracy depends on the full specification of the distribution moments. Many authors have found that higher-order moments (and co-moments) can serve as explanatory variables for modeling stock returns (Harvey and Siddique 2000; Patton 2004; Ranaldo and Favre 2005; Bali et al. 2008; Boyer et al. 2010). The exclusion of higher-order moments information in the asset return model is bound to result in missing variable and misspecification problems.

Second, from the perspective of empirical finance studies, higher-order moments have particular economic meanings. Johnson and Schill (2006) suggest that Fama-French factors (SMB and HML) can be viewed as proxies for higher-order co-skewness and co-kurtosis. They show that Fama-French loadings generally become insignificant when higher-order systematic co-moments are included in cross-sectional regressions of portfolio returns.

Third, for portfolio management, higher-order moments are considered additional risk instruments in constructing the "new" portfolio theory, as argued by Jurczenko and Maillet (2002) and papers cited there. Further, the underlying theory of stochastic dominance (Vinod 2004) suggests that portfolio selection is determined not only by the conditional mean and variance but also by the skewness and kurtosis. The evidence provided by Harvey et al. (2010) and Cvitanic et al. (2008) substantiates the validity of the new portfolio theory. Moreover, in their recent studies, Andersen and Sornette (2001) and Malevergne and Sornette (2005) find that by incorporating higher-order moments risk, it is possible to increase the expected return on the portfolio while lowering its risks. Similarly, Tang (1998) finds that diversification reduces standard deviation but worsens negative skewness and fat tails in his study of the Hong Kong stock market. The evidence thus points to the fact that pricing risk based exclusively on the second moment may be very misleading. In light of this consideration, existing risk management techniques ought to be revised as well.

The significance of higher-order moments has been revealed in a series of dramatic market events, such as the market crash in 1987, the Asian crisis in 1997, the financial collapses of LTCM, the bust of internet bubble, and the subprime loan crisis in 2007. To address excess risk, both financial institutions and regulatory agencies demand risk management techniques to deal with occurrences of extreme values. Although Value at Risk (VaR) has been invented to predict a portfolio's maximum loss over a target horizon in a given confidence interval, the standard VaR models based on normal distribution often underestimate the potential risk.

Three approaches have been developed in the literature to deal with higher-order moments. The first approach is to treat higher-order moments as explanatory variables in the stock return equation. The four-moment CAPM by Jurczenko and Maillet (2002) and Ranaldo and Favre (2005) are the examples. The difficulty of this approach lies in how to generate the explanatory variables. Generating explanatory variables usually relies on higher frequency data or a rolling sample method.

The second approach is to apply the GARCH approach to higher conditional moments. Harvey and Siddique (1999) consider the conditional skewness, Brooks et al. (2005) tackle the autoregressive conditional kurtosis, and Conrad et al. (2009) find that individual securities' volatility, skewness, and kurtosis are strongly related to subsequent returns. Although these studies are capable of extracting information from the higher-order moments and use them to explain the conditional mean, they have not completely resolved the fundamental issue that the dependent variable frequently violates the assumption of a normal distribution.[1] This leads to the third approach: applying non-Gaussian distributions to model stock returns so that higher-order moments are naturally incorporated. This chapter falls into the third category.

The knowledge that stock returns are not following the Gaussian distribution dates back to the papers by Mandelbrot (1963) and Fama (1965). Subsequent research includes Officer (1972), Clark (1973), McCulloch (1985), Bollerslev (1987), Nelson (1991), Hansen (1994), Liu and Brorsen (1995), and Mittnik et al. (1999) among others. These studies propose the t distribution, skewed t distribution, general error distribution (GED, also known as exponential power distribution), and α-stable Levy distributions. Briefly speaking, the t distribution is symmetric so that it inherently fails to describe the issue of skewness. The GED is not flexible enough to allow for larger innovations. The stable distribution has theoretical appeal because of the generalized central limit theorem; however, its moments are not defined for an order greater than α. In particular, the variance is not defined except for one special case, normal distribution; the skewness and kurtosis are always not defined. Finally, the skewed t distribution used in Hansen (1994) is far from being parsimonious, and it is hard to interpret its parameters because transformations are imposed.

Recognizing the weakness of the above distributions, it is necessary to have a model that encompasses the features of asymmetry, a high peak, and fat tails. We find that an exponential generalized beta distribution of the second kind (EGB2) [2] is able to meet the diverse criteria, which forms the research foundation of this chapter.

Results emerging from this chapter show that the EGB2 distribution works very well in dealing with high-order moments of individual stock returns. The evidence indicates that AR(1)-GJR-GARCH(1,1) model based on the EGB2 distribution provides a unique specification in handling the stylized facts of stock return behaviors: autocorrelation, conditional heteroskedasticity, leverage effect, skewness, excess kurtosis, and peakedness.

---

[1]Both Harvey and Siddique (1999) and Brooks et al. (2005) use a t distribution. As shown in this paper, a t distribution has heavy tails but is not a good fit for stock return data with regard to peakedness.

[2]There are other names for the EGB2 distribution in other nonfinancial fields or in non-American journals; for example, generalized logistic distribution in Wu et al. (2000), z-distribution in Barndorff-Nielsen et al. (1982), the Burr-type distribution in actuarial science in Hogg and Klugman (1983), and four-parameter kappa distribution in geology in Hosking (1994).

This study contributes to the literature in the following aspects. We find that using EGB2 distribution is superior to models based on a normal distribution and *t* distribution in handling skewness and kurtosis, as is evident by the goodness-of-fit statistics. Second, the prevalence of the risk management method Value at Risk (VaR) can be handled and updated via the EGB2 distribution. It informs investors that omitting higher moments "...will lead to a systematic underestimate of the true riskiness of a portfolio, where risk is measured as the likelihood of achieving a loss greater than some threshold" (Brooks et al. 2005, p. 400). Third, this chapter systematically examines all 30 stocks in the Dow Jones industrial index. The individual stocks cover a broad range of assets and reveal a variety of fat tail characteristics. The model encompasses a rich spectrum of asset features that help in guiding portfolio decisions. Fourth, we find that the asymmetric effect (leverage effect) has been diminished when the EGB2 distribution is applied. It implies that the so-called leverage effect is, at least, partially attributable to the model's misspecification due to the imposition of a normal distribution of return series.

The remainder of the chapter is organized as follows. Section 80.2 describes the methodology of the EGB2-GARCH model. Section 80.3 discusses the data. Section 80.4 presents the empirical results on the stock returns by applying different distributions; Sect. 80.5 reports the goodness-of-fit tests; Sect. 80.6 contains the probability evaluation using the EGB2 distribution. Section 80.7 contains conclusions.

80.2 The GARCH-Type Model Based on the EGB2 Distribution

80.2.1 General Specification

The AR(1)-GARCH(1,1)-GJR-EGB2 stock return model can be represented by a system given below:

$$R_t = \phi_0 + \phi_1 R_{m,t} + \phi_2 R_{t-1} + \delta D_{87} + \varepsilon_{it} \tag{80.1a}$$

$$\varepsilon_t = \sqrt{h_t} u_t \tag{80.1b}$$

$$h_t = w + \alpha \varepsilon_{t-1}^2 + \beta h_{t-1} + \gamma I(\varepsilon_{t-1} < 0)\varepsilon_{t-1}^2 \tag{80.1c}$$

$$\varepsilon_t | \Im_{t-1} \sim D(0, h_t, z) \tag{80.1d}$$

Equation 80.1a is the mean equation, where R_t is an individual stock's excess return (stock return minus the risk-free rate) at time t; ε_t is an error term. The inclusion of an AR(1) term in the mean equation accounts for the autocorrelation arising from nonsynchronous trading or slow price adjustments

(Lo and MacKinlay 1990; Amihud and Mendelson 1987).[3] The market's equity premium (stock market return minus the risk-free rate), R_{mt}, at time t is included in the equation to capture the market risk as suggested by the CAPM. The dummy variable, D_{87}, takes a value of unity in the week of October 19, 1987, and 0 otherwise. The series, u_t, in Eq. 80.1b is a standardized error by conditional variance.

The conditional variance, h_t, is assumed to follow GARCH(1,1); w,α, and $\beta > 0$ to ensure a strictly positive conditional variance; and I is an indicative function that takes a value of 1 only when the error term has a negative value. The γ is used to capture the asymmetric effect of the extraordinary shock to the variance: bad news usually results in a bigger effect than good news does. In this study, we adopt the asymmetric GARCH approach suggested by Glosten et al. (1993) for its simplicity and effectiveness. The distribution of ε_t is assumed to be a general specification conditional on the distribution captured by the parameter z. For the normal distribution, the error follows that $\varepsilon_t|\Im_{t-1} \sim N(0, h_t)$. As a variant of a normal distribution, in this chapter, we consider two alternatives: t- and EGB2 distributions.

80.2.2 Modeling Financial Time Series Based on Non-Normal Distributions

Student's t distribution is well known for its capacity to capture the fat tail phenomenon. Bollerslev (1987), Bollerslev et al. (1994), and Hueng and McDonald (2005) incorporated t distribution into the GARCH model specification. The pdf of a normalized Student's t distribution takes the form of

$$t(x;\delta,\sigma,v) = \frac{\Gamma\left(\frac{v+1}{2}\right)}{\Gamma\left(\frac{v}{2}\right)\sigma\sqrt{\pi(v-2)}}\left[1+\frac{1}{v-2}\left(\frac{x-\delta}{\sigma}\right)^2\right]^{-\frac{v+1}{2}} \tag{80.2}$$

where x is a random variable, v is the degree of freedom of the t distribution ($v > 2$), and Γ is the gamma function. The excess kurtosis coefficient of the t distribution is given by $\frac{6}{v-4}$ for $v > 4$. In light of system (80.1a–80.1d), the only change is in the error distribution, which is given by $\varepsilon_t|\Im_{t-1} \sim t(0, h_t, v)$. From this perspective,

[3]Dependent on the significance test of the AR(1) coefficient in the AR(1)-GARCH(1,1) model, the AR(1) term is then dropped for some stocks. The following stocks do not have an AR(1) variable: MSFT, HON, DD, GM, IBM, MO, CAT, BA, PFE, AA, DIS, MCD, JPM, and INTC. Stock PG, which is the only one that shows Q(30) any significance, adds an AR(4) variable to ensure that the autocorrelation is removed. The rest of this paper follows this pattern. Recent literature suggests that the sign of the AR(1) coefficient, ϕ_2, can be used to detect feedback trading behavior (Sentana and Wadhwani 1992; Antoniou et al. 2005). Our results show that the coefficient of AR(1) is negative.

both the coefficients and the degree of freedom of the t distribution are estimated simultaneously by maximizing the following log-likelihood function:

$$\log L = T\left[\log\left(\Gamma\left(\frac{v+1}{2}\right) - \log\left(\Gamma\left(\frac{v}{2}\right) - 0.5\log(\pi(v-2))\right]\right.\right. \\ - 0.5\sum\left[\log(h_t) + (v+1)\log\left(1 + \frac{\varepsilon_t{}^2}{h_t(v-2)}\right)\right] \tag{80.3}$$

Although the t distribution is good at modeling fat tails for time data, its shortcoming is the built-in symmetrical nature. The distribution, however, is unable to take care of the skewness characteristic present in the financial time series. Thus, we turn to an exponential generalized beta distribution of the second kind (EGB2) developed by McDonald (1984, 1991) and McDonald and Xu (1995).

EGB2 is attractive because of its simplicity and ease in estimating the parameters.[4] There is a closed-form density function for the EGB2 distribution; its higher-order moments are finite and explicitly expressed by its parameters. Moreover, it is flexible and able to accommodate a wider range of data characteristics, such as thick tails and skewness, than the more commonly used normal and log-normal distributions.

The EGB2 distribution has the probability density function (pdf) given by

$$EGB2(x;\delta,\sigma,p,q) = \frac{\left[e^{\left(\frac{x-\delta}{\sigma}\right)}\right]^p}{|\sigma|B(p,q)\left[1+e^{\left(\frac{x-\delta}{\sigma}\right)}\right]^{p+q}} \tag{80.4}$$

where x is a random variable; δ is a location parameter that affects the mean of the distribution; σ reflects the scale of the density function; p and q ($p > 0$ and $q > 0$) are shape parameters that together determine the skewness and kurtosis of the distribution of the excess return series; $B(p, q)$ is the beta function.[5] As suggested by McDonald

[4]It is not our intention to exhaust all the non-Gaussian models in our study, which is infeasible. Rather, our strategy is to adopt a distribution that is rich enough to accommodate the features of financial data. To our knowledge, there are different types of flexible parametric distributions parallel to the EGB2 distribution to model both third and fourth moments in the literature. One family of such distributions is a skewed generalized t distribution (SGT) (Theodossiou 1998; Hueng and Brooks 2003). Special cases of SGT include generalized t distribution (McDonald and Newey 1988), skewed t distribution (Hansen 1994), and skewed generalized error distribution (SGED) (Nelson 1991). The skewness and excess kurtosis of SGT are in the range $(-\infty, \infty)$ and $(1.8, \infty)$, respectively. Another family is inverse hyperbolic sin distribution (IHS) (Johnson 1949 and Johnson et al. 1994). The skewness and excess kurtosis of IHS is in the range $(3, \infty)$ and $(-\infty, \infty)$. EGB2 has less coverage for skewness and excess kurtosis than SGT and IHS. However, it covers many skewness-kurtosis combinations encountered in practice and performs "impressively" in estimating the slope coefficient in a simulation (Theodossiou et al. 2007). Other families of flexible distributions are also available in the literature. But there isn't any comparison with the EGB2 distribution.

[5]It should be noted that beta function here has nothing to do with the stock's beta.

(1991), the EGB2 is suitable for coefficient of skewness values between −2 and 2 and coefficient of excess kurtosis values up to 6. The distribution is capable of accommodating fat-tailed and skewed error distributions pertinent to stock return modeling.[6]

For the standardized EGB2 distribution with shape parameters p and q, the univariate GARCH-EGB2 log-likelihood function is[7]

$$\log L = T\left[\log\left(\sqrt{\Omega}\right) - \log(B(p,q)) + p\Delta\right] + \sum\left[p\left(\frac{\sqrt{\Omega}\varepsilon_t}{\sqrt{h_t}}\right) - 0.5\log(h_t)\right.$$
$$\left. -(p+q)\log\left(1+\exp\left(\frac{\sqrt{\Omega}\varepsilon_t}{\sqrt{h_t}}+\Delta\right)\right)\right] \tag{80.5}$$

where $\Delta = \psi(p) - \psi(q)$, $\Omega = \psi'(p) + \psi'(q)$, and ψ and ψ' represent digamma and trigamma functions, respectively.[8] The BFGS algorithm is used in RATS to conduct a maximum likelihood estimation. The skewness and excess kurtosis for the EGB2 distribution are given respectively by

$$\text{Skewness} = g(p,q) = \frac{\psi''(p) - \psi''(q)}{\left(\psi'(p) + \psi'(q)\right)^{1.5}} \tag{80.6}$$

$$\text{Kurtosis} = h(p,q) = \frac{\psi'''(p) + \psi'''(q)}{\left(\psi'(p) + \psi'(q)\right)^{2}} \tag{80.7}$$

and ψ'' and ψ''' represent tetragamma and pentagamma functions.

Since the skewness and kurtosis coefficients are based on parameters p and q, the standard deviation of skewness and kurtosis coefficients can be drawn by using a standard delta method (see the appendix for details). By using these measures, we can judge if the EGB2 distribution correctly handles skewness and kurtosis.

80.3 Data and Summary Statistics

When asset returns are analyzed, movements in the Dow Jones Industrial Average (DJIA) are often considered one of the most important pieces of news that indicate

[6]Many distributions are nested in the EGB2 distribution. Wang et al. (2001) show that the EGB2 distribution is very powerful in modeling exchange rates that have fat tails and leptokurtosis features. The EGB2 converges to normal distribution as $p = q$ approaches infinity, to log-normal distribution when only p approaches infinity, to the Weibull distribution when $p = 1$ and q approaches infinity, and to the standard logistic distribution when $p = q = 1$. It is symmetric (called a Gumbel distribution) for $p = q$. The EGB2 is positively (negatively) skewed as $p > q$ ($p < q$) for $\sigma > 0$.

[7]This can be obtained in the appendix of Wang et al. (2001).

[8]The digamma function is the logarithmic derivative of the gamma function; the trigamma function is the derivative of the digamma function.

the health of financial markets and investment performance. This chapter uses the DJIA's 30 stocks as the sample, which represents a group of well-established and diverse companies. The sample covers the period from October 29, 1986, through December 31, 2005. One of the reasons for using this period is its completeness, so we can employ and assess information on all 30 stocks in the sample period.[9] This time period also captures the most vigorous recent stock market advancements while covering several major market crashes and financial crises.

Following the conventional approach, returns from the Standard & Poor's 500 (SP 500) index are used to measure the market return. Both daily returns on the S&P 500 index and data on the 30 Dow Jones firms are compounding and including dividend payments. These data are taken from the CRSP database. The short-term interest rate is measured by the 3-month Treasury bill rate, which is taken from the Federal Reserve's website.[10] The daily risk-free rate is measured by using the annual rate divided by 360. The excess stock return is the difference between actual stock returns and the short-term interest rate.

Weekly data are used in order to be consistent with industrial practice. For example, Value Line, Bloomberg, and Baseline all use weekly data on stocks to calculate the stocks' beta. Daily stock returns are seldom used in industry. It also helps to smooth out the volatility for single date outliers. An additional advantage of using weekly observations is that some calendar effects, such as the Monday effect, can be avoided. The excess returns are measured on a weekly basis. Table 80.1 reports summarized statistics for the weekly excess returns.

Looking at Table 80.1, we find that six stocks have a positive value for the skewness coefficient and two are significant at the 1 % level, while the remaining 24 stocks show negative values and 13 of them are significant at the 1 % level.[11] A negative skewness coefficient means that there are more negative extreme values

[9]Stock C (CitiGroup) in Table 80.1 began its trading data on Oct 29, 1986. Within this period, only one stock has one missing value. Stock MO (Philips Morris Co.) was not traded on May 25, 1994, because of "pending news which could affect the stock price." Philip Morris' board was meeting to announce whether the company would split its food and tobacco units on May 25, 1994. In this sample period, the most striking event is the market crash on October 19, 1987. This paper considers the 1987 market crash as an outlier in the later part. The week of the 9–11 terrorist attacks has only 1 day of trading information and is incorporated into the next week.

[10]http://www.federalreserve.gov/releases/H15/data.htm#top, Treasury bill secondary market rates (serial: tbsm3m) are the averages of the bid rates quoted on a bank discount basis by a sample of primary dealers who report to the Federal Reserve Bank of New York. The rates reported are based on quotes at the official close of the US government securities market for each business day.

[11]The sign of the skewness coefficient is related to data frequency. The skewness of the weekly return has nothing to do with the skewness of the daily return. For example, the stock HON (index = 2) shows significant positive skewness in its daily return but significant negative skewness in its weekly return.

[12]The skewness coefficient is the relation between the second order moment and the third order moment. It is calculated by $\frac{T}{(T-1)(T-2)\sigma^3}\sum(x_i-\mu)^3$ where μ is the mean of the sample. The literature about the positive and negative values of the distribution skewness is confusing. The skewness in our study is based on the distribution's moments (Kenney and Keeping 1962).

Table 80.1 Descriptive statistics for weekly excess stock returns: 1986–2005

Index	Company name	Ticker	Nobs	Mean	Variance	Skewness	Kurtosis	Peakedness	Jarque-Bera	Q(30)	$Q^2(30)$
1	Microsoft Corp.	MSFT	999	0.00623	0.00243	0.1054	1.7434	1.1322	128.3609	40.0341	210.3496
						[1.36]	[11.25]***		0***	0.1	0***
2	Honeywell International Inc.	HON	999	0.00234	0.00189	−0.7458	10.4318	1.0107	4,622.3848	66.6923	138.3476
						[−9.62]***	[67.30]***		0***	0***	0***
3	Coca-Cola Co.	KO	999	0.00264	0.00122	−0.2049	1.5426	1.0786	106.0431	35.541	222.6358
						[−2.64]***	[9.95]***		0***	0.22	0***
4	E.I. DuPont de Nemours & Co.	DD	999	0.00186	0.00141	−0.1613	1.4728	1.0986	94.6198	51.4363	278.872
						[−2.08]**	[9.50]***		0***	0.01***	0***
5	Exxon Mobil Corp.	XOM	999	0.00253	7.84E-04	−0.1485	1.3036	1.1647	74.4108	119.402	208.3203
						[−1.92]*	[8.41]***		0***	0***	0***
6	General Electric Co.	GE	999	0.00297	0.00119	−0.1052	3.2818	1.1407	450.1587	50.5722	197.9357
						[−1.36]	[21.17]***		0***	0.01**	0***
7	General Motors Corp.	GM	999	8.57E-04	0.00176	−0.1895	2.192	1.1856	205.9881	35.2411	26.7299
						[−2.45]**	[14.14]***		0***	0.23	0.64
8	International Business Machines Corp.	IBM	999	0.00168	0.00159	0.0399	2.3767	1.0927	235.388	40.55	198.6334
						[0.51]	[15.33]***		0***	0.09*	0***
9	Altria Group Inc.	MO	999	0.00361	0.00157	−0.3389	3.9432	1.0472	666.3313	37.7432	79.336
						[−4.37]***	[25.44]***		0***	0.16	0***
10	United Technologies Corp.	UTX	999	0.00296	0.00147	−1.4454	12.8243	1.0301	7,193.6328	76.4998	43.2729
						[−18.6]***	[82.74]***		0***	0***	0.06*
11	Procter & Gamble Co.	PG	999	0.00304	0.00125	−2.09	24.1258	1.055	24,955.353	99.2438	45.9705

						[−26.9]***	[155.6]***		0***	0***	0.03**
12	Caterpillar Inc.	CAT	999	0.00319	0.00187	0.1223	3.4189	1.1005	489.0329	49.6413	63.4811
						[1.58]	[22.06]***		0***	0.01**	0***
13	Boeing Co.	BA	999	0.00243	0.00175	−0.9363	8.9094	1.0939	3,450.0172	26.1619	85.6471
						[−12.0]***	[57.48]***		0***	0.67	0***
14	Pfizer Inc.	PFE	999	0.00292	0.00151	−0.2471	1.6742	1.2616	126.8375	46.246	95.4955
						[−3.19]***	[10.80]***		0***	0.03**	0***
15	Johnson & Johnson	JNJ	999	0.00299	0.00109	−0.0269	2.3118	1.18	222.5903	51.99	123.6244
						[−0.35]	[14.92]***		0***	0.01**	0***
16	3 M Co.	MMM	999	0.00229	9.87E-04	−0.0023	2.3585	1.0821	231.5415	43.9957	202.3982
						[−0.03]	[15.22]***		0***	0.05**	0***
17	Merck & Co. Inc.	MRK	999	0.00238	0.00145	−0.3045	2.4453	1.1817	264.3375	34.283	58.3937
						[−3.93]***	[15.78]***		0***	0.27	0***
18	Alcoa Inc.	AA	999	0.00279	0.00208	−0.4607	6.0484	1.1334	1,558.134	56.6534	73.5826
						[−5.95]***	[39.02]***		0***	0***	0***
19	Walt Disney Co.	DIS	999	0.00246	0.00166	−0.2527	2.9907	1.2043	382.9379	38.326	77.1463
						[−3.26]***	[19.30]***		0***	0.14	0***
20	Hewlett-Packard Co.	HPQ	999	0.00322	0.00285	−0.1841	2.1588	1.1304	199.6392	51.4856	186.6745
						[−2.38]**	[13.93]***		0***	0.01***	0***
21	McDonald's Corp.	MCD	999	0.0022	0.00123	−0.0701	1.085	1.2158	49.8201	36.5428	140.4038
						[−0.90]	[7.00]***		0***	0.19	0***
22	JPMorgan Chase & Co.	JPM	999	0.00252	0.00239	−0.0532	1.9238	1.079	154.5247	42.948	353.6849
						[−0.69]	[12.41]***		0***	0.06*	0***
23	Wal-Mart Stores Inc.	WMT	999	0.00324	0.00161	0.0414	1.3849	1.1689	80.1245	47.2936	333.5604

(*continued*)

Table 80.1 (continued)

Index	Company name	Ticker	Nobs	Mean	Variance	Skewness	Kurtosis	Peakedness	Jarque-Bera	Q(30)	$Q^2(30)$
						[0.53]	[8.94]***		0***	0.02**	0***
24	American Express Co.	AXP	999	0.00284	0.0018	−0.2135	2.7292	1.2052	317.6268	45.6515	148.3075
						[−2.75]***	[17.61]***		0***	0.03**	0***
25	Intel Corp.	INTC	999	0.0055	0.00342	−0.632	3.8506	1.2007	683.6673	29.5866	87.1171
						[−8.15]***	[24.84]***		0***	0.49	0***
26	Verizon Communications Inc.	VZ	999	0.00152	0.00116	0.1909	1.8996	1.2103	156.263	55.9734	254.7223
						[2.46]**	[12.26]***		0***	0***	0***
27	AT&T	T	999	0.00196	0.00133	0.2295	3.1149	1.1904	412.6303	51.7092	300.738
						[2.96]***	[20.10]***		0***	0.01***	0***
28	Home Depot Inc.	HD	999	0.00539	0.0023	−0.4238	4.5083	1.1297	875.9315	36.4251	165.8335
						[−5.47]***	[29.09]***		0***	0.19	0***
29	American International Group Inc.	AIG	999	0.00278	0.00139	0.3886	2.8457	1.2107	362.2258	44.6777	112.3984
						[5.01]***	[18.36]***		0***	0.04**	0***
30	Citigroup Inc.	C	999	0.00422	0.00208	0.1921	2.9292	1.1439	363.2863	42.9245	93.5565
						[2.48]**	[18.90]***		0***	0.06*	0***
31	S&P 500 (market)		999	0.00133	4.62E-04	−0.5292	2.8903	1.173	394.3576	58.9947	238.3201
						[−6.83]***	[18.65]***		0***	0***	0***

The 30 stocks are sorted by permanent CRSP number. nobs are the number of observations. The last row market is measured by the S&P 500. Numbers below coefficients are t-values (with bracket). Numbers below tests are p-values. ***Indicates 1 % significance, **5 %, and *10 %. The standard deviations of skewness and excess kurtosis coefficients are given approximately by $(6/T)^{0.5}$ and $(24/T)^{0.5}$, respectively. The peakedness is measured by $f_{0.75}$-$f_{0.25}$, the distance between the values of standardized variables at which the cumulative distribution function equals 0.75 and the value at which the cumulative distribution function equals 0.25. The reference value of the standard normal distribution is 1.35. A number of peakedness less than 1.35 means there is a high peak in the probability density function. The normality test is conducted using a Jarque-Bera statistic. The independence test is conducted using a Ljung-Box Q test up to the order of 30. The Q^2 test up to the order of 30 is to show volatility clustering than positive extreme values in the sample period.[12] With respect to the excess kurtosis (kurtosis coefficient minus 3), all of the estimated values are statistically significant at the 1 % level, suggesting a serious fat-tailed problem. The range of the excess kurtosis coefficient is between 1.08 and 24.13. By checking the range of peakedness measured by the inter-quartile range (i.e., 0.75 fractile minus 0.25 fractile), we found that it lies between 1.01 and 1.26. This range is much lower than the referenced figure, 1.35, indicating the presence of a high peak in the probability density function for all of the stocks under investigation. While testing for dependency, Ljung-Box Q statistics show that ten stocks are serially autocorrelated, and 27 of the 30 stocks are autocorrelated in the squared term as shown by a Q^2 test. The latter suggests a volatility clustering phenomenon and is consistent with a GARCH-type specification. By inspecting the Jarque-Bera statistics, the normality for all 30 stocks is uniformly rejected.[13]

The preliminary statistical results from Table 80.1 clearly indicate that the popular normality assumption does not conform to the weekly returns. The individual stock returns often show positive excess kurtosis (fat tails), accompanied by skewness. The evidence of peakedness is not in agreement with the normal distribution either. Besides the non-Gaussian features, some weekly stock returns present autocorrelation and almost all of them feature volatility clustering.

80.4 Empirical Evidence

In this section, we estimate the system of equations from Eq. 80.1a through 80.1d and present evidence of the GARCH(1,1) model based on different distributions. We analyze the impact of outliers on the EGB2 distribution.

80.4.1 GARCH(1,1) Model Based on the Normal Distribution

Table 80.2 reports the estimates of the GARCH(1,1) model based on the assumption that the error series follows a normal distribution, $\varepsilon_t|\Im_{t-1} \sim N(0, h_t)$.[14] Looking at the t-statistics, the null hypothesis of the absence of skewness is rejected at the 1 % level for 11 out of 30 cases (four positive and seven negative), while the null hypothesis of the absence of excess kurtosis is rejected in all of the cases. Moreover, the Jarque-Bera tests show that all of the return residuals are rejected by assuming

[13]All the non-normality features are more remarkable in daily data but less so in monthly data. This is consistent with Brown and Warner's (1985) report that the non-normal features tend to vanish in low-frequency data, such as monthly observations. Even so, subject to individual monthly stock returns, the Jarque-Bera test rejects the normality for 23 of the 30 stocks at the 1 % level.

[14]The standardized residuals are obtained by dividing the estimated regression residuals by its conditional standard deviation. Standardizing the error term makes the distribution comparison feasible. Mean and variance are not reported in the table due to the use of normalization.

Table 80.2 Estimates of the GARCH(1,1) normal distribution: weekly data, 1986–2005

Index	Skewness	Kurtosis	Peakedness	JB	Q(30)	$Q^2(30)$	ϕ_0	ϕ_1	ϕ_2	δ	w	α	β	γ
1	0.325	1.7464	1.2393	144.3966	33.8396	23.6315	0.0038	1.1478		−0.1035	0.000012	0.043	0.947	0.0064
	[4.19]***	[11.25]***		0***	0.29	0.79	[3.26]***	[19.60]***	[NA]	[−1.92]*	[1.46]	[3.75]***	[68.46]***	[0.30]
2	−0.1386	3.7809	1.1149	597.6384	40.1232	18.262	0.0002	1.1546		0.1514	0.000007	0.0251	0.9483	0.0475
	[−1.79]*	[24.36]***		0***	0.1	0.95	[0.20]	[23.30]***	[NA]	[4.72]***	[2.01]**	[2.73]***	[110.4]***	[2.47]**
3	−0.0996	1.7435	1.1815	128.0553	34.0617	19.7525	0.0019	0.9433	−0.043	0.1308	0.000024	0.0772	0.89	0.0115
	[−1.28]	[11.23]***		0***	0.28	0.92	[2.29]**	[21.38]***	[−1.69]*	[5.26]***	[2.00]**	[3.29]***	[25.35]***	[0.35]
4	0.1178	0.8954	1.1559	35.6498	30.6938	34.7945	0.0007	1.0984		0.0174	0.000008	0.0523	0.9403	−0.0022
	[1.52]	[5.77]***		0***	0.43	0.25	[0.84]	[28.66]***	[NA]	[0.63]	[1.90]*	[3.13]***	[69.26]***	[−0.10]
5	0.1453	0.9415	1.2957	40.3688	36.0064	18.1299	0.0021	0.6815	−0.172	0.1284	0.000004	0.0407	0.968	−0.0334
	[1.87]*	[6.06]***		0***	0.21	0.96	[3.51]***	[24.05]***	[−7.18]***	[0.00]	[8.01]***	[32.48]***	[857.7]***	[−14.1]***
6	0.2082	1.2395	1.226	71.1013	44.0888	28.2673	0.0015	1.1891	−0.0662	0.0525	0.000005	0.0284	0.9557	0.0097
	[2.68]***	[7.99]***		0***	0.05**	0.56	[2.50]**	[39.18]***	[−3.05]***	[2.91]***	[2.10]**	[3.24]***	[90.24]***	[0.52]
7	0.0232	1.7674	1.191	129.9834	37.1809	19.0182	−0.0006	1.0286		0.0042	0.000024	0.008	0.9443	0.0652
	[0.30]	[11.39]***		0***	0.17	0.94	[−0.58]	[22.84]***	[NA]	[0.14]	[1.77]*	[0.81]	[49.51]***	[3.10]***
8	−0.2299	2.4445	1.0648	257.2865	34.2906	28.8303	−0.0008	0.9271		0.0206	0.00001	0.0217	0.9296	0.0956
	[−2.96]***	[15.75]***		0***	0.27	0.53	[−0.87]	[20.08]***	[NA]	[0.93]	[1.99]**	[2.10]**	[66.45]***	[3.41]***
9	−0.6695	4.1205	1.0558	780.568	22.7618	23.094	0.0027	0.8083		−0.0296	0.000011	−0.007	6 0.9666	0.0615
	[−8.63]***	[26.54]***		0***	0.82	0.81	[2.87]***	[15.71]***	[NA]	[−1.09]	[3.62]***	[−0.92]	[127.9]***	[5.18]***
10	−0.4853	3.8586	1.1821	658.2901	36.9234	9.2842	0.0019	1.0402	−0.0725	−0.1718	0.000002	0.0439	0.9599	−0.0111
	[−6.25]***	[24.86]***		0***	0.18	1	[2.31]**	[25.05]***	[−2.97]***	[−5.12]***	[1.07]	[3.93]***	[158.7]***	[−0.67]
11	−0.4783	3.587	1.13	573.0948	46.8546	31.4136	0.0014	0.8322	−0.0987	0.1153	0.000029	0.0764	0.8365	0.14
	[−6.16]***	[23.08]***		0***	0.03**	0.4	[1.84]*	[21.56]***	[−4.11]***	[3.54]***	[2.84]***	[2.69]***	[26.21]***	[3.42]***
12	0.2028	3.7273	1.1323	584.5448	33.3369	16.408	0.0022	1.0168		−0.1516	0.000016	0.0019	0.9663	0.042
	[2.61]***	[24.01]***		0***	0.31	0.98	[1.94]*	[18.30]***	[NA]	[−4.42]***	[2.66]***	[0.27]	[113.5]***	[3.56]***
13	−0.0969	1.7346	1.1864	126.6834	26.0719	44.0327	0.0018	0.9425		0.0029	0.000011	0.0143	0.9523	0.0491
	[−1.25]	[11.17]***		0***	0.67	0.05**	[1.90]*	[17.05]***	[NA]	[0.09]	[2.00]**	[1.45]	[98.07]***	[2.37]**
14	−0.2625	1.7829	1.2095	143.6545	41.7312	32.5686	0.0014	0.9127		−0.0347	0.00005	0.0189	0.9101	0.0545

	[−3.38]***	[11.49]***		0***	0.08*	0.34	[1.39]	[17.52]***	[NA]	[−1.20]	[2.77]***	[1.17]	[36.57]***	[2.38]**
15	0.0521	1.0313	1.2023	44.682	39.197	23.4614	0.0021	0.827	−0.0929	0.0214	0.000028	0.0157	0.9135	0.0717
	[0.67]	[6.64]***		0***	0.12	0.8	[2.59]***	[18.58]***	[−3.50]***	[0.55]	[2.48]**	[0.82]	[34.30]***	[2.76]***
16	0.1481	1.6954	1.1093	123.1715	32.113	43.3573	0.0011	0.8627	−0.0839	−0.0368	0.000003	0.0095	0.9704	0.031
	[1.91]*	[10.92]***		0***	0.36	0.05**	[1.59]	[23.05]***	[−3.39]***	[−1.22]	[1.50]	[1.04]	[117.5]***	[2.37]**
17	−0.5395	5.5791	1.1373	1,342.749	35.2725	13.3108	0.002	0.8975	−0.0696	−0.0272	0.000014	−0.0052	0.9651	0.0536
	[−6.95]***	[35.94]***		0***	0.23	1	[2.04]**	[19.70]***	[−2.47]**	[−0.94]	[2.54]**	[−0.65]	[106.2]***	[3.74]***
18	0.1916	1.3209	1.2232	78.6646	48.917	37.3444	0.0008	1.129		−0.2635	0.00002	0.0399	0.9382	0.0151
	[2.47]**	[8.51]***		0***	0.02**	0.17	[0.76]	[19.86]***	[NA]	[−4.00]***	[2.14]**	[2.99]***	[66.10]***	[0.68]
19	0.0434	1.8057	1.1739	135.8917	31.3465	34.4645	0.0011	1.12		−0.0509	0.000009	0.0236	0.9627	0.0111
	[0.56]	[11.63]***		0***	0.4	0.26	[1.10]	[22.25]***	[NA]	[−1.75]*	[2.01]**	[1.99]**	[111.4]***	[0.64]
20	−0.0463	2.9953	1.1578	373.4356	37.8005	31.4578	0.0017	1.3192	−0.0634	−0.1321	0.000019	0.0214	0.9659	0.0053
	[−0.60]	[19.30]***		0***	0.15	0.39	[1.31]	[19.80]***	[−2.53]**	[−3.67]***	[1.86]*	[2.25]**	[108.9]***	[0.34]
21	0.0039	1.413	1.2017	83.0282	34.6267	27.0469	0.0009	0.8581		0.1522	0.000013	0.0158	0.9529	0.0344
	[0.05]	[9.10]***		0***	0.26	0.62	[0.97]	[17.25]***	[NA]	[5.58]***	[2.41]**	[1.39]	[80.21]***	[1.79]*
22	−0.1535	1.5692	1.1886	106.3135	38.6063	25.3614	−0.0001	1.28		−0.0775	0.000003	0.0212	0.9427	0.0764
	[−1.98]**	[10.11]***		0***	0.13	0.71	[−0.13]	[26.27]***	[NA]	[−1.74]*	[0.75]	[1.61]	[81.87]***	[3.51]***
23	0.11	1.0494	1.2199	47.807	33.1966	20.3827	0.0017	1.1453	−0.0764	0.0414	0.000013	0.0357	0.9436	0.0147
	[1.42]	[6.76]***		0***	0.31	0.91	[2.15]**	[24.08]***	[−2.97]***	[0.94]	[2.02]**	[2.86]***	[67.03]***	[0.85]
24	−0.0935	1.2285	1.2377	64.2169	38.8477	28.2932	0.0008	1.3293	−0.0455	0.0361	0.000004	0.0123	0.9607	0.0468
	[−1.20]	[7.91]***		0***	0.13	0.55	[0.92]	[33.18]***	[−2.12]**	[1.25]	[1.37]	[1.13]	[139.4]***	[2.26]**
25	−0.1157	1.1022	1.1877	52.7396	39.102	18.3382	0.0034	1.4716		−0.3448	0.000139	0.0918	0.8552	−0.0131
	[−1.49]	[7.10]***		0***	0.12	0.95	[2.43]**	[22.73]***	[NA]	[−19.6]***	[1.63]	[2.58]***	[15.05]***	[−0.38]
26	0.1525	1.4721	1.1727	93.9842	18.7153	25.0171	0.0004	0.7143	−0.1018	0.114	0.00002	0.0151	0.915	0.1031
	[1.97]**	[9.48]***		0***	0.95	0.72	[0.43]	[16.06]***	[−3.38]***	[4.37]***	[2.31]**	[1.02]	[42.86]***	[2.93]***
27	0.1414	1.7771	1.1934	134.6485	28.2358	46.9571	0.0011	0.8291	−0.0536	0.1442	0.000005	0.0527	0.9445	−0.0015
	[1.82]*	[11.45]***		0***	0.56	0.03**	[1.25]	[18.36]***	[−1.94]*	[5.09]***	[1.48]	[3.83]***	[96.88]***	[−0.08]
28	−0.24	1.7515	1.2185	137.1534	35.0204	49.3694	0.0039	1.328	−0.0679	−0.1298	0.000022	0.0583	0.9209	0.0117
	[−3.09]***	[11.28]***		0***	0.24	0.01**	[3.83]***	[24.83]***	[−2.93]***	[−3.15]***	[2.29]**	[3.74]***	[65.73]***	[0.62]

(*continued*)

Table 80.2 (continued)

Index	Skewness	Kurtosis	Peakedness	JB	Q(30)	$Q^2(30)$	ϕ_0	ϕ_1	ϕ_2	δ	w	α	β	γ
29	0.1677	2.281	1.1977	221.0298	21.721	34.2298	0.0018	1.1269	−0.0865	0.0229	0.000022	0.0359	0.9123	0.0516
	[2.16]**	[14.69]***		0***	0.86	0.27	[2.19]**	[24.40]***	[−3.53]***	[0.72]	[2.33]**	[2.68]***	[36.49]***	[1.60]
30	0.741	7.2363	1.1258	2,268.791	32.4082	15.2051	0.0024	1.4348	−0.0655	−0.0534	0.000096	0.0605	0.837	0.0388
	[9.55]***	[46.62]***		0***	0.35	0.99	[2.44]**	[29.36]***	[−2.77]***	[−1.36]	[2.43]**	[2.01]**	[15.74]***	[0.90]

The 30 stocks are sorted by permanent CRSP number. Numbers below coefficients are t-values (with bracket). Numbers below tests are p-values. ***Indicates 1 % significance, **5 %, and *10 %. The standard deviations of skewness and excess kurtosis coefficients are given approximately by $(6/T)^{0.5}$ and $(24/T)^{0.5}$, respectively. The peakedness is measured by $f_{0.75} - f_{0.25}$, the distance between the values of standardized variables at which the cumulative distribution function equals 0.75 and the value at which the cumulative distribution function equals 0.25. The reference value of the standard normal distribution is 1.35. A number of peakedness less than 1.35 means there is a high peak in the probability density function. The normality test is conducted using Jarque-Bera (JB) statistics. The independence test is conducted using a Ljung-Box Q test up to the order of 30. The Q^2 test up to the order of 30 is to show volatility clustering. The model is

(1.a) $R_t = \phi_0 + \phi_1 R_{m,t} + \phi_2 R_{t-1} + \delta D_{87} + \varepsilon_{it}$

(1.b) $\varepsilon_t = \sqrt{h_t} z_t$

(1.c) $h_t = w + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1} + \gamma\, I(\varepsilon_{t-1} < 0)\varepsilon_{t-1}^2$

(1.d) $\varepsilon_t | \Im t{-}1 \sim N(0, h_t)$. The following stocks do not have AR(1) variable: MSFT, HON, DD, GM, IBM, MO, CAT, BA, PFE, AA, DIS, MCD, JPM, and INTC. Stock PG is the only one to have an AR(4) variable in the mean equation, which ensures that autocorrelation has been removed

Gaussian distribution. If we check further into the measure of peakedness, the estimate values range from 1.06 to 1.30. All of these figures are lower than the reference point of a standard normal distribution, 1.35, indicating that all of the returns are leptokurtic. It is apparent that assuming that residuals for the estimated financial data are normally distributed is invalid.

80.4.2 GARCH(1,1) Model Based on the Student's *t* Distribution

Estimating the model by using a *t* distribution indicates that the excess kurtosis has mostly been removed from the estimated residuals. As shown in Table 80.3, 29 stocks show that the coefficients of excess kurtosis are insignificant. This demonstrates the effectiveness of *t* distribution in modeling the excess kurtosis. However, the problem of skewness has not been resolved at all. The evidence shows that 18 out of 30 are significant at the 5 % level or higher. There are four significant positive and eight significant negative skewness coefficients in the standardized residuals at the 1 % level.[15]

Another problem emerging from this model is insufficient peakedness of the distribution. The range of the estimated degree of freedom is (3.9 ~ 11.1), which corresponds to the range of peakedness (1.53 ~ 1.39). Note that the actual peakedness measurement from Table 80.3 is in the range of (1.02 ~ 1.29), indicating the presence of leptokurtosis. The evidence shows that the *t* distribution is worse than the normal distribution in modeling peakedness (see Fig. 80.1).

80.4.3 GARCH(1,1) Model Based on the EGB2 Distribution

To advance our study, we reestimate the GARCH(1,1) model by employing the EGB2 distribution. Table 80.4 reports the comparable statistics based on the standardized residuals from GARCH(1,1) cum EGB2 distribution: $\varepsilon_t | \Im_{t-1} \sim EGB2(0, h_t, p, q)$. The results show that the skewness problem for most cases has been alleviated by using the EGB2 distribution. The evidence indicates that only five stocks show the presence of skewness. Turning to the statistics of excess kurtosis, we find that the EGB2 distribution works well on some stocks' kurtosis but not all of them. The evidence in Table 80.4 indicates that nine stocks still show excess kurtosis.

Table 80.4 also contains the range of p (0.334 ~ 1.776) and of q (0.348 ~ 1.669). The reported p- and q-values suggest that the residuals' distributions are far from

[15] To deal with the skewness, a number of skewed *t* distributions have been proposed (Theodossiou 1998; Hueng and Brooks 2003). One obvious drawback of a skewed *t* distribution in our study is the outcome of its peakedness measurement, which displays platykurtosis (flat topped in density). This appears to be the opposite of the leptokurtic stock returns. For this reason, we do not report results from the GARCH model based on a skewed *t* distribution in order to focus on the EGB2 distribution. However, the results are available upon request.

Table 80.3 Statistics of the standardized errors on the GARCH(1,1)-*t* distribution: weekly data, 1986–2005

Index	Skewness	Kurtosis	Peakedness	Q(30)	$Q^2(30)$	ϕ_0	ϕ_1	ϕ_2	δ	w	α	β	γ	v
1	0.343	1.7838	1.2436	33.2195	23.0715	0.0025	1.1206		−0.1049	1.1E-05	0.0476	0.942	0.0103	6.4635
	[4.42]***	[−0.53]	(1.44)	0.31	0.81	[2.37]**	[20.40]***	[NA]	[−2.37]**	[1.04]	[3.09]***	[44.63]***	[0.37]	
2	−0.2319	4.7297	1.1348	40.0766	21.0042	0	1.0971		0.1458	1.9E-05	0.0298	0.939	0.0277	4.5243
	[−2.99]***	[−0.50]	(1.48)	0.1	0.89	[−0.00]	[23.66]***	[NA]	[7.48]***	[2.36]**	[2.30]**	[62.86]***	[1.21]	
3	−0.1337	1.994	1.1909	33.3153	20.5154	0.0019	0.9345	−0.0596	0.1281	9E-06	0.0687	0.9273	−0.0097	6.7609
	[−1.72]*	[−0.17]	(1.44)	0.31	0.9	[2.49]**	[23.49]***	[−2.41]**	[6.46]***	[0.99]	[3.12]***	[28.16]***	[−0.30]	
4	0.1218	0.9512	1.1426	30.255	33.7479	0.0004	1.0819		0.016	9E-06	0.0598	0.929	0.0065	6.7749
	[1.57]	[−1.02]	(1.44)	0.45	0.29	[0.55]	[27.61]***	[NA]	[0.84]	[1.60]	[2.55]**	[49.02]***	[0.22]	
5	0.1278	1.0005	1.2934	34.3873	16.7991	0.002	0.6895	−0.1679	0.1297	6E-06	0.0501	0.9537	−0.0284	11.1396
	[1.65]*	[0.44]	(1.39)	0.27	0.97	[3.09]***	[22.69]***	[−6.81]***	[0.00]	[5.48]***	[20.20]**	[425.1]***	[−6.14]***	
6	0.2187	1.3807	1.2375	45.1364	27.5533	0.0012	1.1779	−0.061	0.0524	4E-06	0.0295	0.9515	0.0224	8.6249
	[2.82]***	[0.14]	(1.41)	0.04**	0.59	[2.04]**	[38.69]***	[−3.49]***	[2.98]***	[1.49]	[2.43]**	[79.40]***	[0.97]	
7	0.0013	2.0779	1.1981	36.1023	18.2453	−0.001	1.0089		0.0026	1.1E-05	0.0154	0.9586	0.0433	6.2797
	[0.02]	[−0.39]	(1.44)	0.2	0.95	[−1.01]	[21.65]***	[NA]	[0.10]	[1.23]	[1.31]	[78.94]***	[2.03]**	
8	−0.2701	2.8559	1.0195	34.8834	28.569	−0.001	0.9383		0.0218	5E-06	0.0297	0.9489	0.0491	3.9057
	[−3.48]***	NA	(1.53)	0.25	0.54	[−1.20]	[24.54]***	[NA]	[1.35]	[1.06]	[2.05]**	[72.36]***	[1.89]*	
9	−0.706	4.5322	1.0486	22.5016	21.7429	0.0035	0.804		−0.0309	1.4E-05	0.0176	0.9518	0.0397	4.0701
	[−9.10]***	[−0.12]	(1.48)	0.84	0.86	[3.79]***	[16.44]***	[NA]	[−1.61]	[2.07]**	[0.81]	[57.86]***	[1.75]*	
10	−0.8208	6.9991	1.1804	33.9934	7.8299	0.0021	1.0041	−0.0877	−0.1768	1.5E-05	0.0539	0.9273	0.001	5.9383
	[−10.5]***	[2.28]**	(1.45)	0.28	1	[2.69]***	[26.12]***	[−3.61]***	[−7.95]***	[1.58]	[2.17]**	[31.36]***	[0.04]	
11	−0.8627	8.5655	1.1322	48.1859	21.2743	0.0022	0.8264	−0.1159	0.1119	0.000015	0.0492	0.9235	0.0176	5.1337
	[−11.1]***	[0.88]	(1.45)	0.02**	0.88	[3.01]***	[21.67]***	[−4.51]***	[5.75]***	[2.88]***	[2.69]***	[61.63]***	[0.65]	
12	0.1705	4.2164	1.1414	34.2222	14.6933	0.0017	1.0317		−0.1496	1.4E-05	0.0201	0.9592	0.0231	4.8239
	[2.20]**	[−0.53]	(1.48)	0.27	0.99	[1.69]*	[21.07]***	[NA]	[−5.66]***	[1.48]	[1.34]	[65.73]***	[1.30]	
13	−0.1782	2.0983	1.1979	26.1595	46.3158	0.0015	0.9043		−0.0005	1.9E-05	0.0132	0.9502	0.0414	6.4902
	[−2.30]**	[−0.25]	(1.44)	0.67	0.03**	[1.50]	[18.78]***	[NA]	[−0.02]	[2.23]**	[1.03]	[72.45]***	[1.70]*	

14	−0.2855	2.0056	1.1918	40.2125	30.7673	0.0018	0.92		−0.0344	2.9E-05	0.03	0.9247	0.044	6.1476
	[−3.68]***	[−0.52]	(1.44)	0.1	0.43	[1.94]*	[20.01]***	[NA]	[−1.64]*	[1.81]*	[1.56]	[36.67]***	[1.55]	
15	0.0652	1.0808	1.1834	39.0689	25.1389	0.0018	0.8325	−0.0783	0.0243	2.3E-05	0.0226	0.9211	0.059	6.828
	[0.84]	[−0.94]	(1.44)	0.12	0.72	[2.34]**	[21.67]***	[−3.08]***	[0.98]	[2.06]**	[1.02]	[32.82]***	[2.15]**	
16	0.1556	1.6957	1.0909	32.1196	42.3807	0.0011	0.8569	−0.0758	−0.0366	4E-06	0.0121	0.9688	0.0277	4.863
	[2.00]**	[−0.86]	(1.48)	0.36	0.07*	[1.60]	[24.26]***	[−3.35]***	[−2.23]**	[1.47]	[0.79]	[106.0]***	[1.81]*	
17	−0.8697	10.312	1.1498	31.9551	7.3149	0.0019	0.9572	−0.0651	−0.0208	2.6E-05	0.0285	0.9234	0.0484	5.5646
	[−11.2]***	[2.80]***	(1.45)	0.37	1	[2.16]**	[21.44]***	[−2.47]**	[−0.96]	[2.27]**	[1.57]	[44.28]***	[1.69]*	
18	0.1839	1.3727	1.2229	49.1214	37.5822	0.0004	1.1146		−0.2645	1.6E-05	0.0326	0.9427	0.0282	7.4977
	[2.37]**	[−0.43]	(1.42)	0.02**	0.16	[0.43]	[22.39]***	[NA]	[−7.62]***	[1.50]	[2.23]**	[58.66]***	[1.14]	
19	0.0404	1.8821	1.1772	30.5184	31.6679	0.0011	1.1197		−0.0509	1.2E-05	0.0295	0.9512	0.0162	5.9915
	[0.52]	[−0.69]	(1.45)	0.44	0.38	[1.19]	[26.38]***	[NA]	[−2.38]**	[1.57]	[1.90]*	[62.59]***	[0.70]	
20	−0.0672	3.2296	1.1420	37.5776	30.4423	0.0017	1.2807	−0.0657	−0.1361	1.6E-05	0.0294	0.9602	0.0057	4.8937
	[−0.87]	[−0.69]	(1.48)	0.16	0.44	[1.39]	[22.11]***	[−2.79]***	[−4.78]***	[1.42]	[2.29]**	[85.02]***	[0.28]	
21	−0.0207	1.6059	1.2007	34.3777	27.0955	0.0005	0.865		0.1532	9E-06	0.0179	0.955	0.0394	6.6886
	[−0.27]	[−0.55]	(1.44)	0.27	0.62	[0.55]	[20.89]***	[NA]	[6.54]***	[1.51]	[1.39]	[73.68]***	[1.61]	
22	−0.1468	1.6039	1.1651	38.9648	25.0082	0.0001	1.3264		−0.0731	4E-06	0.0217	0.9422	0.0748	6.3138
	[−1.89]*	[−0.72]	(1.44)	0.13	0.72	[0.12]	[27.83]***	[NA]	[−2.56]**	[0.81]	[1.28]	[63.96]***	[2.90]***	
23	0.1166	1.0596	1.2152	33.3934	21.1235	0.0014	1.1359	−0.0768	0.0406	1.3E-05	0.0379	0.9393	0.0202	7.765
	[1.50]	[−0.71]	(1.42)	0.31	0.88	[1.71]*	[26.23]***	[−3.28]***	[1.28]	[1.30]	[2.49]**	[46.50]***	[0.92]	
24	−0.1277	1.3495	1.2377	37.6296	28.6486	0.0006	1.3136	−0.0494	0.0342	8E-06	0.0138	0.952	0.0545	8.2353
	[−1.65]*	[−0.10]	(1.41)	0.16	0.54	[0.71]	[30.37]***	[−1.99]**	[1.35]	[1.59]	[0.98]	[82.43]***	[2.10]**	
25	−0.1312	1.2702	1.1790	39.0144	28.3397	0.0038	1.45		−0.2712	9E-06	0.0514	0.9609	−0.0311	6.8554
	[−1.69]*	[−0.82]	(1.44)	0.13	0.55	[2.90]***	[20.87]***	[NA]	[−7.00]***	[0.53]	[16.39]**	[60.67]***	[−1.56]	
26	0.2108	1.9665	1.1407	19.1298	26.7937	0.0002	0.7986	−0.0871	0.1231	1.9E-05	0.0515	0.8996	0.0672	5.8863
	[2.72]***	[−0.67]	(1.45)	0.94	0.63	[0.23]	[19.30]***	[−3.28]***	[5.48]***	[1.94]*	[2.05]**	[36.66]***	[1.60]	
27	0.1706	1.9752	1.1781	27.9549	46.7514	0.0009	0.8697	−0.05	0.1487	8E-06	0.0621	0.9383	−0.015	6.2806
	[2.20]**	[−0.48]	(1.44)	0.57	0.03**	[1.20]	[20.74]***	[−1.97]**	[6.38]***	[1.51]	[3.27]***	[68.32]***	[−0.51]	

(continued)

Table 80.3 (continued)

Index	Skewness	Kurtosis	Peakedness	Q(30)	$Q^2(30)$	ϕ_0	ϕ_1	ϕ_2	δ	w	α	β	γ	v
28	−0.2452	1.9163	1.2308	34.7149	53.8793	0.004	1.2955	−0.067	−0.1331	2.5E-05	0.0515	0.9275	0.0041	6.8976
	[−3.16]***	[−0.16]	(1.44)	0.25	0***	[3.78]***	[24.05]***	[−2.83]***	[−4.47]***	[1.94]*	[3.10]***	[54.01]***	[0.18]	
29	0.1626	2.4257	1.1944	21.4156	31.1321	0.0012	1.1447	−0.0772	0.0254	0.00002	0.0403	0.9188	0.0332	6.849
	[2.10]**	[0.33]	(1.44)	0.87	0.41	[1.55]	[26.33]***	[−3.46]***	[1.07]	[1.84]*	[2.29]**	[33.77]***	[1.10]	
30	0.7595	10.0335	1.1622	31.9395	10.6979	0.0016	1.406	−0.0759	−0.0565	1.1E-05	0.0189	0.9579	0.0278	5.316
	[9.79]***	[1.90]*	(1.45)	0.37	1	[1.85]*	[30.66]***	[−3.58]***	[−1.87]*	[1.16]	[1.82]*	[59.90]***	[1.27]	

The 30 stocks are sorted by permanent CRSP number. Numbers below coefficients are t-values (with bracket). Numbers below tests are p-values. ***Indicates 1 % significance, **5 %, and *10 %. The standard deviations of skewness coefficients are given approximately by $(6/T)^{0.5}$. The excess kurtosis coefficient of t distribution is given by $\frac{6}{v-4}$ for $v > 4$. Its standard deviation is obtained using the delta method. The peakedness is measured by $f_{0.75} - f_{0.25}$, the distance between the values of standardized variable at which the cumulative distribution function equals 0.75 and the value at which the cumulative distribution function equals 0.25. The reference value of the standard normal distribution is 1.35. The reference value for the estimated t distribution is reported below actual peakedness (in parenthesis), which is in the range (1.39, 1.53). A number of peakedness less than the reference value means there is a high peak in the probability density function. The normality test is omitted, since the assumption is a Student's t distribution. The independence test is conducted using a Ljung-Box Q test up to the order of 30. The Q^2 test up to the order of 30 is to show volatility clustering. The model is

(1.a) $R_t = \phi_0 + \phi_1 R_{m,t} + \phi_2 R_{t-1} + \delta D_{87} + \varepsilon_{it}$

(1.b) $\varepsilon_t = \sqrt{h_t} z_t$

(1.c) $h_t = w + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1} + \gamma\, I(\varepsilon_{t-1} < 0)\varepsilon_{t-1}^2$

(1.d) $\varepsilon_t | \Im_{t-1} \sim t(0, h_t, v)$

The following stocks do not have an AR(1) variable: MSFT, HON, DD, GM, IBM, MO, CAT, BA, PFE, AA, DIS, MCD, JPM, and INTC. Stock PG is the only one to have an AR(4) variable in the mean equation, which ensures that autocorrelation has been removed

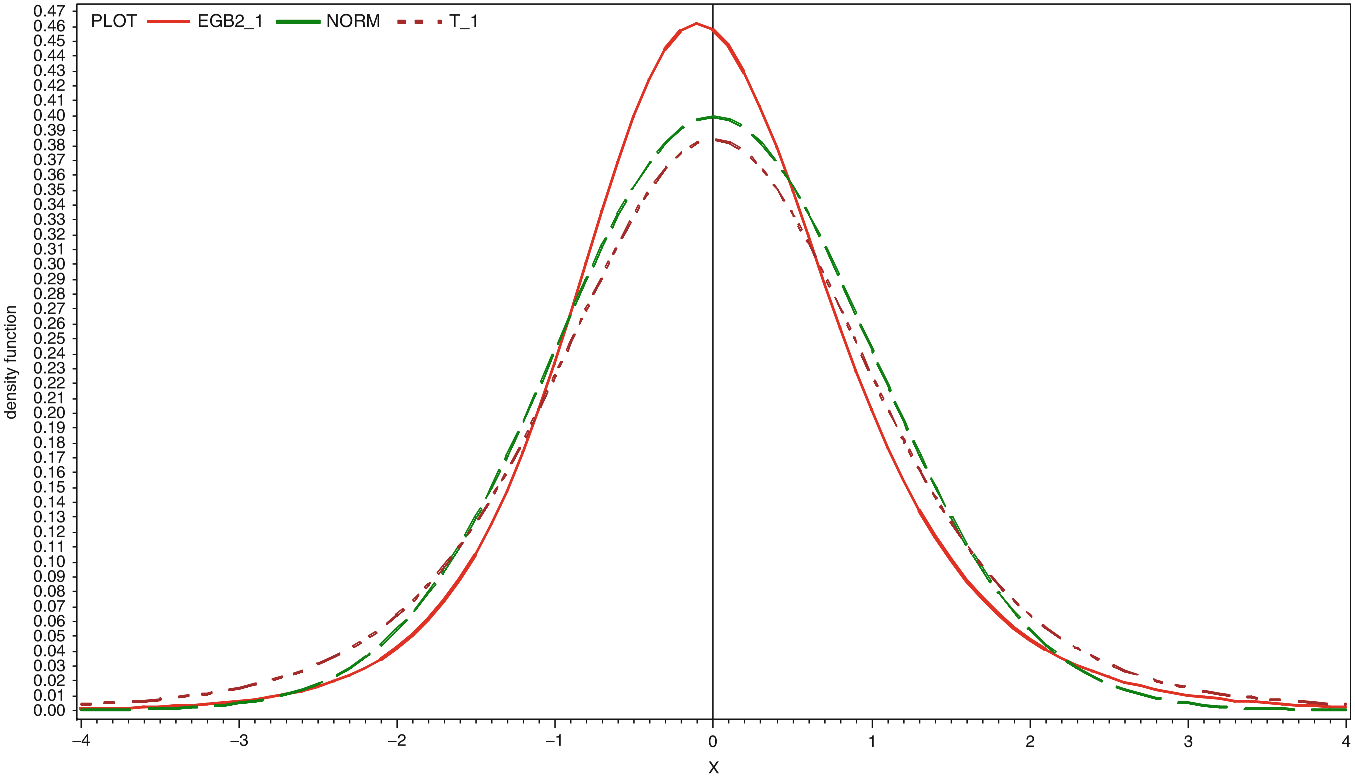


Fig. 80.1 Comparisons of probability density function of three distributions. Note: This chart compares the probability density function (pdf) of three distributions. The *solid line* is for the EGB2 distribution, *dashed line* for the normal distribution, and *dash-dot line* is for the Student's t distribution. Distribution-estimated parameters are from stock MSFT (index = 1): $p = 1.0233$; $q = 0.7971$; $v = 6.4635$

Table 80.4 Statistics of the standardized errors on the GARCH(1,1)-EGB2 estimates: weekly data, 1986–2005

Index	Skewness	Kurtosis	Peakedness	Q(30)	$Q^2(30)$	ϕ_0	ϕ_1	ϕ_2	δ	w	α	β	γ	p	q
1	0.3412	1.7765	1.2455	33.4102	22.9306	0.0034	1.1176		−0.101	1.1E-05	0.0473	0.9435	0.0045	1.0233	0.7971
	[0.43]	[1.18]	(1.20)	0.31	0.82	[3.23]***	[21.84]***	[NA]	[−2.27]**	[1.08]	[3.23]***	[45.21]***	[0.17]		
2	−0.2048	4.4376	1.1664	40.0458	19.8209	0	1.1021		0.1467	1.4E-05	0.0269	0.9437	0.029	0.5436	0.5274
	[−1.87]*	[9.86]***	(1.13)	0.1	0.92	[−0.01]	[26.13]***	[NA]	[7.30]***	[2.67]***	[2.49]**	[79.79]***	[1.50]		
3	−0.1272	1.9475	1.1934	33.3921	20.4258	0.002	0.9353	−0.056	0.1291	1.1E-05	0.0694	0.9241	−0.0101	0.8898	0.8338
	[−1.65]*	[1.70]*	(1.19)	0.31	0.9	[2.94]***	[21.09]***	[−2.19]**	[6.63]***	[0.97]	[2.89]**	[23.65]***	[−0.29]		
4	0.1248	0.953	1.1425	30.344	33.8905	0.0008	1.0782		0.0169	9E-06	0.0642	0.9277	−0.0007	0.7613	0.6589
	[−0.49]	[−1.75]*	(1.17)	0.45	0.29	[0.95]	[28.14]***	[NA]	[0.91]	[1.45]	[2.61]***	[44.76]***	[−0.02]		
5	0.1341	0.9897	1.2981	34.7639	17.1257	0.0021	0.6903	−0.1673	0.1301	5E-06	0.0481	0.9584	−0.0317	1.776	1.6686
	[0.72]	[1.18]	(1.26)	0.25	0.97	[2.95]***	[19.25]***	[−6.31]***	[6.99]***	[1.00]	[2.81]***	[32.60]***	[−1.26]		
6	0.2233	1.3534	1.2350	45.1204	27.1924	0.0014	1.1795	−0.0632	0.0532	4E-06	0.032	0.9529	0.0141	1.378	1.1263
	[0.28]	[1.06]	(1.23)	0.04**	0.61	[2.13]**	[39.92]***	[−3.10]***	[3.53]***	[1.49]	[2.42]**	[81.42]***	[0.58]		
7	0.0048	1.9897	1.2069	36.3661	18.4223	−0.0007	1.0121		0.0036	1.3E-05	0.0136	0.956	0.0463	0.8337	0.7498
	[−0.98]	[1.52]	(1.18)	0.2	0.95	[−0.65]	[21.31]***	[NA]	[0.15]	[1.33]	[1.28]	[69.88]***	[2.16]**		
8	−0.257	2.7399	1.0606	34.8276	28.5692	−0.0012	0.9455		0.0223	6E-06	0.0262	0.944	0.056	0.3342	0.3477
	[−1.22]	[1.48]	(1.08)	0.25	0.54	[−1.35]	[23.95]***	[NA]	[1.61]	[1.24]	[1.96]**	[64.61]***	[2.13]**		
9	−0.6971	4.3728	1.0869	22.6857	21.929	0.0026	0.8023		−0.0322	1.3E-05	0.0054	0.9573	0.0486	0.3736	0.4322
	[−3.07]***	[7.39]***	(1.09)	0.83	0.86	[2.64]***	[18.33]***	[NA]	[−1.97]**	[2.46]**	[0.35]	[71.58]***	[2.73]***		
10	−0.7098	5.8553	1.1884	34.6538	8.0929	0.0017	1.009	−0.087	−0.1769	0.00001	0.0458	0.9423	−0.003	0.6849	0.7425
	[−4.88]***	[12.84]***	(1.17)	0.26	1	[2.23]**	[27.68]***	[−3.83]***	[−8.36]***	[1.72]*	[2.60]***	[54.62]***	[−0.13]		
11	−0.6741	6.0631	1.1326	46.4982	26.043	0.0017	0.8242	−0.1203	0.11	1.9E-05	0.0518	0.9004	0.0504	0.5305	0.6015
	[−3.52]***	[12.84]***	(1.14)	0.03**	0.67	[2.32]**	[22.70]***	[−4.71]***	[6.16]***	[2.44]**	[2.32]**	[34.82]***	[1.31]		
12	0.1834	4.1096	1.1571	34.3886	15.0915	0.0024	1.0259		−0.1484	1.6E-05	0.017	0.9585	0.0245	0.609	0.5251
	[−0.26]	[7.69]***	(1.14)	0.27	0.99	[2.18]**	[20.40]***	[NA]	[−6.07]***	[1.68]*	[1.22]	[64.75]***	[1.45]		

13	−0.1586	2.0019	1.2016	26.1503	45.7982	0.0016	0.912		0.0008	1.7E-05	0.0128	0.9503	0.0442	0.866	0.8183
	[−1.67]*	[2.05]**	(1.19)	0.67	0.03**	[1.70]*	[17.47]***	[NA]	[0.03]	[2.22]**	[1.00]	[74.53]***	[1.84]*		
14	−0.2815	1.9679	1.2027	40.3874	30.9281	0.0013	0.9224		−0.0352	3.2E-05	0.0276	0.9212	0.0483	0.748	0.8582
	[−0.80]	[1.47]	(1.18)	0.1	0.42	[1.38]	[18.75]***	[NA]	[−1.67]*	[2.04]**	[1.52]	[37.57]***	[1.63]		
15	0.0796	1.0827	1.1947	39.075	24.6498	0.0021	0.8394	−0.0769	0.0268	2.4E-05	0.0306	0.9128	0.0564	0.8794	0.7645
	[−0.74]	[−1.16]	(1.19)	0.12	0.74	[2.53]**	[21.19]***	[−3.00]***	[1.03]	[2.08]**	[1.14]	[30.49]***	[1.92]*		
16	0.1496	1.703	1.1076	32.4643	43.0084	0.0011	0.8548	−0.074	−0.0368	3E-06	0.0089	0.9711	0.0299	0.4835	0.4921
	[1.29]	[−1.12]	(1.12)	0.35	0.06*	[1.57]	[25.54]***	[−3.22]***	[−2.62]***	[1.42]	[0.82]	[105.5]***	[1.88]*		
17	−0.799	9.2326	1.1596	32.7319	8.4127	0.0019	0.9539	−0.0652	−0.021	2.4E-05	0.0206	0.9335	0.0452	0.5532	0.5478
	[−5.64]***	[21.66]***	(1.14)	0.33	1	[2.01]**	[23.75]***	[−2.52]**	[−1.07]	[2.20]**	[1.25]	[46.75]***	[1.68]*		
18	0.2006	1.3321	1.2287	49.0945	38.2149	0.0009	1.1161		−0.262	1.9E-05	0.0349	0.9417	0.0183	1.1292	0.9509
	[0.15]	[0.43]	(1.21)	0.02**	0.14	[0.87]	[20.81]***	[NA]	[−8.05]***	[1.67]*	[2.31]**	[56.20]***	[0.72]		
19	0.0419	1.8661	1.1842	30.6913	32.2694	0.0012	1.1209		−0.0504	1.1E-05	0.0287	0.9535	0.014	0.7224	0.6921
	[−0.11]	[0.85]	(1.17)	0.43	0.36	[1.26]	[25.84]***	[NA]	[−2.66]***	[1.59]	[1.67]*	[69.68]***	[0.55]		
20	−0.0622	3.1716	1.1715	37.46	30.6583	0.0017	1.277	−0.0667	−0.1363	1.7E-05	0.0261	0.9614	0.0057	0.6121	0.5983
	[−0.66]	[5.14]***	(1.15)	0.16	0.43	[1.30]	[20.89]***	[−2.69]***	[−5.28]***	[1.62]	[2.14]**	[96.11]***	[0.30]		
21	−0.0083	1.5496	1.2055	34.3959	27.1835	0.0008	0.8637		0.1543	9E-06	0.0194	0.9549	0.0332	0.8679	0.7485
	[−1.47]	[0.21]	(1.18)	0.27	0.61	[0.88]	[21.01]***	[NA]	[7.51]***	[1.66]*	[1.45]	[75.20]***	[1.57]		
22	−0.1528	1.6012	1.1773	38.8848	24.9758	−0.0001	1.322		−0.0747	4E-06	0.0197	0.9422	0.0788	0.7846	0.8586
	[−0.31]	[0.50]	(1.19)	0.13	0.73	[−0.12]	[29.64]***	[NA]	[−2.45]**	[0.75]	[1.24]	[66.75]***	[2.76]***		
23	0.122	1.0602	1.2103	33.3633	21.3757	0.0018	1.135	−0.0739	0.043	1.5E-05	0.0397	0.936	0.0198	1.0471	0.879
	[−0.53]	[−0.68]	(1.20)	0.31	0.88	[1.98]**	[26.40]***	[−2.96]***	[1.47]	[1.63]	[2.48]**	[47.76]***	[0.88]		
24	−0.1166	1.321	1.2361	37.8084	28.7249	0.0007	1.3141	−0.0483	0.0357	7E-06	0.0151	0.9543	0.0485	1.1216	0.9945
	[−1.86]*	[0.48]	(1.22)	0.15	0.53	[0.84]	[33.10]***	[−2.29]**	[1.45]	[1.36]	[1.07]	[86.97]***	[1.84]*		
25	−0.1305	1.2595	1.1869	39.0536	28.8414	0.0032	1.4509		−0.2732	8E-06	0.0493	0.9632	−0.0312	0.7835	0.9016
	[0.30]	[−0.52]	(1.19)	0.12	0.53	[2.37]**	[23.06]***	[NA]	[−6.75]***	[0.44]	[2.99]***	[50.58]***	[−1.42]		
26	0.2048	1.9035	1.1586	19.1195	26.5286	0.0004	0.7923	−0.0884	0.1229	1.8E-05	0.0454	0.9039	0.0669	0.7541	0.7163
	[0.99]	[1.20]	(1.17)	0.94	0.65	[0.44]	[20.62]***	[−3.31]***	[6.16]***	[1.86]*	[1.99]**	[37.62]***	[1.67]*		

(*continued*)

Table 80.4 (continued)

Index	Skewness	Kurtosis	Peakedness	Q(30)	$Q^2(30)$	ϕ_0	ϕ_1	ϕ_2	δ	w	α	β	γ	p	q
27	0.1627	1.9351	1.1779	27.9691	46.5729	0.001	0.873	−0.0499	0.149	7E-06	0.0595	0.9396	−0.0114	0.724	0.7236
	[1.19]	[1.01]	(1.17)	0.57	0.03**	[1.09]	[19.65]***	[−1.82]*	[6.82]***	[1.51]	[3.19]***	[69.05]***	[−0.40]		
28	−0.2448	1.873	1.2342	34.7132	52.5435	0.0037	1.3029	−0.0683	−0.1334	2.4E-05	0.0531	0.9257	0.0055	0.8967	0.9744
	[−1.27]	[2.01]**	(1.20)	0.25	0.01**	[3.47]***	[26.02]***	[−2.77]***	[−4.64]***	[1.86]*	[3.16]***	[51.93]***	[0.25]		
29	0.1727	2.3833	1.2032	21.5452	32.2833	0.0016	1.1419	−0.0754	0.0273	2.2E-05	0.0409	0.9138	0.0348	0.9941	0.8369
	[−0.21]	[3.43]***	(1.20)	0.87	0.35	[1.90]*	[27.31]***	[−3.26]***	[1.04]	[1.91]*	[2.28]**	[30.42]***	[1.10]		
30	0.754	9.3273	1.1708	33.0307	11.0894	0.0023	1.4123	−0.0674	−0.0517	1.9E-05	0.0271	0.9441	0.02	0.6879	0.5446
	[3.64]***	[29.70]***	(1.15)	0.32	1	[2.43]**	[31.60]***	[−3.11]***	[−2.13]**	[1.17]	[1.67]*	[34.58]***	[0.86]		

The 30 stocks are sorted by permanent CRSP number. Numbers below coefficients are t-values (with bracket). Numbers below Q test and Q^2 test are p-values. ***Indicates 1 % significance, **5 %, and *10 %. The predicted higher moments are given by the formula: Skewness $= \frac{\psi''(p)-\psi''(q)}{(\psi'(p)+\psi'(q))^{1.5}}$, Kurtosis $= \frac{\psi'''(p)+\psi'''(q)}{(\psi'(p)+\psi'(q))^{2}}$. Their standard deviations are obtained using the delta method. The peakedness is measured by $f_{0.75}$-$f_{0.25}$, the distance between the values of standardized variable at which the cumulative distribution function equals 0.75 and the value at which the cumulative distribution function equals 0.25. The reference value of the standard normal distribution is 1.35. A number of peakedness less than 1.35 means there is a high peak in the probability density function. The reference value of the EGB2 distribution is reported below peakedness with parenthesis and is in the range of (1.07, 1.26). The normality test is omitted, since the assumption is the EGB2 distribution. The independence test is conducted using a Ljung-Box Q test up to the order of 30. The Q^2 test up to the order of 30 is to show volatility clustering. The estimated model is

(1.a) $R_t = \phi_0 + \phi_1 R_{m,t} + \phi_2 R_{t-1} + \delta D_{87} + \varepsilon_{it}$
(1.b) $\varepsilon_t = \sqrt{h_t} z_t$
(1.c) $h_t = w + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1} + \gamma\, I(\varepsilon_{t-1} < 0)\varepsilon_{t-1}^2$
(1.d) $\varepsilon_t | \Im_{t-1} \sim EGB2(0, h_t, p, q)$

An AR(1) variable is excluded from the following stocks in the estimated equation: MSFT, HON, DD, GM, IBM, MO, CAT, BA, PFE, AA, DIS, MCD, JPM, and INTC. Stock PG contains the only AR(4) variable in the mean equation the normal distribution that requires that both p- and q-values approach infinity. Based on the estimated shape parameters, the expected peakedness for the 30 stocks is in the range of (1.07 ∼ 1.26). The peakedness obtained from residuals of the mean equation is in the range of (1.06 ∼ 1.30), conforming to the existence of high peak implied by the EGB2 distribution.

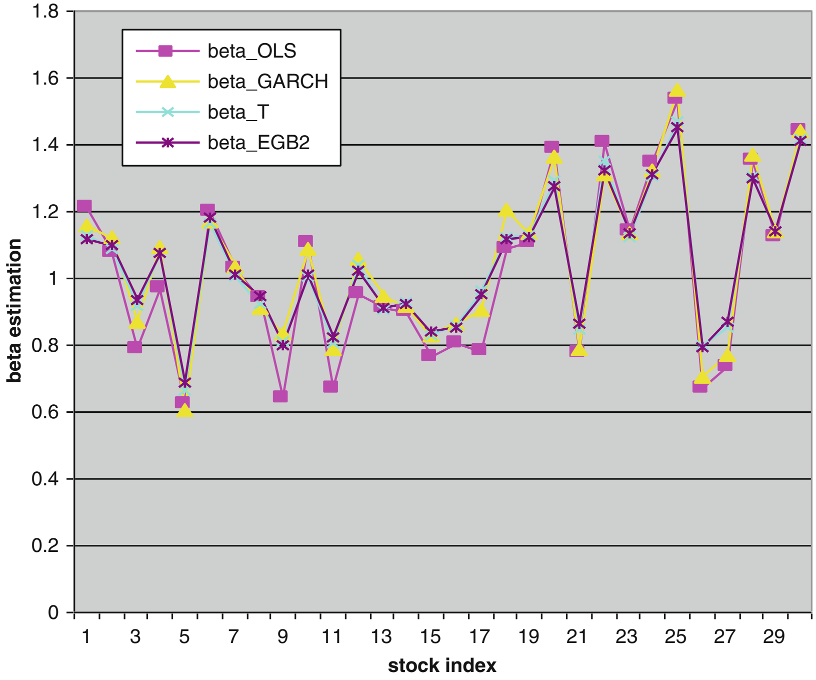
With respect to the beta coefficients, we find that the estimated values are highly significant, ranging from 0.69 to 1.32. The evidence suggests that the market risk is still one of the most influential factors for predicting individual stocks. It is of interest to compare the beta values and the associated standard errors across different distributions. As may be seen from Fig. 80.2, where the figures are mainly reproduced from Tables 80.2, 80.3 and 80.4, we find no significant difference among them for the estimated betas. This is not surprising, since the estimations of the betas are obtained from the average effect based on the whole probability space. Our finding is consistent with the results form Nelson (1991) and Hansen (1994).

While inspecting the lagged individual stock return variable, about half of them present negative signs and are statistically significant, indicating that a mean reversion process is present in the weekly data. Turning to the 1987 market crash dummy, the testing results show that 20 out of 30 stocks are significant at the 5 % level, although the signs are mixed. The diverse movements signify the profound impact due to an influential observation. Consistent with most financial data, with a few exceptions, the coefficients of the GARCH equation for each stock are found to be highly significant.

One of the most striking results emerging from the estimations is that while testing the leverage effect, only four stocks are found to be statistically significant at the 5 % level. The number of stocks that present asymmetric effects has been reduced dramatically, as compared with the statistics reported in Table 80.2, where 15 stocks show a significant asymmetric effect. It can be argued that the so-called asymmetric effect may result from the fact that empirical analysis was built on a misleading assumption by imposing a normal distribution on the financial data.

One fact in Table 80.4 is a bit disturbing: three stocks show their kurtosis coefficient being greater than 6, which is beyond the scope of EGB2 distribution. Despite this shortcoming and the abovementioned nine stocks that have significant kurtosis, we find a significant improvement compared with the model by assuming the normal distribution or the t distribution. The predicted skewness and excess kurtosis of the EGB2 distribution are much closer to the observed skewness and kurtosis. Thus, the EGB2 distribution has a good fit, although the results are not perfect.[16] Finally, we test the independence of the correlation on statistics for both return level and return squares, and we find that in only three cases in each Q and Q^2 test can the null hypothesis be rejected at the 5 % level but none at the 1 % level. In general, the models are adequate.

[16] Some refinement of the model is contained in the following section.



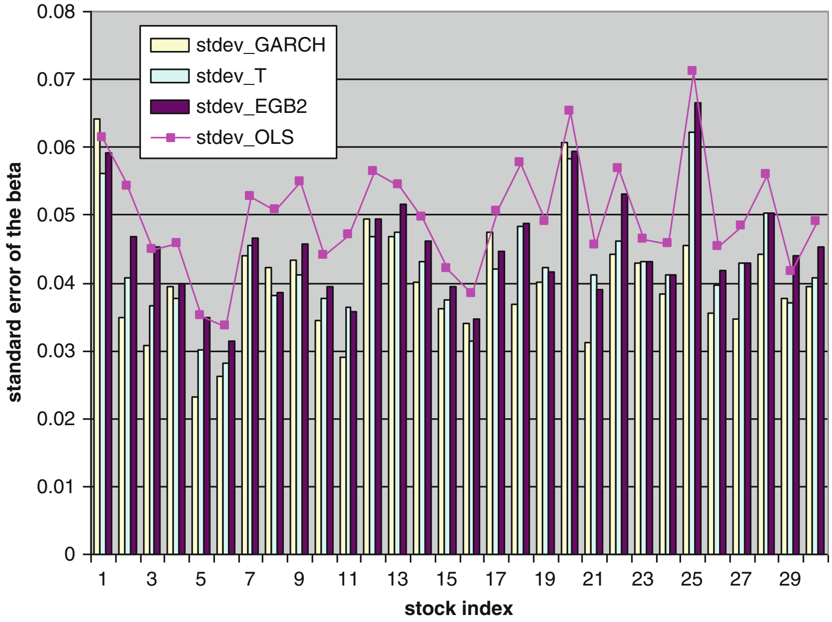


Fig. 80.2 Comparisons of the beta estimation in different models. Note: The *upper* figure contains the plots of beta coefficients. The *lower* figure presents the corresponding standard deviations of the beta coefficients

80.4.4 The Impact of Outliers

Theoretically, the EGB2 distribution is feasible for coefficients of skewness in a range of (−2, 2) and the coefficients of excess kurtosis in (0, 6). However, the statistics in Table 80.4 are not completed in these desired ranges. Two possible reasons might contribute to these problems. First, the residual series was contaminated by the presence of outliers. As pointed out by Peña et al. (2001), an outlier can have very serious effects on the properties of the observed time series and affect the estimated residuals and the parameter values. Second, the mean equation and/or the variance equation may be misspecified, although an asymmetric effect has been considered.[17] To address this issue, we further investigate the stock return series on which the outliers might more seriously impinge.

By investigating the microstructure of the nine stocks with excess kurtosis, we find a common phenomenon: multiple outliers are present. This means that a 1987 market crash dummy is incapable of accommodating multiple extreme values in the data series. For instance, stock UTX (index = 10) has an extreme value of −38 % during the week of the 9–11 terrorist attacks in 2001. To address this issue, we identify the outliers and patch the outliers by using intervention analysis as in Box and Tiao (1975) and the extension by Tsay et al. (2000) and Peña et al. (2001). Table 80.5 reports the statistics of the residual analysis for these nine stocks by adding different dummies in the mean equation. This result is rather encouraging, as evidenced by the reduction of the significance of the kurtosis coefficient. It reveals that the kurtosis problem is somehow related to a failure to account for extraordinary events that disturb the data structure, rather than the failure of the EGB2 distribution. It is evident that after removing the effect of the outliers in a given time series, the EGB2 distribution is capable of addressing the financial data with skewness and kurtosis in an appropriate range.[18]

80.5 Distributional Fit Test

Previous sections emphasized estimates of parameters pertinent to modeling the skewness and kurtosis of the standardized residuals by applying non-Gaussian distributions. As part of the modeling process, model checking in terms of goodness of fit is also important. Table 80.6 and Fig. 80.3 compare a GARCH(1,1) model based on three distributions: normal, Student's *t*, and EGB2. The reported log-likelihood function values (negative) clearly show that the EGB2 distribution outperforms the

[17]Engle et al. (1987) suggest putting a conditional volatility variable in the mean equation, which is called the GARCH-M model. However, the expected sign of the conditional variance variable is uncertain, according to literature surveys. Since we do not find its statistical significance in our empirical experiment (not reported), the conditional mean term is excluded from our test equation.

[18]Longin (1996) proposes the use of a Frechet distribution, which is able to highlight those extreme price movements. However, his model does not cover whole return distributions but only extreme values.

Table 80.5 Statistics of the nine stocks' standardized errors on the GARCH(1,1)-EGB2 estimates: weekly data, 1986–2005

Index	Ticker	Skewness	Kurtosis	Peakedness	Q(30)	$Q^2(30)$	ϕ_0	ϕ_1	ϕ_2	w	α	β	γ	p	q	N
2	HON	0.1578	2.3953	1.2182	50.5229	30.1429	0.0006	1.0566		1.3E-05	0.0408	0.935	0.0211	0.9103	0.7813	10
		[−0.20]	[3.25]***	(1.19)	0.01**	0.46	[0.77]	[27.88]***	[NA]	[1.93]*	[2.20]**	[55.37]***	[0.80]			
9	MO	−0.5069	2.7862	1.1174	22.8401	27.9385	0.0027	0.8159		1.3E-05	0.0407	0.9407	0.0142	0.4568	0.5192	3
		[−2.07]**	[2.28]**	(1.12)	0.82	0.57	[2.88]***	[18.11]***	[NA]	[1.76]*	[2.24]**	[58.26]***	[0.64]			
10	UTX	−0.1411	1.2787	1.2052	32.3131	29.7857	0.0021	0.9843	−0.0853	1.1E-05	0.0414	0.9334	0.0221	0.8765	0.8795	2
		[−1.02]	[−0.19]	(1.20)	0.35	0.48	[2.64]***	[25.53]***	[−4.20]***	[1.65]*	[2.03]**	[42.05]***	[0.92]			
11	PG	−0.0746	1.5127	1.1577	44.1755	25.669	0.0021	0.7953	−0.1211	1.7E-05	0.0561	0.8984	0.0527	0.6161	0.6576	1
		[0.13]	[−0.66]	(1.16)	0.05*	0.69	[2.57]***	[22.15]***	[−4.74]***	[1.92]*	[2.20]**	[31.66]***	[1.45]			
12	CAT	0.2477	2.0579	1.1768	32.5068	19.5369	0.0022	1.0234		1.7E-05	0.0198	0.9527	0.0284	0.7285	0.6271	3
		[0.32]	[1.17]	(1.16)	0.34	0.93	[2.07]**	[21.42]***	[NA]	[1.50]	[1.07]	[47.31]***	[1.44]			
17	MRK	−0.0976	1.905	1.1634	34.2624	29.7861	0.002	0.9621	−0.0609	0.00002	0.0183	0.9364	0.0532	0.6352	0.6109	1
		[−1.09]	[0.43]	(1.15)	0.27	0.48	[2.13]**	[22.28]***	[−2.49]**	[2.34]**	[1.17]	[53.32]***	[2.28]**			
20	HPQ	0.0938	1.8507	1.2149	38.0405	51.1062	0.0019	1.2568	−0.0687	1.3E-05	0.0219	0.9602	0.0199	0.8271	0.7859	9
		[0.21]	[1.33]	(1.19)	0.15	0.01**	[1.71]*	[22.32]***	[−3.00]***	[1.44]	[1.71]*	[87.05]***	[0.96]			

29	AIG	0.0815	1.777	1.232	20.983	28.3133	0.0016	1.1323	−0.0779	2.1E-05	0.0434	0.9149	0.0283	1.2246	1.0228	2
		[−0.81]	[2.09]**	(1.22)	0.89	0.55	[2.08]**	[30.66]***	[−3.69]***	[1.88]*	[2.31]**	[31.32]***	[0.95]			
30	C	0.2551	1.3999	1.1966	30.3054	29.5047	0.0024	1.4095	−0.0761	1.1E-05	0.0374	0.9496	0.002	1.0202	0.7612	3
		[−0.65]	[−0.22]	(1.19)	0.45	0.49	[2.43]**	[32.79]***	[−3.62]***	[1.18]	[2.24]**	[50.63]***	[0.09]			

The nine stocks have significant excess coefficients in Table 80.4. Numbers below coefficients are t-values (with bracket). Numbers below tests are *p*-values. ***Indicates 1 % significance, **5 %, and *10 %. The predicted higher moments are given by the formula: Skewness $= \frac{\psi''(p)-\psi''(q)}{(\psi'(p)+\psi'(q))^{1.5}}$, Kurtosis $= \frac{\psi'''(p)+\psi'''(q)}{(\psi'(p)+\psi'(q))^{2}}$. Their standard deviations are obtained using the delta method. The peakedness is measured by $f_{0.75}$-$f_{0.25}$, the distance between the values of standardized variable at which the cumulative distribution function equals 0.75 and the value at which the cumulative distribution function equals 0.25. The reference value of the standard normal distribution is 1.35. A number of peakedness less than 1.35 means there is a high peak in the probability density function. The reference value of the EGB2 distribution is reported below peakedness in parenthesis and is in the range of (1.12, 1.22). The normality test is omitted, since the assumption is the EGB2 distribution. The independence test is conducted using a Ljung-Box Q test up to order 30. The Q^2 test of order of 30 is to show volatility clustering. The model is

(1.a) $R_t = \phi_0 + \phi_1 R_{m,t} + \phi_2 R_{t-1} + \Delta_n D_{extreme} + \varepsilon_{it}$

(1.b) $\varepsilon_t = \sqrt{h_t} z_t$

(1.c) $h_t = w + \alpha\varepsilon_{t-1}^2 + \beta h_{t-1} + \gamma\, I(\varepsilon_{t-1} < 0)\varepsilon_{t-1}^2$

(1.d) $\varepsilon_t|\Im_{t-1} \sim EGB2(0, h_t, p, q)$. N in the table represents the number of dummies in the mean equation (at most 10). Those dummies represent the extreme values in the individual stocks' return series. The following stocks do not have an AR(1) variable: MSFT, HON, DD, GM, IBM, MO, CAT, BA, PFE, AA, DIS, MCD, JPM, and INTC. Stock PG is the only one to have an AR(4) variable in the mean equation, which ensures that autocorrelation has been removed

Table 80.6 Fitness comparisons among alternative distributions

		Likelihood (-lnL)			Chi-square statistics		
Index	Ticker	Normal	t	EGB2	Normal	t	EGB2
1	MSFT	2,727.663	2,405.698	1,836.975	56.32**	70***	38.6
2	HON	2,939.66	2,655.106	2,078.661	77.76***	129.92***	60.12***
3	KO	3,070.488	2,747.057	2,175.794	45.6	89.76***	30.2
4	DD	3,080.986	2,748.706	2,180.607	50.32	70.88***	41.04
5	XOM	3,286.659	2,949.103	2,378.235	46.24	49.6*	38
6	GE	3,330.797	2,998.795	2,428.805	33.36	54.72**	23.32
7	GM	2,851.589	2,529.761	1,958.933	40.56	78.32***	36.4
8	IBM	2,940.611	2,646.765	2,076.628	87.84***	151.04***	128.68***
9	MO	2,870.27	2,589.11	2,016.484	76.56***	138.72***	49*
10	UTX	3,086.131	2,777.7	2,204.927	56.24**	86.4***	42.64
11	PG	3,099.136	2,798.107	2,226.256	77.6***	106.4***	47.92
12	CAT	2,808.847	2,515.001	1,941.912	71.04***	110.88***	50.52*
13	BA	2,869.513	2,546.216	1,974.935	54.72**	77.36***	44.12
14	PFE	2,906.508	2,585.459	2,014.814	66***	84.08***	31.2
15	JNJ	3,094.648	2,764.042	2,194.88	68.72***	86.72***	53**
16	MMM	3,213.133	2,899.268	2,330.127	69.92***	99.68***	54.88**
17	MRK	2,921.886	2,622.322	2,049.385	54.88**	91.52***	54.68**
18	AA	2,820.431	2,491.263	1,921.461	43.92	69.84***	34
19	DIS	2,947.914	2,628.138	2,057.377	45.84	76.56***	47.64
20	HPQ	2,640.579	2,341.888	1,768.097	68.56***	117.92***	67.28***
21	MCD	3,021.621	2,694.789	2,125.111	65.36***	117.12***	45.52
22	JPM	2,851.61	2,527.93	1,957.228	54.72**	78.32***	48.72*
23	WMT	2,993.255	2,660.808	2,091.682	48.64	64.48***	36.68
24	AXP	3,013.207	2,681.583	2,111.124	42.32	72.8***	40.72
25	INTC	2,560.727	2,232.273	1,663.033	58.8**	63.04***	26.32
26	VZ	3,055.394	2,733.302	2,162.317	64.4***	106.32***	50.28*
27	T	3,023.173	2,700.77	2,129.957	51.84*	85.84***	45.16
28	HD	2,838.915	2,515.001	1,943.85	50.24	77.92***	42.24
29	AIG	3,097.629	2,777.491	2,206.573	56**	63.28***	39.72
30	C	2,918.77	2,633.718	2,060.1	75.44***	91.04***	39.76

This table compares the GARCH(1,1) model based on three distributions: normal, Student's *t*(T), and EGB2 based on a negative logarithm of the likelihood function value (Left) and the χ^2 goodness-of-fit test statistic value (Right). The quantiles are computed via 40 intervals. The degree of freedom (d.f.) is 37 for the EGB2, 38 for the *t* distribution and 39 for the normal distribution. The chi-square critical values at the 1 %, 5 %, and 10 % levels are 59.89, 52.19, and 48.36, respectively, with d.f. being 37; 61.16, 53.39, and 49.51 with d.f. being 38; and 62.43, 54.57, and 50.66 with d.f. being 39. ***Indicates 1 % significance, **5 %, and *10 %

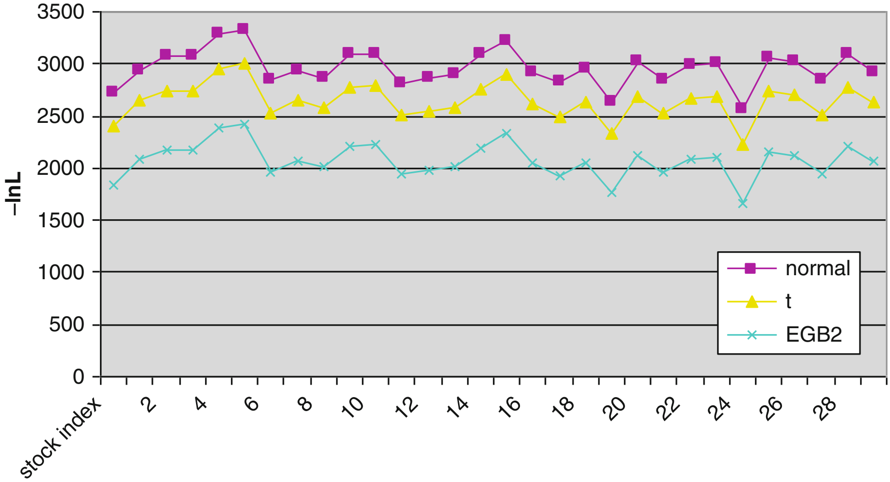


Fig. 80.3 Comparisons of log-likelihood function values in different models. The greater the function value the better of the fit the model is. The figure plots the negative logarithm value of the likelihood function

rival distributions: the normal distribution and the t distribution. However, as noted by Boothe and Glassman (1987), making non-nested distribution comparisons based on log-likelihood values can lead to spurious conclusions.[19] Consequently, we calculate the goodness-of-fit (GoF) statistics[20] to compare differences between observed distribution of standardized residuals and theoretical distribution based on estimated shape parameters following Snedecor and Cochran (1989).

The null hypothesis tested by the GoF statistics is that the observed and predicted distribution functions are identical. The statistic is calculated by

$$GoF = \sum_{i=1}^{k} \frac{(f_i - F_i)^2}{F_i} \tag{80.8}$$

where f_i is the observed count of actual standardized residuals in the i th data class (interval), F_i is the predicted count derived from the estimated values for the distribution parameters, and k is the number of data intervals used in distributional comparisons. GoF has an asymptotic chi-square distribution with degrees of freedom equal to the number of intervals minus the number of estimated distribution

[19] Normal distribution is a special case of the EGB2 distribution. A likelihood ratio rest suggests that there is significant improvement in the fit of the EGB2 distribution over that of the normal distribution.

[20] The chi-square test is an alternative to the Anderson-Darling and Kolmogorov-Smirnov goodness-of-fit tests. The chi-square test and Anderson-Darling test make use of the specific distribution in calculating critical values. This has the advantage of allowing a more sensitive test and the disadvantage that critical values must be calculated for each distribution.

parameters minus one. For EGB2 distribution, two parameters are estimated; for Student's t distribution, one parameter is estimated; for the normal distribution, no parameter is required, since the error term has been standardized.

Table 80.6 reports the results of the χ^2 test for three distributions used in the GARCH(1,1) model. The test power is maximized by choosing a data class equiprobably (equal probability). The rule of thumb of the chi-square test is to choose the number of groups starting at 2 $T^{0.4}$.[21] The test results show that the null hypothesis is rejected by 12 stocks on the normal distribution at the 1 % level, 28 stocks on the t distribution and only three stocks on the EGB2 distribution. Furthermore, the χ^2 statistic also shows that the EGB2 distribution yields lower absolute values. We can conclude that the model based on the EGB2 distribution has the least deviation of the residuals from the theoretical distribution. The evidence suggests that the Student's t distribution is able to solve the kurtosis problem, but it could not fit the whole error distributions due to peakedness. Putting the evidence together, it is clear that the EGB2 distribution is superior to the t distribution and the normal distribution in our empirical analysis.

80.6 Implication of EGB2 Distribution

One of the main objectives of analyzing financial data for risk management purposes is to provide an answer to the question: how should we evaluate the probability of the extreme values by using statistical distributions? According to the normal distribution, the 1987 market crash with more than -17σ (daily data) would have never happened. However, recent market crashes indicate that big market swings or significant declines in asset prices happen more frequently than we expect. Although VaR is one of the most prevalent risk measures under normal conditions, it cannot deal with those extreme values, since extreme values are *not* normal. From this perspective, the EGB2 distribution provides a management tool for calculating risk.

Table 80.7 reports the probability of the semivolatility of shocks. Here, we concentrate on the probability of the error term having negative shocks. From this table, we see that the predicted probability for extreme values (beyond -2σ) is greater than that of the normal distribution. For instance, probabilities of a -5σ and -7σ shock for MSFT (index = 1) are 4.9E-5 and 8.4E-7, much greater than 2.8E-7 and 1.3E-12 based on the normal distribution.

Yet, the probabilities for the EGB2 distribution under a moderate range (within $\pm 2\sigma$) are less than that of the normal distribution. This is an alternative way to determine the peakedness and fat tails of portfolio returns. Notice that the crossing point between the EGB2 distribution and the normal distribution is in

[21] Our sample contains 999 observations; 40 intervals are used. Each group (data class) has theoretically 25 observations. The degrees of freedom are 37, 38, and 39 for the EGB2 distribution, the t distribution and the normal distribution, respectively. (The chi-squared critical values are given in the note in Table 80.6.)

Table 80.7 The probability of negative extreme shocks in the error term

Stock		Shocks						
		−7σ	−6σ	−5σ	−4σ	−3σ	−2σ	−1σ
1	MSFT	8.41E-07	6.42E-06	4.90E-05	0.000373	0.002832	0.021062	0.137138
2	HON	7.31E-06	3.76E-05	0.000193	0.000984	0.005031	0.025703	0.12933
3	KO	3.42E-06	2.03E-05	0.00012	0.000708	0.004177	0.024489	0.135087
4	DD	2.83E-06	1.73E-05	0.000105	0.000642	0.003902	0.023582	0.134136
5	XOM	1.05E-06	8.00E-06	6.07E-05	0.000459	0.003429	0.024407	0.143861
6	GE	4.52E-07	3.96E-06	3.47E-05	0.000303	0.002615	0.021451	0.141932
7	GM	3.30E-06	1.96E-05	0.000117	0.000693	0.004112	0.024242	0.134684
8	IBM	1.83E-05	8.05E-05	0.000352	0.001534	0.006686	0.029148	0.126866
9	MO	2.35E-05	9.97E-05	0.00042	0.001761	0.007391	0.031013	0.129726
10	UTX	1.17E-05	5.64E-05	0.000269	0.001287	0.006139	0.029219	0.135227
11	PG	1.53E-05	7.00E-05	0.000318	0.00144	0.006526	0.029553	0.132183
12	CAT	5.00E-06	2.74E-05	0.00015	0.000818	0.004465	0.024337	0.129658
13	BA	3.28E-06	1.97E-05	0.000118	0.000708	0.004233	0.025055	0.137264
14	PFE	9.08E-06	4.58E-05	0.00023	0.001154	0.00578	0.028808	0.137298
15	JNJ	1.90E-06	1.25E-05	8.18E-05	0.000536	0.003512	0.022751	0.135477
16	MMM	1.07E-05	5.16E-05	0.000248	0.001193	0.005729	0.027508	0.130414
17	MRK	9.13E-06	4.53E-05	0.000224	0.001105	0.005452	0.026874	0.130475
18	AA	2.66E-06	1.66E-05	0.000103	0.00064	0.00397	0.024331	0.137113
19	DIS	4.92E-06	2.74E-05	0.000152	0.000842	0.004664	0.025708	0.134949
20	HPQ	7.93E-06	4.05E-05	0.000206	0.001045	0.005301	0.02685	0.132606
21	MCD	2.13E-06	1.36E-05	8.67E-05	0.000552	0.003515	0.022249	0.132611
22	JPM	5.72E-06	3.12E-05	0.00017	0.00092	0.004991	0.026883	0.136911
23	WMT	1.21E-06	8.74E-06	6.28E-05	0.000451	0.003228	0.022585	0.138697
24	AXP	1.09E-06	8.05E-06	5.92E-05	0.000435	0.003184	0.022668	0.139658
25	INTC	1.14E-05	5.56E-05	0.000269	0.0013	0.006275	0.030129	0.138371
26	VZ	3.65E-06	2.13E-05	0.000124	0.00072	0.004185	0.024206	0.133382
27	T	6.62E-06	3.50E-05	0.000184	0.000964	0.005062	0.026487	0.133843
28	HD	6.65E-06	3.56E-05	0.00019	0.00101	0.005372	0.028274	0.139248
29	AIG	1.35E-06	9.52E-06	6.69E-05	0.000469	0.003282	0.022544	0.137761
30	C	1.91E-06	1.24E-05	7.97E-05	0.000514	0.003315	0.021299	0.13028
If normal		1.28E-12	9.87E-10	2.87E-07	3.17E-05	0.00135	0.02275	0.158655

The probability is calculated based on estimated p- and q-values of the EGB2 distribution. It tells how often the error terms have negative extreme values. The probability values based on the normal distribution are the same for all 30 stocks

the neighborhood of $\pm 2\sigma$, where the probabilities of both distributions are about the same value. This feature implies that VaR at the 95 % confidence level based on the normal distribution is by chance consistent with reality. However, beyond this critical level, the VaR method based on the normal distribution leads an underestimation in forecasts of losses. Nevertheless, the EGB2 distribution in this regard provides a broader spectrum of risk information for guiding risk management.

80.7 Conclusions

In this chapter, we present empirical evidence on the stock return equation based on market risk, time series pattern, and asymmetric conditional variance for the 30 Dow Jones stocks. Special attention is placed on the issue of presenting skewness, kurtosis, and outlier effects. Although we find no significant difference over the estimated betas and the corresponding standard errors of the distributions, the evidence shows that the exponential generalized beta distribution of the second kind (EGB2) is superior to the Student's *t* distribution and the normal distribution in dealing with data that demonstrate skewness and excess kurtosis simultaneously. The superiority of the EBG2 distribution in modeling financial data is not only due to its flexibility but also to its closed-form density function for the distribution. Its higher-order moments are finite and explicitly expressed by its parameters. Thus, the EGB2 model provides a useful tool for forecasting variances involving extreme values. As a result, this model can have practical use for risk management.

Consistent with the finding in the literature, the asymmetric effects are highly significant in the standard GJR-GARCH specification by assuming normal distributions. However, by incorporating the heavy tail information into the distributions, we can reduce the asymmetric effects. Our study confirms that the EBG2 distribution has the capacity to deal with the asymmetric effects. Since excess kurtosis is often caused by outliers, our finding suggests that removing the contamination of outliers from the residuals enhances the performance of EGB2 distributions. In short, the GJR-GARCH-type model based on the EGB2 distribution provides a richer framework for modeling stock return volatility. It accommodates several special stock return features, including fat tails, skewness, peakedness, autocorrelation, volatility clustering, and leverage effect. As a result, this model is effective for empirical estimation and suitable for risk management.

Appendix 1

Delta Method and Standard Errors of the Skewness and Kurtosis Coefficients of the EGB2 Distribution

The delta method, in its essence, expands a function of a random variable about its mean, usually with a one-step Taylor approximation, and then takes the variance. For example, if we want to approximate the variance of $G(x)$ where x is a random variable with mean μ and $G(x)$ is differentiable, we can try

$$G(x) \approx G(\mu) + (x - \mu)G'(\mu) \tag{80.9}$$

so that

$$\mathrm{Var}(G(x)) \approx G'(\mu)^2 \mathrm{Var}(x) \tag{80.10}$$

where $G'() = dG/dX$. This is a good approximation only if x has a high probability of being close enough to its mean so that the Taylor approximation is still good.

The nth central moments of the EGB2 distribution is given by

$$\text{The nth moment} = \sigma^n \left(\psi^{n-1}(p) + (-1)^n \psi^{n-1}(q)\right) \tag{80.11}$$

where ψ^n is nth order a polygamma function. Correspondingly, the skewness coefficient is given by

$$\text{Skewness} = g(p,q) = \frac{\psi''(p) - \psi''(q)}{(\psi'(p) + \psi'(q))^{1.5}} \tag{80.12}$$

The variance of the skewness coefficient by the delta method is given by

$$\begin{aligned}\mathrm{Var}(\text{Skewness}) = {} & g_p^{'}(p,q)^2 \mathrm{var}(p) + g_q^{'}(p,q)^2 \mathrm{var}(q) \\ & + 2g_p^{'}(p,q) g_q^{'}(p,q) \mathrm{cov}(p,q)\end{aligned} \tag{80.13}$$

where

$$g_p^{'}(p,q) = \frac{\psi'''(p)(\psi'(p) + \psi'(q)) - 1.5\psi''(p)(\psi''(p) - \psi''(q))}{(\psi'(p) + \psi'(q))^{2.5}} \tag{80.14}$$

$$g_q^{'}(p,q) = \frac{-\psi'''(q)(\psi'(p) + \psi'(q)) - 1.5\psi''(q)(\psi''(p) - \psi''(q))}{(\psi'(p) + \psi'(q))^{2.5}} \tag{80.15}$$

Similarly, the excess kurtosis coefficient is given by

$$\text{Kurtosis} = h(p,q) = \frac{\psi'''(p) + \psi'''(q)}{(\psi'(p) + \psi'(q))^2} \tag{80.16}$$

The variance of the kurtosis coefficient by the delta method is given by

$$\begin{aligned}\mathrm{Var}(\text{Kurtosis}) = {} & h_p^{'}(p,q)^2 \mathrm{var}(p) + h_q^{'}(p,q)^2 \mathrm{var}(q) \\ & + 2h_p^{'}(p,q) h_q^{'}(p,q) \mathrm{cov}(p,q)\end{aligned} \tag{80.17}$$

where

$$h_p'(p,q) = \frac{\psi''''(p)(\psi'(p)+\psi'(q)) - 2\psi''(p)(\psi'''(p)+\psi'''(q))}{(\psi'(p)+\psi'(q))^3} \tag{80.18}$$

$$h_q'(p,q) = \frac{\psi''''(q)(\psi'(p)+\psi'(q)) - 2\psi''(q)(\psi'''(p)+\psi'''(q))}{(\psi'(p)+\psi'(q))^3} \tag{80.19}$$

In the equations above, high-order polygamma functions are involved. A polygamma function is the nth normal derivative of the logarithmic derivative of $\Gamma(z)$:

$$\psi^n(z) = \frac{d^{n+1}}{dz^{n+1}} \ln(\Gamma(z)) \tag{80.20}$$

which, for $n > 0$, can be written as

$$\psi^n(z) = (-1)^{n+1} n! \sum_{k=0}^{\infty} \frac{1}{(z+k)^{n+1}} \tag{80.21}$$

which is used to calculate polygamma functions in this chapter.

Note: This appendix is based on the paper by Wang et al. (2001). However, there are typos and errors in that paper. The standard deviation formula for skewness used by Wang et al. (2001) is incorrect (see g′q(p,q) equation in the paper at http://www.econ.queensu.ca/jae/2001-v16.4/wang-fawson-barrett-mcdonald/Appendix4_delta_derivations.pdf). We therefore provide this appendix. Accordingly, as we reviewed and replicated that paper by using the data supplied by the *Journal of Applied Econometrics*, the EGB2 distribution doesn't remove the skewness problem completely as shown in their Table 3. In addition, there is a computational error in the JPY series, so that its kurtosis has not been resolved either.

References

Akgiray, V. (1989). Conditional heteroskedasticity in time series of stock returns: Evidence and forecasts. *Journal of Business, 62*, 55–80.

Amihud, Y., & Mendelson, H. (1987). Trading mechanism and stock returns: An empirical investigation. *Journal of Finance, 42*, 533–553.

Andersen, J. V., & Sornette, D. (2001). Have your cake and eat it too: Increasing returns while lowering large risks! *Journal of Risk Finance, 2*, 70–82.

Antoniou, A., Koutmos, G., & Pericli, A. (2005). Index futures and positive feedback trading: Evidence from major stock exchanges. *Journal of Empirical Finance, 12*, 219–238.

Avramov, D. (2002). Stock return predictability and model uncertainty. *Journal of Financial Economics, 64*, 423–458.

Baillie, R. T., & DeGennaro, R. P. (1990). Stock returns and volatility. *Journal of Financial and Quantitative Analysis, 25*, 203–215.
Bali, T. G., Mo, H., & Tang, Y. (2008). The role of autoregressive conditional skewness and kurtosis in the estimation of conditional VaR. *Journal of Banking & Finance, 32*, 269–282.
Barndorff-Nielsen, O., Kent, J., & Sorensen, M. (1982). Normal variance-mean mixtures and z distributions. *International Statistical Review, 50*, 145–159.
Bollerslev, T. (1987). A conditionally heteroskedastic time series model for speculative prices and rates of return. *Review of Economics and Statistics, 69*, 542–547.
Bollerslev, T., Chou, R. Y., & Kroner, K. F. (1992). ARCH modeling in finance: A review of the theory and empirical evidence. *Journal of Econometrics, 52*, 5–59.
Bollerslev, T. R., Engle, F., & Nelson, D. B. (1994). ARCH model. In Engle & McFadden (Eds.), *Handbook of econometrics* (Vol. IV). Amsterdam: North Holland.
Boothe, P., & Glassman, D. (1987). The statistical distribution of exchange rates. *Journal of International Economics, 22*, 297–319.
Box, G. E. P., & Tiao, G. C. (1975). Intervention analysis with applications to economic and environmental problem. *Journal of the American Statistical Association, 70*, 70–79.
Boyer, B., Mitton, T., & Vorkink, K. (2010). Expected idiosyncratic skewness. *Review of Financial Studies, 23*, 169–202.
Brooks, C., Burke, S., Heravi, S., & Persand, G. (2005). Autoregressive conditional kurtosis. *Journal of Financial Econometrics, 3*, 399–421.
Brown, S., & Warner, J. (1985). Using daily stock returns: The case of event studies. *Journal of Financial Economics, 14*, 3–31.
Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica, 41*, 135–155.
Conrad, J. S., Dittmar, R. F., & Ghysels, E. (2009). Ex ante skewness and expected stock returns. Available at SSRN: http://ssrn.com/abstract=1522218 or http://dx.doi.org/10.2139/ssrn.1522218 (Dec 11).
Cvitanic, J., Polimenis, V., & Zapatero, F. (2008). Optimal portfolio allocation with higher moments. *Annals Finance*, *4*, 1–28.
Engle, R., Lilien, D., & Robins, R. (1987). Estimating time-varying risk premia in the term structure: The ARCH-M model. *Econometrica, 55*, 391–408.
Fama, E. (1965). The behavior of stock-market prices. *The Journal of Business, 38*, 34–105.
Fama, E., & French, K. (1996). Multifactor explanations of asset pricing anomalies. *Journal of Finance, 51*, 55–84.
French, K., Schwert, G. W., & Stambaugh, R. (1987). Expected stock returns and volatility. *Journal of Financial Economics, 19*, 3–29.
Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the normal excess return on stocks. *Journal of Finance, 48*, 1779–1801.
Hansen, B. (1994). Autoregressive conditional density estimation. *International Economic Review, 35*, 705–730.
Hansen, C. B., McDonald, J. B., & Theodossiou, P. (2006). *Some flexible parametric models for partially adaptive estimators of econometric models* (Working Paper). Brigham Young University.
Harvey, C. R., & Siddique, A. (1999). Autoregressive conditional skewness. *Journal of Financial and Quantitative Analysis, 34*, 465–487.
Harvey, C. R., & Siddique, A. (2000). Conditional skewness in asset pricing tests. *Journal of Finance, 55*, 1263–1295.
Harvey, C. R., Liechty, J. C., Liechty, M. W., & Muller, P. (2003). Portfolio selection with higher moments. *Quantitative Finance*, *10*(5), 469–485.
Hogg, R. V., & Klugman, S. A. (1983). On the estimation of long tailed skewed distributions with actuarial applications. *Journal of Econometrics, 23*, 91–102.
Hosking, J. R. M. (1994). The four-parameter kappa distribution. *IBM Journal of Research and Development, 38*, 251–258.

Hueng, C. J., & Brooks, R. (2003). Forecasting asymmetries in aggregate stock market returns: Evidence from higher moments and conditional densities (Working Paper, WP02-08-01). The University of Alabama. http://scholar.google.com.hk/citations?view_op=view_citation&hl=en&user=Rw22qOoAAAAJ&citation_for_view=Rw22qOoAAAAJ:IjCSPb-OGe4C
Hueng, C. J., & McDonald, J. (2005). Forecasting asymmetries in aggregate stock market returns: Evidence from conditional skewness. *Journal of Empirical Finance, 12*, 666–685.
Johnson, N. L. (1949). Systems of frequency curves generated by methods of translation. *Biometrica, 36*, 149–176.
Johnson, H., & Schill, M. J. (2006). Asset pricing when returns are nonnormal: Fama-French factors versus higher order systematic comoments. *Journal of Business, 79*, 923–940.
Johnson, N. L., Kotz, S., & Balakrishnan, N. (1994). *Continuous univariate distributions* (2nd ed., Vol. 1). New York: Wiley.
Jurczenko, E., & Maillet, B. (2002). *The four moment capital asset pricing model: Some basic results* (Working Paper). Paris: Ecole supérieure de commerce de Paris.
Kenney, J. F., & Keeping, E. S. (1962). Skewness. §7.10 In *Mathematics of statistics* (pp. 100–101). Princeton: Van Nostrand.
Liu, S. M., & Brorsen, B. (1995). Maximum likelihood estimation of a GARCH – STABLE model. *Journal of Applied Econometrics, 10*, 273–285.
Lo, A., & MacKinlay, A. C. (1990). Data-snooping biases in tests of financial asset pricing models. *Review of Financial Studies, 3*, 431–467.
Longin, F. (1996). The asymptotic distribution of extreme stock market returns. *Journal of Business, 63*, 383–408.
Malevergne, Y., & Sornette, D. (2005). Higher moment portfolio theory. *The Journal of Portfolio Management*, *31*(4), 49–55.
Mandelbrot, B. B. (1963). The variation of certain speculative prices. *Journal of Business, 36*, 394–419.
McCulloch, J. H. (1985). Interest-risk sensitive deposit insurance premia: Stable ACH estimates. *Journal of Banking and Finance, 9*, 137–156.
McDonald, J. B. (1984). Some generalized functions for the size distribution of income. *Econometrica, 52*, 647–663.
McDonald, J. B. (1991). Parametric models for partially adaptive estimation with skewed and leptokurtic residuals. *Economics Letters, 37*, 273–278.
McDonald, J. B., & Newey, W. K. (1988). Partially adaptive estimation of regression models via the generalized *t* distribution. *Econometric Theory, 4*, 428–457.
McDonald, J. B., & Xu, Y. J. (1995). A generalization of the beta distribution with applications. *Journal of Econometrics, 66*, 133–152.
Mittnik, S., Rachev, S., Doganoglu, T., & Chenyao, D. (1999). Maximum likelihood estimation of stable paretian models. *Mathematical and Computer Modelling, 29*, 275–293.
Nelson, D. B. (1991). Conditional heteroskedasticity in asset return: A new approach. *Econometrica, 59*, 347–370.
Officer, R. (1972). The distribution of stock returns. *Journal of the American Statistical Association, 67*, 807–812.
Patton, A. J. (2004). On the out-of-sample importance of skewness and asymmetric dependence for ass *et al* location. *Journal of Financial Econometrics, 2*, 130–168.
Pena, D., Tiao, G. C., & Tsay, R. S. (2001). *A course in time series analysis*. New York: Wiley.
Ranaldo, A., & Favre, L. (2005). How to price hedge funds: From two- to four-moment CAPM (UBS Research Paper). SSRN: http://ssrn.com/abstract=474561.
Sentana, E., & Wadhwani, S. (1992). Feedback traders and stock return autocorrelations: Evidence from a century of daily data. *Economic Journal, 102*, 415–425.
Snedecor, G. W., & Cochran, W. G. (1989). *Statistical methods*. Ames: Iowa State Press.
Tang, G. Y. N. (1998). Monthly pattern and portfolio effect on higher moments of stock returns: Empirical evidence from Hong Kong. *Asia-Pacific Financial Markets, 5*, 275–307.

Theodossiou, P. (1998). Financial data and the skewed generalized *t* distribution. *Management Science, 44*, 1650–1661.

Theodossiou, P., McDonald, J. B., & Hansen, C. B. (2007). Some flexible parametric models for partially adaptive estimators of econometric models. *Economics: The Open-Access, Open-Assessment E-Journal*, *1*(2007-7), 1–20. http://dx.doi.org/10.5018/economics-ejournal.ja.2007-7.

Tsay, R. C., Peña, D., & Pankratz, A. (2000). Outliers in multivariate time series. *Biometrica, 87*, 789–804.

Vinod, H. D. (2004). Ranking mutual funds using unconventional utility theory and stochastic dominance. *Journal of Empirical Finance, 11*, 353–377.

Wang, K. L., Fawson, C., Barrett, C. B., & McDonald, J. B. (2001). A flexible parametric GARCH model with an application to exchange rates. *Journal of Applied Econometrics, 16*, 521–536.

Wu, J. W., Hung, W. L., & Lee, H. M. (2000). Some moments and limit behaviors of the generalized logistic distribution with applications. *Proceedings of the National Science Council ROC, 24*, 7–14.

Does Revenue Momentum Drive or Ride Earnings or Price Momentum?

81

Hong-Yi Chen, Sheng-Syan Chen, Chin-Wen Hsin, and Cheng-Few Lee

Contents

This article is a reprint of the article entitled "Does revenue momentum drive or ride earnings or price momentum?" published in the Journal of Banking & Finance (Hong-Yi Chen, Sheng-Syan Chen, Chin-Wen Hsin, and Cheng-Few Lee, Vol. 38, 2014, pp. 166-185). We would like to thank the editor of the Journal of Banking & Finance and Elsevier for permission to reprint the paper.

H.-Y. Chen (✉)
Department of Finance, National Central University, Taoyuan, Taiwan
e-mail: fnhchen@ncu.edu.tw

S.-S. Chen
National Central University, Zhongli City, Taiwan
e-mail: fnschen@management.ntu.edu.tw

C.-W. Hsin
Yuan Ze University, Zhongli City, Taiwan
e-mail: fncwhsin@saturn.yzu.edu.tw

C.-F. Lee
Department of Finance and Economics, Rutgers Business School, Rutgers, The State University of New Jersey, Piscataway, NJ, USA

Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
e-mail: lee@business.rutgers.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_81,

Abstract

This study examines the profits of revenue, earnings, and price momentum strategies in an attempt to understand investor reactions when facing multiple information of firm performance in various scenarios. We first offer evidence that there is no dominating momentum strategy among the revenue, earnings, and price momentums, suggesting that revenue surprises, earnings surprises, and prior returns each carry some exclusive unpriced information content. We next show that the profits of momentum driven by firm fundamental performance information (revenue or earnings) depend upon the accompanying firm market performance information (price), and vice versa. The robust monotonicity in multivariate momentum returns is consistent with the argument that the market does not only underestimate the individual information but also the joint implications of multiple information on firm performance, particularly when they point in the same direction. A three-way combined momentum strategy may offer monthly return as high as 1.44 %. The information conveyed by revenue surprises and earnings surprises combined account for about 19 % of price momentum effects, which finding adds to the large literature on tracing the sources of price momentum.

Keywords

Revenue surprises • Earnings surprises • Post-earnings-announcement drift • Momentum strategies

81.1 Introduction

Financial economists have long been puzzled by two robust and persistent anomalies in the stock market: price momentum (see Jegadeesh and Titman 1993, 2001; Rouwenhorst 1998), and post-earnings-announcement drift (see Ball and Brown 1968; Foster et al. 1984; Bernard and Thomas 1989; Chan et al. 1996). More recently, Jegadeesh and Livnat (2006b) also find that price reactions to revenue surprises on announcement dates only partially reflect the incremental information conveyed by the surprises. The information contents carried by revenue, earnings and stock prices are intrinsically linked through firm operations and investor evaluation, and there is evidence of mutual predictability for respective future values (e.g., see Jegadeesh and Livnat 2006b). Nonetheless, investors, aware of the linkages among the information content conveyed by revenue, earnings and prices (see Ertimur et al. 2003; Raedy et al. 2006; and Heston and Sadka 2008), may still fail to take full account of their joint implications when pricing the stocks.

This study investigates how investors price securities when facing multiple information contents of a firm, particularly those firm performance information that are most accessible for investors – price, earnings, and revenue.[1] The long-short strategy of momentums, widely used in the literature, provides a venue to detect market reactions toward individual and multiple information contents. Accordingly, this study will start with documenting the revenue momentum profits and re-confirming the earnings and price momentums profits. Explorations with momentum strategies expect to yield implications that answer our two research questions. First, among the performance information of revenue surprises, earnings surprises, and prior returns, does each carry some exclusive information content that is not priced by the market? Second, do investors mis-react toward the joint implications as well as individual information of firm revenue, earnings, and price?

Our first research question is explored by testing momentum dominance. One momentum strategy is said to be dominated if its payoffs can be fully captured by the information measure serving as the sorting criterion of another momentum strategy. Note that our emphasis here is not asset pricing tests; instead, as in Chan et al. (1996) and Heston and Sadka (2008), we focus on the return anomalies based on revenue surprises, earnings surprises, and prior returns. Results from both a pairwise nested comparison and a regression analysis indicate that revenue surprises, earnings surprises, and prior returns each lead to significant momentum returns that cannot be explained away by one another. That is, revenue momentum neither drives nor rides earnings or price momentum. Following the information diffusion hypothesis of Hong and Stein (1999), our evidence then suggests that revenue surprises, earnings surprises, and prior returns each contribute to the phenomenon of gradual information flow, or that each have some exclusive information content that is not priced by the market.[2] Further regression tests indicate

[1]Researches in the literature offer some evidence on the information linkage among revenue, earnings and prices. For example, Lee and Zumwalt (1981) find that revenue information is complementary to earnings information in security rate of return determination. Bagnoli et al. (2001) find that revenue surprises but not earnings surprises can explain stock prices both during and after the internet bubble. Swaminathan and Weintrop (1991) and Ertimur et al. (2003) suggest that the market reacts significantly more strongly to revenue surprises than to expenses surprises. Rees and Sivaramakrishnan (2001) and Jegadeesh and Livnat (2006b) also find that, conditional on earnings surprises, there is still a certain extent of market reaction to the information conveyed by revenue surprises. Ghosh et al. (2005) find that sustained increases in earnings are supported by sustained increases in revenues rather than by cost reductions.

[2]The asset pricing tests of Chordia and Shivakumar (2006) support that price momentum is subsumed by the systematic component of earnings momentum, even though they also find earnings surprises and past returns have independent explanatory power for future returns. This latter finding is consistent with the results of Chan et al. (1996) and our results, as is reported later. In comparison, Chan et al. (1996), Jegadeesh and Livnat (2006b), and we focus on whether and

that earnings surprise and revenue surprise information each accounts for about 14 % and 10 % of price momentum returns, and that these two fundamental performance information combined account for just about 19 % of price momentum effects. These results provide additional evidence in the literature on the sources of price momentum (e.g., see Moskowitz and Grinblatt 1999; Lee and Swaminathan 2000; Piotroski 2000; Grundy and Martin 2001; Chordia and Shivakumar 2002, 2005; Ahn et al. 2003; Griffin et al. 2003; Bulkley and Nawosah 2009; Chui et al. 2010; Novy-Marx 2012).

Our second research question inquires how the market reacts to the joint implications of multiple information measures. The three measures under our study all carry important messages on innovations in firm performance, and therefore expect to trigger investor reactions. They become ideal target to be studied to entail implications on how investors process multiple information interactively in pricing stocks. The results from two-way sorted portfolios find that the market anomalies vary monotonically with the joint condition of revenue surprises, earnings surprises, and prior returns, and anomalies tend to be strongest when stocks show the strongest signals in the same direction. The cross-contingencies of momentums are observed in that the momentum returns driven by *fundamental performance information* (revenue surprises or earnings surprises) change with the accompanying *market performance information* (prior returns), and vice versa. Such finding, as interpreted by the gradual-information-diffusion model, is consistent with the suggestion that the market not only underreacts to individual firm information but also underestimates the significance of the joint implications of revenue, earnings, and price information.[3] These results also have interesting implications for investment strategies that the fundamental performance information plays an important role in differentiating future returns among price winners, while the market performance information is particularly helpful in predicting future returns for stock with high surprises in revenue or earnings. Specifically, price winners, compared to price losers, yield higher returns from revenue/earnings momentum strategies; stock with greater surprises in fundamentals yield greater returns from price momentums.

The results of our dominance tests and multivariate momentum suggest that a combined momentum strategy should yield better results over single-criterion momentum strategies. A combined momentum strategy using all three performance measures is found to yield monthly returns as high as 1.44 %, which amounts to an annual return of 17.28 %. Such a combined momentum strategy outperforms single-criterion momentum strategies by at least 0.72 percentage points in monthly return.

how firm characteristics, such as revenue surprises, earnings surprises, and prior returns, are related to future cross-sectional returns, while Chordia and Shivakumar (2006) also conduct asset pricing tests.

[3]The firm performance measures, revenue, earnings, and stock price, do not only share common origins endogenously but also have added implications for future values of one another. Jegadeesh and Livnat (2006b) have documented evidence on the temporal linkages among these variables. In this study, we focus on the further inquiry that whether investors fully exploit such temporal linkages among these firm performance information in pricing stocks.

Our conclusions remain robust whether we use raw returns or risk-adjusted returns, whether we include January results or not, and whether we use dependent or independent sorts. Chan et al. (1996), Piotroski (2000), Griffin et al. (2005), Mohanram (2005), Sagi and Seasholes (2007), Asem (2009), and Asness et al. (2013) conduct similar tests on combined momentum strategies using alternative sorting criteria.[4] In comparison, our study is the first to document results considering these three firm performance information, revenue surprises, earnings surprises, and prior returns altogether.

In terms of persistency, the earnings momentum strategy is found to exhibit the strongest persistence, while the revenue momentum strategy is relatively short-lived. All the same, the short-lived revenue momentum effect is prolonged when the strategy is executed using stocks with the best prior price performance and more positive earnings surprises. In fact, the general conclusion supports our claim of cross-contingencies of momentum as applied to momentum persistence.

This study contributes to the finance literature in several respects. First, we specifically identify the profitability of revenue momentum and its relation with earnings surprises and prior returns in terms of momentum strength and persistence. A revenue momentum strategy executed with a 6-month formation period and 6-month holding-period strategy yields an average monthly return of 0.61 % for the period between 1974 and 2009. Second, this study identifies empirical interrelations of anomalies arising from three firm performance information – revenue, earnings and price. To the best of our knowledge, we are the first to offer evidence that there is no dominating momentum strategy among the three, and that the profits of momentum driven by firm fundamental performance information (revenue or earnings) depend upon the accompanying firm market performance information (price), and vice versa.[5] Third, aside from academic interest, the aforementioned findings may well serve as useful guidance for asset managers seeking profitable investment strategies. Fourth, this study also adds to the large literature attempting to trace the sources of price momentum. Our numbers indicate that the information conveyed by revenue surprises and earnings surprises combined account for about 19 % of price momentum effects. Last, our results offer additional evidence to the literature using the behavioral explanation for momentums.[6] Our empirical results

[4]Chan et al. (1996) and Griffin et al. (2005) find that when sorting prior price performance and earnings surprises together, the profits of a zero-investment portfolio are higher than those of single sorting. Piotroski (2000) and Mohanram (2005) develop fundamental indicators, FSCORE and GSCORE, to separate winners from losers. Sagi and Seasholes (2007) find that price momentum strategy becomes even more profitable when applied to stocks with high revenue growth volatility, low costs, or valuable growth options. Asness et al. (2013) find that the combination of value strategy and momentum strategy can perform better than either one alone. Asem (2009) find the momentum profits can be enhanced combining prior price returns and dividend behaviors.

[5]Heston and Sadka (2008) and Novy-Marx (2012) also provide evidence that earnings surprises are unable to explain price momentum. However, this study is the first to consider earnings surprises and revenue surprises at the same time in explaining price momentum.

[6]Barberis et al. (1998), Daniel et al. (1998), Hong and Stein (1999), Jackson and Johnson (2006), Verardo (2009), and Moskowitz et al. (2012) provide evidence in support of behavioral

are consistent with the suggestion that revenue surprises, earnings surprises, and prior returns each carry some exclusive unpriced information content. Moreover, the monotonicity of abnormal returns found in multivariate momentums suggests that the market does not only underestimate the individual information but also the joint implications of multiple information on firm performance. Such suggestion is new to the literature, and may also present a venue to track the sources of price momentum.

The study is organized as follows. In Sect. 81.2, we develop our models and describe the methodologies. In Sect. 81.3, we describe the data. In Sect. 81.4, we report the results on momentum strategies based on a single criterion. In Sect. 81.5, we discuss the empirical results of exploration of inter-relations among revenue, earnings, and price momentums using strategies built on multiple sorting criteria. In Sect. 81.6, we test the persistency and seasonality of momentum strategies. Section 81.7 concludes.

81.2 Revenue, Earnings, and Price Momentum Strategies

81.2.1 Measures for Earnings Surprises and Revenue Surprises

We follow Jegadeesh and Livnat (2006a, b) and measure revenue surprises and earnings surprises based on historical revenues and earnings.[7] Assuming that both quarterly revenue and quarterly earnings per share follow a seasonal random walk with a drift, we define the measure of revenue surprises for firm i in quarter t, standardized unexpected revenue growth (SURGE), as

$$SURGE_{i,t} = \frac{Q_{i,t}^{R} - E\left(Q_{i,t}^{R}\right)}{\sigma_{i,t}^{R}}, \tag{81.1}$$

where $Q_{i,t}^{R}$ is the quarterly revenue of firm i in quarter t, $E(Q_{i,t}^{R})$ is the expected quarterly revenue prior to earnings announcement, and $\sigma_{i,t}^{R}$ is the standard deviation of quarterly revenue growth.

The same method is applied to measure earnings surprises, specifically standardized unexpected earnings (SUE), defined as

$$SUE_{i,t} = \frac{Q_{i,t}^{E} - E\left(Q_{i,t}^{E}\right)}{\sigma_{i,t}^{E}}, \tag{81.2}$$

explanation to momentum effect, while Grundy and Martin (2001), Johnson (2002), Ahn et al. (2003), Sagi and Seasholes (2007), Li et al. (2008), Liu and Zhang (2008), and Wang and Wu (2011) attribute momentum effect to missing risk factors. In addition, Korajczyk and Sadka (2004) and Lesmond et al. (2004) re-examine the profitability of momentum strategies after taking the transaction cost into account and get mixed results.

[7]See Appendix for a detailed discussion of measures to estimate revenue and earnings surprises.

where $Q_{i,t}^{E}$ is the quarterly earnings per share from continuing operations, $E(Q_{i,t}^{E})$ is the expected quarterly earnings per share prior to earnings announcement, and $\sigma_{i,t}^{E}$ is the standard deviation of quarterly earnings growth.

81.2.2 Measuring the Profitability of Revenue, Earnings, and Price Momentum Strategies

We construct all three momentum strategies based on the approach suggested by Jegadeesh and Titman (1993). To evaluate the information effect of earnings surprises on stock returns, we form an earnings momentum strategy analogous to the one designed by Chordia and Shivakumar (2006). At the end of each month, we sort sample firms by SUE and then group the firms into ten deciles.[8] Dec 1 includes stocks with the most negative earnings surprises, and Dec 10 includes those with the most positive earnings surprises. The SUEs used in every formation month are obtained from the most recent earnings announcements, made within three months before the formation date.

We hold a zero-investment portfolio, long the most positive earnings surprises portfolio and short the most negative earnings surprises portfolio, for K ($K = 3$, 6, 9, and 12) subsequent months, not rebalancing the portfolios during the holding period. Such positive minus negative strategy (PMN) holds K different long-positive and short-negative portfolios each month. Accordingly, we obtain a series of zero-investment portfolio returns, which are the monthly returns to this earnings momentum strategy. Similarly, we apply this positive-minus-negative method to construct a revenue momentum strategy.

In the case of price momentum, we form a zero-investment portfolio each month by taking a long position in the top decile portfolio (winner) and a short position in the bottom decile portfolio (loser), and we hold this winner minus loser portfolio (WML) for subsequent K months. We thus obtain a series of zero-investment portfolio returns, i.e., the returns to the price momentum strategy.

81.3 Data and Sample Descriptions

81.3.1 Data

We collect from Compustat the firm basic information, earnings announcement dates, and firm accounting data. Stock prices, stock returns, share codes, and

[8]Note that we sort the sample firms into five quintile portfolios on each criterion in our later construction of multivariate momentum strategies. To conform to the same sorting break points, we also test the single momentum strategies based on quintile portfolios and find the results remain similar to those based on decile portfolios.

exchange codes come retrieved from the Center for Research in Security Prices (CRSP) files. The sample period is from 1974 through 2009. Only common stocks (SHRCD = 10, 11) and firms listed on New York Stock Exchange, American Stock Exchange, or Nasdaq (EXCE = 1, 2, 3, 31, 32, 33) are included in our sample. We exclude from the sample regulated industries (SIC = 4,000–4,999) and financial institutions (SIC = 6,000–6,999). We also exclude firms with stock prices below $5 on the formation date, considering that investors generally pay only limited attention to such stocks.

For the purpose of estimating their revenue surprises (SURGE), earnings surprises (SUE), and prior price performance, firms in the sample should have at least eight consecutive quarterly earnings announcements and six consecutive monthly returns before each formation month. To examine the return drift following the estimated SURGE, SUE, and prior price performance, firms in the sample need to have at least 12 consecutive monthly returns following each formation month. Firms in the sample should also have corresponding SURGE, SUE, size and book-to-market factors available in each formation month.

81.3.2 Sample Descriptions

Table 81.1 presents the summary statistics for firm size, estimates of revenue surprises and estimates of earnings surprises for our sample firms between year 1974 and year 2009. Panel A shows that there are 223,831 firm-quarters during the sample period. Median firm market capitalization is $235 million. Panel B and Panel C describe the distributions the revenue surprises (SURGE) and the earnings surprises (SUE) across firms of different market capitalization and different book-to-market ratio. Around 54 % of revenue surprises and 50 % of earnings surprises are positive.[9]

The values of SURGE and SUE are expected to be positively correlated. After all, a firm's income statement starts with revenue (sales) and ends with earnings; these two attributes share common firm operational information to a great extent, and their innovations, SURGE and SUE, should be correlated as well. Table 81.2 shows the time-series average of the cross-sectional correlations between 1974 and 2009. Panel A and Panel B present, respectively, the Pearson correlations and Spearman rank correlations. The average of both types of correlations between SURGE and SUE is 0.32, while prior price performance is not as significantly correlated with SURGE or SUE, with average correlations equal to about 0.15 and 0.19, respectively.

We then partition the sample by book-to-market ratio (B/M) and size. Value firms and small firms are found to exhibit slightly higher correlations among

[9]To ensure that firm accounting information is available to public investors at the time the stock returns are recorded, we follow the approach of Fama and French (1992) and match the accounting data for all fiscal years ending in calendar year $t-1$ with the returns for July of year t through June of $t+1$. The market capitalization is calculated by the closing price on the last trading day of June of a year times the number of outstanding shares at the end of June of that year.

Table 81.1 Summary statistics of sample firm characteristics

Panel A: sample size and firm market capitalization									
	Number of firm-quarters				Market cap (million dollars)				
					Mean	Median	Min	Max	
ALL	223,831				2,276	235	0.91	602,433	
Panel B. Descriptive statistics of SURGE									
	Positive SURGE				Negative SURGE				Zero SURGE
	N	Mean	Median	STD	N	Mean	Median	STD	N
ALL	121,525	3.31	2.84	2.34	102,306	−3.00	−2.56	2.21	0
Growth	45,670	3.63	3.25	2.40	27,829	−2.84	−2.35	2.21	0
Mid-BM	50,881	3.21	2.73	2.32	46,309	−3.05	−2.62	2.25	0
Value	24,974	2.91	2.41	2.20	28,168	−3.06	−2.69	2.15	0
Small	61,827	3.19	2.70	2.31	54,935	−2.96	−2.57	2.14	0
Mid-size	38,338	3.41	2.98	2.37	30,591	−3.02	−2.56	2.28	0
Large	21,360	3.45	2.99	2.40	16,780	−3.06	−2.56	2.33	0
Panel C. Descriptive statistics of SUE									
	Positive SUE				Negative SUE				Zero SUE
	N	Mean	Median	STD	N	Mean	Median	STD	N
ALL	112,068	2.42	1.89	1.94	111,330	−2.92	−2.11	2.59	433
Growth	37,928	2.47	1.98	1.92	35,407	−2.83	−2.10	2.43	164
Mid-BM	48,767	2.41	1.88	1.94	48,221	−2.92	−2.09	2.60	202
Value	25,373	2.37	1.79	1.95	27,702	−3.04	−2.17	2.76	67
Small	56,746	4.42	1.87	1.94	59,765	−2.86	−2.04	2.54	251
Mid-size	35,031	2.43	1.91	1.93	33,773	−2.98	−2.18	2.65	125
Large	20,291	2.42	1.92	1.92	17,792	−3.01	−2.21	2.66	57

This table presents the descriptive statistics for major characteristics of our sample stocks. Our sample includes stocks listed on the NYSE, the AMEX, and Nasdaq with data available to compute book-to-market ratios, revenue surprises, and earnings surprises. All financial service operations and utility companies are excluded. Firms with prices below $5 as of the earnings announcement date are also excluded. Panel A lists numbers of firm-quarter observations between January 1974 and December 2009. Panel B and Panel C respectively list the mean and median values the measure of revenue surprises (SURGE) and for the measure of earnings surprises (SUE) across all firm-quarters in our sample. Statistics for positive surprises, negative surprises, and zero surprises are presented separately. Sample firms are also classified into bottom 30 %, middle 40 %, and top 30 % groups by their respective market capitalizations or book-to-market ratios. The breakpoints for the size subsamples are based on ranked values of market capitalization of NYSE firms. The breakpoints for the book-to-market subsamples are based on ranked values of book-to-market ratio of all sample firms

SURGE, SUE, and prior price performance than growth firms and large firms, although the differences in correlations across B/M and size groups are not significant. Table 81.2 also shows the fractions of months where non-zero correlations are significant at the 1 % level. These numbers again confirm that the correlations between SURGE and SUE tend to be strongest across various classifications of firms, followed by correlations between SURGE and prior returns, and then those between SUE and prior returns.

Table 81.2 Correlation among revenue surprises, earnings surprises, and prior price performance

Correlated variables	All firms	Subsample by B/M			Subsample by size		
Panel A. Pearson correlations among *SURGE*, *SUE*, and *prior 6-month returns*							
		Value	Mid	Growth	Small	Mid	Large
(*SURGE, SUE*)	0.3200***	0.3331***	0.3361***	0.2818***	0.3641***	0.2917***	0.2362***
	(101.17)	(84.46)	(107.04)	(65.93)	(118.69)	(69.91)	(42.64)
	[100 %]	[100 %]	[100 %]	[100 %]	[100 %]	[100 %]	[71.1 %]
(*SURGE, Prior returns*)	0.1458***	0.1272***	0.1263***	0.1353***	0.1686***	0.1304***	0.1061***
	(44.09)	(33.86)	(35.67)	(35.36)	(55.44)	(29.78)	(17.78)
	[88.7 %]	[41.5 %]	[64.6 %]	[62.7 %]	[86.9 %]	[55.4 %]	[35.9 %]
(*SUE, Prior returns*)	0.1868***	0.2120***	0.2015***	0.1496***	0.2330***	0.1523***	0.0959***
	(65.54)	(57.68)	(54.40)	(47.01)	(75.82)	(40.74)	(20.93)
	[98.4 %]	[81.9 %]	[92.7 %]	[68.1 %]	[98.8 %]	[67.1 %]	[23.7 %]
Panel B. Spearman rank correlations among *SUE*, *SURGE*, and *prior 6-month-returns*							
Correlated variables	All firms	Subsample by B/M			Subsample by size		
		Value	Mid	Growth	Small	Mid	Large
(*SURGE, SUE*)	0.3231***	0.3367***	0.3397***	0.2828***	0.3652***	0.2952***	0.2407***
	(106.09)	(93.92)	(112.08)	(68.22)	(124.45)	(72.92)	(45.40)
	[100 %]	[100 %]	[100 %]	[99.8 %]	[100 %]	[100 %]	[74.4 %]
(*SURGE, Prior returns*)	0.1426***	0.1227***	0.1255***	0.1315***	0.1647***	0.1285***	0.1032***
	(42.61)	(33.68)	(36.33)	(33.09)	(55.45)	(29.37)	(17.58)
	[86.6 %]	[41.8 %]	[63.4 %]	[58.2 %]	[87.8 %]	[53.3 %]	[35.0 %]
(*SUE, Prior returns*)	0.1834***	0.2117***	0.1980***	0.1383***	0.2314***	0.1501***	0.0959***
	(63.98)	(59.29)	(54.79)	(41.56)	(76.68)	(39.50)	(20.21)
	[97.2 %]	[84.0 %]	[91.1 %]	[62.0 %]	[99.3 %]	[64.8 %]	[23.2 %]

This table presents the correlations among SURGE, SUE and prior returns of our sample firms. At the end of each month, each sample firm should have its corresponding most current SUE, most current SURGE, and previous 6-month return. SURGE and SUE are winsorized at 5 % and 95 %, setting all SURGE and SUE values greater than the 95th percentile to the value of the 95th percentile and all SURGE and SUE values smaller than the 5th percentile to the value of the 5th percentile. Panel A lists the average Pearson correlations among SUE, SURGE, and prior returns between 1974 and 2009. Panel B lists the average Spearman rank correlations, where all sample firms are grouped into ten portfolios based on SURGE, SUE, and prior-6-month-returns independently at the end of each month. Decile 1 portfolio consists of firms with the lowest value of the attribute (SURGE, SUE, or prior 6-month returns), and Decile 10 consists of firms with the highest value of the attribute. The correlations are calculated at the end of each month. The values reported in the table are monthly averages of those correlations. Sample firms are further classified into bottom 30 %, middle 40 %, and top 30 % groups by their respective market capitalizations or book-to-market ratios at the end of the formation months. The breakpoints for the size subsamples are based on ranked values of market capitalization of NYSE firms. The breakpoints for the book-to-market subsamples are based on ranked values of book-to-market ratio of all sample firms. The numbers in parentheses are the average t-statistics under the null hypothesis that the correlation is zero. ***, **, and * indicate statistical significance at 1 %, 5 %, and 10 %, respectively. Percentages in brackets represent the fraction of the months with non-zero correlations that are significant at the 1 % level

These preliminary results suggest that revenue surprises and earnings surprises share highly correlated information, while each still have a distinctive content, a conclusion consistent with Swaminathan and Weintrop (1991) and Jegadeesh and Livnat (2006b). The information content conveyed by market information, i.e., prior returns, differs more from that carried by the two fundamental information measures, SURGE and SUE.

81.3.3 Descriptive Statistics for Stocks Grouped by SURGE, SUE, and Prior Returns

We next compare the firm characteristics for portfolios characterized by different revenue surprises (SURGE), earnings surprises (SUE) and prior returns. All sample stocks are sorted into quintiles based on their SURGE, SUE, and prior 6-month returns independently. The characteristics of those quintile portfolios are reported in Table 81.3. Several interesting observations emerge.

The price level, as expected, is found to be lowest for the price losers (P1). Stocks with negative revenue surprises (R1) or negative earnings surprises (E1) also have lower price levels, while the trend is not as obvious as for price losers. We also find price losers (P1) and price winners (P5) tend to be smaller stocks. Another interesting observation revealed in the book-to-market ratios is that stocks with the most positive SURGE or the most winning returns tend to be growth stocks. Stocks with the most positive SUE also have lower B/M ratios, but to much less of a degree. This suggests that growth stocks are characterized by strong revenue but not necessarily strong earnings.

The last three sections of Table 81.3 list the SURGE, SUE, and prior returns for those sorted portfolios. Stocks with strong SURGE also tend to have higher SUE and higher prior returns. A similar pattern is seen for stocks with high SUE or high prior returns. Stocks with strong SURGE, strong SUE, or winning prior returns tend to excel on all three information dimensions. This relation is consistent with the positive correlations reported in Table 81.2.

81.4 Empirical Results of Univariate Momentum Strategies

Table 81.4 presents the monthly returns to momentum strategies based on firms' revenue surprises (SURGE), earnings surprises (SUE), and prior price performance, respectively termed as revenue momentum, earnings momentum, and price momentum strategies. Decile portfolio results are reported here.

We first examine the profitability of revenue momentum. We are interested in knowing whether the well-documented post-announcement revenue drift also enables a profitable investment strategy. Following a similar strategy of earnings momentum by Chordia and Shivakumar (2006), we define a revenue momentum portfolio as a zero-investment portfolio by buying stocks with the most positive revenue surprises and selling stocks with the most negative revenue surprises.

Table 81.3 Descriptive statistics of characteristics of various portfolio groups

	SURGE					SUE					Prior 6-month returns				
	R1	R2	R3	R4	R5	E1	E2	E3	E4	E5	P1	P2	P3	P4	P5
Price															
Mean	22.05	23.50	24.13	25.03	27.27	23.16	23.59	23.93	25.26	26.03	16.86	23.39	26.62	28.29	26.83
Median	16.38	17.75	18.13	19.00	21.13	17.80	17.38	17.75	19.50	19.63	12.63	18.12	21.13	22.50	20.13
STD	25.92	26.41	27.11	30.71	29.78	27.27	32.49	26.73	25.36	28.05	21.33	27.84	27.77	29.17	31.85
Mkt Cap (million dollars)															
Mean	2,117	2,312	2,483	2,400	2,567	2,247	2,111	2,189	2,771	2,561	1,316	2,524	3,018	3,120	1,902
Median	218	239	239	250	286	238	227	236	275	253	169	256	310	322	222
STD	12,173	13,316	14,720	12,593	14,426	12,883	12,774	11,838	15,122	14,516	8,718	14,233	15,384	16,527	10,874
B/M															
Mean	0.7426	0.7130	0.6782	0.6389	0.5529	0.6868	0.6762	0.6765	0.6485	0.6378	0.7774	0.7317	0.6793	0.6148	0.5223
Median	0.6085	0.5769	0.5381	0.4948	0.4133	0.5446	0.5389	0.5371	0.5103	0.4988	0.6408	0.6027	0.5471	0.4822	0.3836
STD	0.5284	0.5215	0.5119	0.5034	0.4695	0.5250	0.5226	0.5157	0.4975	0.4948	0.5595	0.5209	0.4991	0.4760	0.4566
Prior-6-month-returns															
Mean	−0.0026	0.0056	0.0107	0.0153	0.0220	−0.0037	0.0054	0.0107	0.0156	0.0229	−0.0443	−0.0109	0.009	0.0297	0.0676
Median	−0.0025	0.0052	0.0100	0.0145	0.0209	−0.0038	0.0047	0.0098	0.0145	0.0216	−0.0403	−0.0093	0.0096	0.0294	0.0637
STD	0.0452	0.0446	0.0457	0.0464	0.0488	0.0460	0.0461	0.0459	0.0453	0.0469	0.0322	0.0236	0.0220	0.0236	0.0359

SUE															
Mean	−1.9168	−0.7191	−0.0497	0.5215	1.0800	−5.2584	−1.5994	0.0029	1.4750	4.3051	−1.4463	−0.6197	−0.1473	0.2722	0.8561
Median	−1.5894	−0.4592	0.1297	0.6600	1.1971	−4.9957	−1.4842	0.0169	1.4329	4.1362	−1.0550	−0.3205	0.0709	0.4132	0.8788
STD	3.5535	3.2176	3.1100	3.1184	3.4312	2.2130	0.8896	0.6252	0.6952	1.5130	3.5826	3.3691	3.3150	3.2623	3.2648
SURGE															
Mean	−4.7404	−1.6453	0.3942	2.4354	5.7428	−1.2127	−0.2628	0.4276	1.0688	2.1560	−0.6392	−0.0739	0.3424	0.8732	1.6730
Median	−4.5579	−1.5386	0.5175	2.5045	5.6739	−1.5664	−0.4180	0.4178	1.1661	2.3349	−0.7582	−0.1127	0.3592	0.9336	1.7770
STD	1.8859	1.3789	1.2745	1.2251	1.7428	4.0386	3.7035	3.5492	3.5099	3.6600	3.8580	3.7885	3.7861	3.7634	3.7450

This table presents the descriptive statistics of firm characteristics for stocks sorted on SURGE, SUE, and prior returns. All sample stocks are sorted independently according to their SURGE, SUE, and prior 6-month returns. R1 (E1) represents the quintile portfolio of stocks with the most negative SURGE (SUE), and R5 (E5) represents the quintile portfolio of stocks with the most positive SURGE (SUE). Similarly, P1 denotes the quintile portfolio of stocks with the lowest prior 6-month returns while P5 denotes the portfolio of stocks with the highest prior 6-month returns. Reported characteristics include price level, market capitalization, B/M ratio, SURGE, SUE and prior 6-month returns for component stocks in each corresponding quintile portfolio. The reported mean values are the equally weighted averages for stocks in each quintile portfolio

Panel A of Table 81.4 reports significant returns to the revenue momentum strategies. These strategies yield average monthly returns of 0.94 %, 0.93 %, and 0.84 %, respectively, by holding the relative-strength portfolios for 3, 6, and 9 months. This research, to the best of our knowledge, is the first to document specific evidence on the profitability of revenue momentum.

We also test with more recent data the profitability of earnings momentum and price momentum strategies, which have both been studied in the literature. Panel B of Table 81.4 reports the results for the earnings momentum strategies. We again find that these positive-minus-negative (PMN) zero-investment portfolios yield significantly positive returns for holding periods ranging from 3 to 12 months. The profit is strongest when the PMN portfolios are held for 3 months, leading to an average monthly return of 0.99 %, significant at the 1 % level. The results are consistent with those of Bernard and Tomas (1989) and Chordia and Shivakumar (2006). Chordia and Shivakumar (2006) find a significant monthly return of 0.96 % on a 6-month holding-period earnings momentum strategy executed over

Table 81.4 Returns to revenue momentum, earnings momentum, and price momentum strategies

Panel A. Revenue momentum returns					
Holding period	Low	High	PMN	CAPM_Adj. (1)	FF3_Adj. (2)
3 months	0.0074***	0.0163***	0.0089***	0.0084***	0.0105***
	(2.56)	(5.37)	(7.19)	(6.88)	(9.22)
6 months	0.0097***	0.0158***	0.0061***	0.0056***	0.0079***
	(3.34)	(5.17)	(5.10)	(4.71)	(7.32)
9 months	0.0118***	0.0154***	0.0036***	0.0030***	0.0054***
	(4.01)	(5.03)	(3.03)	(2.58)	(5.16)
12 months	0.0131***	0.0145***	0.0014	0.0010	0.0034***
	(4.43)	(4.78)	(1.24)	(0.87)	(3.36)
Panel B. Earnings momentum returns					
Holding period	Low	High	PMN	CAPM_Adj. (1)	FF3_Adj. (2)
3 months	0.0079***	0.0178***	0.0099***	0.0099***	0.0102***
	(2.71)	(6.14)	(9.77)	(9.71)	(9.90)
6 months	0.0098***	0.0169***	0.0071***	0.0070***	0.0077***
	(3.35)	(5.81)	(7.82)	(7.71)	(8.42)
9 months	0.0116***	0.0164***	0.0048***	0.0048***	0.0056***
	(3.92)	(5.65)	(5.68)	(5.59)	(6.63)
12 months	0.0127***	0.0155***	0.0028***	0.0028***	0.0037***
	(4.28)	(5.37)	(3.60)	(3.64)	(4.47)
Panel C. Price momentum returns					
Holding period	Loser	Winner	WML	CAPM_Adj. (1)	FF3_Adj. (2)
3 months	0.0085**	0.0179***	0.0094***	0.0101***	0.0113***
	(2.18)	(4.81)	(3.23)	(3.48)	(3.80)
6 months	0.0088**	0.0182***	0.0093***	0.0098***	0.0112***
	(2.29)	(4.94)	(3.47)	(3.62)	(4.09)
9 months	0.0099***	0.0183***	0.0084***	0.0085**	0.103***
	(2.62)	(5.02)	(3.57)	(3.62)	(4.32)

(*continued*)

Table 81.4 (continued)

12 months	0.0109***	0.0171***	0.0061***	0.0062***	0.0085***
	(2.94)	(4.72)	(2.93)	(2.94)	(4.06)

This table presents monthly returns and associated *t*-statistics from revenue, earnings, and price momentum strategies executed during the period from 1974 through 2009. For the revenue momentum strategy, firms are grouped into ten deciles based on the measure SURGE during each formation month. Decile 1 represents the most negative revenue surprises, and Decile 10 represents the most positive revenue surprises. The values of SURGE for each formation month are computed using the most recent revenue announcements made within three months before the formation date. The zero-investment portfolios—long the most positive revenue surprises portfolio and short the most negative revenue surprises portfolio (PMN) —are held for K ($K = 3$, 6, 9, and 12) subsequent months and are not rebalanced during the holding period. Panel A lists the average monthly returns earned from the portfolio of those firms with the most negative SURGE (low), from the portfolio of those with the most positive SURGE (high), and from the earnings momentum strategies (PMN). Earnings momentum strategies are developed with the same approach of revenue momentum strategies, by buying stocks with the most positive earnings surprises and selling stocks with the most negative earnings surprises. The zero investment portfolios are then held for K subsequent months. Panel B lists the average monthly returns earned from the portfolio of those firms with the most negative SUE (low), from the portfolio of those with the most positive SUE (high), and from the earnings momentum strategies (PMN). For the price momentum strategy, firms are sorted into 10 ascending deciles on the basis of previous 6 months returns. Portfolios of buying Decile 1 (winner) and selling Decile 10 (loser) are held for K subsequent months and not rebalanced during the holding period. The average monthly returns of winner, loser, and price momentum strategies are presented in Panel C. Risk-adjusted momentum returns are also provided in this table. Adj. (1) is momentum returns adjusted by CAPM, and Ad.j (2) is momentum returns adjusted by the Fama-French 3-factor model.***, **, and * indicate statistical significance at 1 %, 5 %, and 10 %, respectively

1972–1999, while we show a significant monthly return of 0.71 % for a sample period extending to 2009.

Panel C shows the performance of price momentum strategies. Similar to the results in Jegadeesh and Titman (1993), price momentum strategies yield average monthly returns of 0.94 %, 0.93 %, 0.84 %, and 0.61 %, for the 3, 6, 9, and 12 months holding-period respectively.

A comparison of the three momentum strategies indicates that the highest returns are for price momentum, followed by earnings momentum and revenue momentum. Meanwhile, the profitability for earnings momentum portfolio deteriorates faster than for price momentum as the holding period extends from 3 to 12 months.[10] The revenue momentum strategy yields the smallest and the shortest-lived profits, with returns diminishing to an insignificant level when the holding period is extended to 12 months.

Following a similar approach by Fama and French (1996) and Jegadeesh and Titman (2001), we implement the capital asset pricing model and a Fama-French three factor (FF-3) model to examine whether the momentum returns can be

[10]We show later that earnings momentum actually demonstrates stronger persistence than price momentum when the momentum portfolios are held over 2 years.

explained by pricing factors.[11] The last two columns in Panel A of Table 81.4 list the risk-adjusted returns to revenue momentum, which remain significant. The market risk premium, size factor, and book-to-market factor, while serving to capture partial effects of the revenue momentum strategy, are still unable to explain away abnormal returns entirely. The FF-3 factor adjusted return for 6 months remains strong at 0.79 % with a *t*-statistic equal to 7.32. The risk-adjusted returns to earnings momentum and price momentum in Panels B and C of Table 81.4 are similar to those in the literature (see Jegadeesh and Titman 1993; and Chordia and Shivakumar 2006) and generally confirm the conclusion of Fama (1998) that post-earnings-announcement drift and price momentum profits remain significant.

81.5 Interrelation of Revenue, Earnings, and Price Momentum

We further examine the interrelation of momentum strategies through tests of dominance, cross-contingencies, and combined strategies. The objective is to find empirical support for hypotheses for our two research questions. First, we hypothesize that revenue surprises, earnings surprises, and prior returns each have some exclusive information content that is not captured by the market. Under this hypothesis, a particular univariate momentum strategy should not be subsumed by another strategy, which we examine through dominance tests. Second, we hypothesize that the market not only underreacts to individual firm information, but also underestimates the significance of the joint implications of revenue, earnings, and price information. Under this hypothesis, return anomalies are likely to be most pronounced when the information variables all point in the same direction.

81.5.1 Testing for Dominance Among the Momentum Strategies

To tackle the interrelation of momentums, we first explore whether any of the three momentum strategies is entirely subsumed by another strategy. Stock price represents the firm value evaluated by investors in the aggregate, given their available information. The most important firm fundamental information for investors is undoubtedly firm earnings, which summarize firm performance. Jegadeesh and Livnat (2006b) point out that an important reference for investors regarding the persistence of firm earnings is offered by firm revenue information. Obviously, these three pieces of firm-specific information, revenue, earnings and stock price, share significant information content with each other. The anomalies of their corresponding momentums therefore may arise from common sources. That is, payoffs to a momentum strategy based on one measure, being revenue

[11] We obtain monthly data on market return, the risk-free rate, and SMB and HML from Kenneth French's website (http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/).

surprises, earnings surprises, or prior returns, may be fully captured by another measure. The dominance tests serve to test for such a possibility.

We first apply the pairwise nested comparison model introduced by George and Hwang (2004) and test whether one particular momentum strategy dominates another. Table 81.5 reports the results in three panels. Panel A compares the revenue momentum and earnings momentum strategies. In Panel A.1, stocks are first sorted on earnings surprises, with each quintile further sorted on revenue surprises. We find that, when controlling for the level of earnings surprises, the revenue momentum strategy still yields significant profits. The zero-investment portfolio returns for 6-month holding periods range from 0.26 % to 0.36 %. In Panel A.2, stocks are first sorted on revenue surprises, and then on earnings surprises. Likewise, the returns to an earnings momentum strategy, when controlling for the level of revenue surprises, are still significantly positive. These paired results indicate that neither earnings momentum nor revenue momentum dominates one another.

We follow the same process in comparing revenue momentum and price momentum strategies. Results in Panel B indicate that all the nested revenue momentum strategies and the nested price momentum strategies are found profitable, with the exception of revenue momentum in the loser stock group. In general, we still conclude that neither revenue momentum nor price momentum is dominated by the other. Panel C of Table 81.5 presents the results of the nested momentum strategies based on two-way sorts on earnings surprises and prior returns. Returns to all these nested momentum strategies remain significantly positive.

The pairwise nested comparisons suggest that revenue surprises, earnings surprises, and prior returns each convey some unpriced information which is not shared by each other, and therefore further contributes to a momentum effect.

A second approach allows us to simultaneously isolate the returns contributed by each momentum portfolio. Taking advantage of George and Hwang's (2004) model, we implement a panel data analysis with six performance dummies.

$$\begin{aligned} R_{it} = \alpha_{jt} &+ \beta_{1jt} R_{i,t-1} + \beta_{2jt} size_{i,t-1} + \beta_{3jt} R1_{i,t-j} + \beta_{4jt} R5_{i,t-j} \\ &+ \beta_{5jt} E1_{i,t-j} + \beta_{6jt} E5_{i,t-j} + \beta_{7jt} P1_{i,t-j} + \beta_{8jt} P5_{i,t-j} + e_{it} \end{aligned} \tag{81.3}$$

where $j = 1, \ldots, 6$. We first regress firm i's return in month t on control variables and six dummies for the portfolio ranks. We include the previous month return $R_{i,t-1}$ to control for the bid-ask bounce effect and the market capitalization $size_{i,t-1}$ to control for the size effect in the cross-sectional regressions. Momentum portfolio dummies, $R1_{i,t-j}$, $R5_{i,t-j}$, $E1_{i,t-j}$, $E5_{i,t-j}$, $P1_{i,t-j}$, and $P5_{i,t-j}$, indicate whether firm i is included in one or more momentum portfolios based on their scores in month $t - j$. To obtain momentum profits corresponding to the Jegadeesh and Titman (1993) strategies, we average the estimated coefficients of the independent variable over $j = 1, \ldots, 6$, and then subtract the coefficient average for the bottom quintile portfolio from that for the top quintile portfolio. These are the returns

Table 81.5 Momentum strategies: two-way dependent sorts by revenue surprises, earnings Surprises, and prior returns

Panel A. Revenue Momentum vs. Earnings Momentum							
A.1 Revenue momentum in various SUE groups				*A.2 Earnings momentum in various SURGE groups*			
Portfolios classified by SUE	Portfolios classified by SURGE	Ave. Monthly Return	*t*-stats	Portfolios classified by SURGE	Portfolios classified by SUE	Ave. Monthly Return	*t*-stats
E1 (Low)	R1 (Low)	0.0065		R1 (Low)	E1 (Low)	0.0064	
	R5 (High)	0.0101			E5 (High)	0.0104	
	R5-R1	0.0036	(3.24)		E5-E1	0.0040	(4.66)
E2	R1 (Low)	0.0086		R2	E1 (Low)	0.0079	
	R5 (High)	0.0115			E5 (High)	0.0113	
	R5-R1	0.0028	(2.85)		E5-E1	0.0034	(4.91)
E3	R1 (Low)	0.0090		R3	E1 (Low)	0.0089	
	R5 (High)	0.0119			E5 (High)	0.0131	
	R5-R1	0.0029	(3.29)		E5-E1	0.0042	(6.03)
E4	R1 (Low)	0.0096		R4	E1 (Low)	0.0096	
	R5 (High)	0.0122			E5 (High)	0.0140	
	R5-R1	0.0026	(2.70)		E5-E1	0.0043	(5.59)
E5 (High)	R1 (Low)	0.0116		R5 (High)	E1 (Low)	0.0112	
	R5 (High)	0.0149			E5 (High)	0.0152	
	R5-R1	0.0033	(3.22)		E5-E1	0.0040	(4.74)
Panel B. Revenue momentum vs. Price momentum							
B.1 Revenue momentum in various PriorRet groups				*B.2 Price momentum in various SURGE groups*			
Portfolios classified by Prior Ret	Portfolios classified by SURGE	Ave. Monthly Return	*t*-stats	Portfolios classified by SURGE	Portfolios classified by Prior Ret	Ave. Monthly Return	*t*-stats
P1 (Loser)	R1 (Low)	0.0070		R1 (Low)	P1 (Loser)	0.0072	
	R5 (High)	0.0077			P5 (Winner)	0.0095	
	R5-R1	0.0008	(0.67)		P5-P1	0.0024	(1.35)
P2	R1 (Low)	0.0083		R2	P1 (Loser)	0.0084	
	R5 (High)	0.0099			P5 (Winner)	0.0110	
	R5-R1	0.0015	(1.82)		P5-P1	0.0026	(1.51)
P3	R1 (Low)	0.0091		R3	P1 (Loser)	0.0092	
	R5 (High)	0.0123			P5 (Winner)	0.0135	
	R5-R1	0.0032	(4.33)		P5-P1	0.0042	(2.29)
P4	R1 (Low)	0.0089		R4	P1 (Loser)	0.0092	
	R5 (High)	0.0132			P5 (Winner)	0.0149	
	R5-R1	0.0042	(5.53)		P5-P1	0.0057	(3.35)
P5 (Winner)	R1 (Low)	0.0106		R5 (High)	P1 (Loser)	0.0080	
	R5 (High)	0.0175			P5 (Winner)	0.0176	
	R5-R1	0.0070	(7.03)		P5-P1	0.0096	(4.82)

(*continued*)

Table 81.5 (continued)

Panel C. Earnings momentum vs. Price momentum							
C.1 Earnings momentum in various PriorRet groups				*C.2 Price momentum in various SUE groups*			
Portfolios classified by SURGE	Portfolios classified by Prior Ret	Ave. Monthly return	*t*-stats	Portfolios classified by Prior Ret	Portfolios classified by SURGE	Ave. Monthly return	*t*-stats
P1 (Loser)	E1 (Low)	0.0063		E1 (Low)	P1 (Loser)	0.0066	
	E5 (High)	0.0096			P5 (Winner)	0.0097	
	E5-E1	0.0034	(3.73)		P5-P1	0.0031	(1.62)
P2	E1 (Low)	0.0082		E2	P1 (Loser)	0.0083	
	E5 (High)	0.0106			P5 (Winner)	0.0118	
	E5-E1	0.0024	(3.67)		P5-P1	0.0035	(1.80)
P3	E1 (Low)	0.0090		E3	P1 (Loser)	0.0081	
	E5 (High)	0.0126			P5 (Winner)	0.0134	
	E5-E1	0.0036	(5.96)		P5-P1	0.0052	(2.87)
P4	E1 (Low)	0.0091		E4	P1 (Loser)	0.0096	
	E5 (High)	0.0137			P5 (Winner)	0.0143	
	E5-E1	0.0046	(7.69)		P5-P1	0.0047	(2.65)
P5 (Winner)	E1 (Low)	0.0104		E5 (High)	P1 (Loser)	0.0100	
	E5 (High)	0.0178			P5 (Winner)	0.0177	
	E5-E1	0.0073	(8.78)		P5-P1	0.0077	(4.16)

This table presents the results of pairwise nested comparison between momentum strategies. Panel A shows the comparison between revenue momentum and earnings momentum during the period from 1974 to 2009. In each month, stocks are first sorted into five groups by earnings surprises (revenue surprises), then further sorted by revenue surprises (earnings surprises) in each group. All portfolios are held for 6 months. The monthly returns to 10 extreme portfolios and 5 conditional earnings (revenue) momentum strategies are presented. Pair tests are provided under the hypothesis that conditional earnings (revenue) momentum profits are the same. Panel B shows the comparison between revenue and price momentum strategies, and Panel C shows the comparison between earnings and price momentum strategies

contributed by each momentum strategy when the contributions from other momentum strategies are controlled for.

Panel A of Table 81.6 reports the regression results. The returns isolated for revenue momentum, earnings momentum, and price momentum are listed in the last three rows. The results are all significant in terms of either raw returns or FF-3 factor adjusted returns when all months are included or when all non-January months are included. Note, however, that the isolated returns to revenue momentum (R5−R1) and to price momentum (P5−P1) strategies are no longer significantly positive in January. The insignificant returns in January are consistent with the tax-loss-selling hypothesis, proposing that investors sell poorly performing stocks in October through December and buy them back in January (e.g., see Keim 1989; Odean 1998; Grinblatt and Moskowitz 2004).

The overall significant profits contributed by $R5 - R1$ ($E5 - E1$ or $P5 - P1$) indicate market underreactions with respect to the information content of revenue

Table 81.6 Comparison of revenue, earnings, and price momentum strategies

Panel A. Contribution of momentum returns solely from prior performance information						
	Raw returns			Risk-adjusted returns		
	All months	Jan.	Feb.–Dec.	All months	Jan.	Feb.–Dec.
Intercept	0.0130	0.0354	0.0110	0.0051	0.0075	0.0048
	(4.96)	(3.13)	(4.14)	(6.31)	(2.79)	(5.64)
$R_{i,t-1}$	−0.0412	−0.1146	−0.0346	−0.0371	−0.0851	−0.0327
	(−8.95)	(−6.04)	(−7.55)	(−8.54)	(−4.60)	(−7.48)
Size	>−0.0001	>−0.0001	>−0.0001	>−0.0001	>−0.0001	>−0.0001
	(−2.94)	(−2.64)	(−1.88)	(−1.12)	(−0.02)	(−0.99)
R1 Dummy	−0.0015	0.0042	−0.0020	−0.0020	0.0015	−0.0023
	(−3.39)	(2.81)	(−4.43)	(−4.68)	(1.02)	(−5.20)
R5 Dummy	0.0013	−0.0002	0.0014	0.0021	0.0019	0.0021
	(2.18)	(−0.07)	(2.31)	(4.09)	(1.06)	(3.81)
E1 Dummy	−0.0018	−0.0037	−0.0017	−0.0016	−0.0028	−0.0015
	(−5.07)	(−2.89)	(−4.42)	(−4.43)	(−1.96)	(−4.03)
E5 Dummy	0.0024	0.0045	0.0023	0.0026	0.0055	0.0023
	(6.86)	(3.61)	(6.09)	(6.81)	(4.05)	(6.06)
P1 Dummy	−0.0026	0.0125	−0.0040	−0.0036	0.0072	−0.0048
	(−2.14)	(1.96)	(−3.35)	(−3.17)	(1.13)	(−4.43)
P5 Dummy	0.0040	0.0023	0.0041	0.0044	0.0033	0.0045
	(2.92)	(0.53)	(2.88)	(3.39)	(0.67)	(3.33)
R5-R1	0.0028	−0.0044	0.0035	0.0041	0.0004	0.0044
	(3.23)	(−1.50)	(3.82)	(5.45)	(0.18)	(5.48)

E5-E1			0.0043	0.0082	0.0039	0.0041	0.0083	0.0038
			(7.89)	(4.64)	(6.92)	(7.45)	(4.34)	(6.66)
P5-P1			0.0066	−0.0102	0.0081	0.0080	−0.0039	0.0092
			(3.35)	(−1.15)	(4.10)	(3.95)	(−0.39)	(4.59)
Panel B. Univariate price momentum return and conditional price momentum returns								
	Raw returns				Risk-adjusted returns			
Intercept	0.0131	0.0129	0.0131	0.0130	0.0053	0.0051	0.0053	0.0051
	(4.97)	(4.92)	(5.03)	(4.96)	(6.62)	(6.36)	(6.58)	(6.31)
$R_{i,t-1}$	−0.0404	−0.0409	−0.0407	−0.0412	−0.0363	−0.0368	−0.0368	−0.0371
	(−8.72)	(−8.83)	(−8.87)	(−8.95)	(−8.29)	(−8.41)	(−8.46)	(−8.54)
Size	−0.0003	−0.0003	−0.0003	>−0.0001	−0.0001	−0.0001	−0.0001	>−0.0001
	(−2.72)	(−2.75)	(−2.97)	(−2.94)	(−0.83)	(−0.88)	(−1.16)	(−1.12)
R1 Dummy			−0.0021	−0.0015			−0.0026	−0.0020
			(−4.53)	(−3.39)			(−5.69)	(−4.68)
R5 Dummy			0.0018	0.0013			0.0026	0.0021
			(2.88)	(2.18)			(4.87)	(4.09)
E1 Dummy		−0.0022		−0.0018		−0.0021		−0.0016
		(−5.78)		(−5.07)		(−5.53)		(−4.43)
E5 Dummy		0.0028		0.0024		0.0030		0.0026
		(6.95)		(6.86)		(7.47)		(6.81)
P1 Dummy	−0.0034	−0.0029	−0.0030	−0.0026	0.0044	−0.0038	−0.0039	−0.0036
	(−2.74)	(−2.31)	(−2.44)	(−2.14)	(−3.83)	(−3.36)	(−3.48)	(−3.17)
P5 Dummy	0.0046	0.0042	0.0043	0.0040	0.0052	0.0047	0.0047	0.0044
	(3.33)	(3.01)	(3.14)	(2.92)	(3.98)	(3.59)	(3.64)	(3.39)

(*continued*)

Table 81.6 (continued)

R5-R1			0.0039	0.0028			0.0051	0.0041
			(4.26)	(3.23)			(6.45)	(5.45)
E5-E1		0.0050		0.0043		0.0051		0.0041
		(8.07)		(7.89)		(8.26)		(7.45)
P5-P1	0.0081	0.0070	0.0073	0.0066	0.0096	0.0085	0.0086	0.0080
	(4.02)	(3.52)	(3.70)	(3.35)	(4.70)	(4.18)	(4.29)	(3.95)

This table presents returns to relative strength portfolios and momentum strategies. Each month during the period from 1974 through 2009, six cross-sectional regressions are estimated for revenue, earnings, and price momentum strategies:
$R_{it} = \alpha_{jt} + \beta_{1jt}R_{i,t-1} + \beta_{2jt}size_{i,t-1} + \beta_{3jt}R1_{i,t-j} + \beta_{4jt}R5_{i,t-j} + \beta_{5jt}E1_{i,t-j} + \beta_{6jt}E5_{i,t-j} + \beta_{7jt}P1_{i,t-j} + \beta_{8jt}P5_{i,t-j} + e_{it}$,
where R_{it} and $size_{i,t}$ are the return and the market capitalization of stock i in month t; and $R1_{i,t-j}$ ($R5_{i,t-j}$) is the most negative (positive) revenue surprise dummy that takes the value of 1 if revenue surprises for stock i is ranked in the bottom (top) quintile in month $t-j$, and zero otherwise. The dummies with respect to earnings surprises ($E1_{i,t-j}$ and $E5_{i,t-j}$), and the dummies with respect to prior 6 month price returns ($P1_{i,t-j}$ and $P5_{i,t-j}$) are similar to the settings of $R1_{i,t-j}$ and $R5_{i,t-j}$. The estimated coefficients of independent variable are averaged over $j = 1, \ldots, 6$. The numbers reported for raw returns are the time-series average of these averages. The t-statistics calculated from the time series are in parentheses. The risk adjusted returns are intercepts from Fama-French 3-factor regressions on raw returns; their t-statistics are in parentheses. Panel A presents returns to relative strength portfolios and momentum strategies solely belong to each of prior price performance, earnings surprise, and revenue surprises. Panel B presents raw return and conditional returns of price momentum strategy

surprises (earnings surprises or prior price performance) unrelated to the other two information measures. The isolated returns are greatest for price momentum (0.66 %), followed by earnings momentum (0.43 %) and then revenue momentum (0.28 %). This is similar to our earlier results on single-criterion momentum. Such a finding again rejects the existence of a dominating momentum strategy among the three.

We do not find that information leading to revenue momentum or earnings momentum fully captures the price momentum returns. Similar findings are documented by Chan et al. (1996), Heston and Sadka (2008), and Novy-Marx (2012) for the relation between earnings surprises and price momentum. We would like to examine specifically how much of the price momentum can be explained by revenue surprises and/or earnings surprises information. For this reason, we perform similar regressions by including only a subset of portfolio dummies. The results are reported in Panel B of Table 81.6. In the case of raw returns, the return to price momentum without isolating other momentum sources is 0.81 %, while it is only reduced to 0.73 % after controlling for revenue momentum, to 0.70 % after controlling for earnings momentum, and to 0.66 % after controlling for both. In other words, information leading to revenue momentum and earnings momentum each accounts for about 10 % and 14 % of price momentum, and the two pieces of information combined account for just about 19 % of price momentum effects. The results for risk-adjusted returns are similar. This conclusion adds to the large literature attempting to trace the sources of price momentum. Our numbers indicate that the information conveyed by revenue surprises or earnings surprises seems to make only a limited contribution to price momentums.

Results of the pairwise nested comparisons in Table 81.5 and the regression analysis in Table 81.6 both support the hypothesis that revenue surprises, earnings surprises, and prior returns each have some unpriced information content that is exclusive to each measure itself. This conclusion also suggests the possibility that one can improve momentum strategies by using all three information measures.

81.5.2 Two-Way Sorted Portfolio Returns and Momentum Ccross-Contingencies

Here and in the next section, we examine the momentum strategies using multiple sorting criteria. These results serve to answer the research question of whether investors underestimate the implications of joint information of revenue surprises, earnings surprises, and prior returns.

Given that the market usually informs investors with not just a single piece but multiple pieces of firm information, the incremental information content of additional firm data is likely to be contingent upon other information for the stock. Jegadeesh and Livnat (2006b) suggest that the information content of SURGE has implications for the future value of SUE and such information linkage is particularly significant when both measures point in the same direction. Jegadeesh and

Livnat (2006a) further find that the market, including financial analysts, underestimates the joint implications of these measures and thus firm market value.

Our second research question extends Jegadeesh and Livnat (2006b) by additionally considering the information of prior price performance. We hypothesize that return anomalies should be most pronounced when the joint implications of multiple measures are most underestimated by the market, and this likely occurs when all information variables point in the same direction. In addition, a different but related issue is that any momentum profits driven by one measure may well depend on the accompanying alternative information, which we call the *cross-contingencies of momentum*. We use multivariate sorted portfolios to test this hypothesis.

81.5.2.1 Two-Way Sorts on Revenue Surprises and Earnings Surprises

We start by testing the performance of investment strategies based on the joint information of revenue surprises and earnings surprises. We sort stocks into quintiles on the basis of their revenue surprises and then independently into quintiles based on earnings surprises during the 6-month formation period on each portfolio formation date. Panel A of Table 81.7 presents the raw returns of these 25 two-way sorted portfolios. The intersection of R1 and E1, labeled as $R1 \times E1$, is the portfolio formed by the stocks with both the lowest SURGE and the lowest SUE, and the intersection of R5 and E5 labeled as $R5 \times E5$, represents the portfolio formed by the stocks with both the highest SURGE and the highest SUE.

We first note that the next-period returns of the 25 two-way sorted portfolios increase monotonically with SURGE as well as with SUE. The return to the portfolio with a similar level of SURGE increases with SUE (e.g., the return increases from 0.88 % for $R1 \times E1$ to 1.21 % for $R1 \times E5$). Similarly, the payoffs to the portfolio of stocks with a similar level of SUE increase with SURGE (e.g., the return increases from 1.23 % for $R1 \times E5$ to 1.70 % for $R5 \times E5$). That is, stocks that have performed well in terms of revenue and earnings continue to outperform expectations and yield higher future returns.

Panel D of Table 81.7 shows the corresponding risk-adjusted abnormal returns for each of the 5×5 double-sorted portfolios based on SURGE and SUE. The monotonicity we see in raw returns in Panel A persists for the risk-adjusted returns. The most positive abnormal returns are for the portfolio of high-SURGE and high-SUE stocks ($R5 \times E5$) while the most negative abnormal returns are for the portfolio of low-SURGE and low-SUE stocks ($R1 \times E1$). This provides direct and robust evidence that the return anomalies tend to be most pronounced when SURGE and SUE point in the same direction.

The evidence of monotonicity suggests that the market underreaction is at its extreme when different elements of stock performance information signal in the same direction, i.e., the scenarios of $R1 \times E1$ or $R5 \times E5$. These are the scenarios where the information of SURGE and SUE are expected to have the most significant joint implications for firm value, while market underestimation of their joint implications is found to be strongest, leading to the most pronounced return drifts in the next period. This observation is consistent with the suggestion by Jegadeesh and Livnat (2006a, b).

Table 81.7 Momentum strategies: two-way sorts by revenue surprises, earnings surprises, and prior returns

Panel A. Raw returns sorted on revenue surprises (SURGE) and earnings surprise (SUE)								
		SUE					*Arbitrage returns on portfolios sorted by earnings*	
		E1(Low)	E2	E3	E4	E5(High)		
SURGE	R1(Low)	0.0088	0.0107	0.0109	0.0112	0.0121	0.0049	(4.63)
	R2	0.0098	0.0112	0.0117	0.0129	0.0139	0.0041	(4.71)
	R3	0.0106	0.0121	0.0134	0.0142	0.0154	0.0048	(5.40)
	R4	0.0108	0.0124	0.0133	0.0137	0.0165	0.0057	(5.84)
	R5(High)	0.0123	0.0141	0.0141	0.0146	0.0170	0.0039	(3.43)
Arbitrage returns on portfolios sorted by revenue	0.0043	0.0034	0.0032	0.0034	0.0036			
	(2.86)	(2.59)	(2.84)	(2.84)	(2.70)			
Revenue-earnings combined momentum strategy: R5 × E5 – R1 × E1							**0.0081**	**(6.25)**
Panel B. Raw returns sorted on revenue surprises (SURGE) and prior price performance								
		Prior price performance					*Arbitrage returns on portfolios sorted by price*	
		P1(Loser)	P2	P3	P4	P5(Winner)		
SURGE	R1(Low)	0.0089	0.0104	0.0109	0.0109	0.0122	0.0034	(1.45)
	R2	0.0099	0.0112	0.0121	0.0121	0.0135	0.0036	(1.59)
	R3	0.0108	0.0125	0.0133	0.0139	0.0161	0.0053	(2.26)
	R4	0.0100	0.0125	0.0131	0.0141	0.0176	0.0076	(3.66)
	R5(High)	0.0090	0.0112	0.0143	0.0156	0.0198	0.0108	(4.67)
Arbitrage returns on portfolios sorted by price		0.0001	0.0008	0.0033	0.0048	0.0078		
		(0.06)	(0.79)	(3.69)	(5.21)	(6.43)		
Revenue-Price combined momentum strategy: R5 × P5–R1 × P1							**0.0109**	**(4.53)**
Panel C. Raw returns sorted on earnings surprises (SUE) and prior price performance								
		Prior price performance					*Arbitrage returns on portfolios sorted by price*	
		P1(Loser)	P2	P3	P4	P5(Winner)		
SUE	E1(Low)	0.0083	0.0103	0.0109	0.0107	0.0105	0.0045	(1.94)
	E2	0.0098	0.0115	0.0119	0.0126	0.0141	0.0044	(1.89)
	E3	0.0099	0.0117	0.0127	0.0134	0.0162	0.0062	(2.73)
	E4	0.0106	0.0120	0.0133	0.0138	0.0168	0.0062	(2.81)
	E5(High)	0.0107	0.0127	0.0149	0.0160	0.0201	0.0092	(4.01)
Arbitrage returns on portfolios sorted by earnings		0.0030	0.0023	0.0040	0.0053	0.0078		
		(2.66)	(3.10)	(5.49)	(7.51)	(7.79)		

(*continued*)

Table 81.7 (continued)

Price-revenue combined momentum strategy: E5 × P5 – E1 × P1							**0.0118**	**(5.47)**
Panel D. Risk-adjusted returns sorted on revenue surprises (SURGE) and earnings surprise (SUE)								
		SUE					*Risk-adjusted returns on portfolios sorted by earnings*	
		E1(Low)	E2	E3	E4	E5(High)		
	R1(Low)	−0.0043	−0.0026	−0.0020	−0.0013	0.0005	0.0049	(4.52)
	R2	−0.0029	−0.0017	−0.0008	0.0002	0.0013	0.0043	(4.79)
SURGE	R3	−0.0016	−0.0004	0.0008	0.0017	0.0032	0.0049	(5.38)
	R4	−0.0008	0.0006	0.0011	0.0018	0.0044	0.0052	(5.21)
	R5(High)	0.0021	0.0027	0.0023	0.0033	0.0054	0.0033	(2.86)
Risk-adjusted returns on portfolios sorted by revenue		0.0064	0.0053	0.0043	0.0046	0.0045		
		(4.78)	(4.32)	(3.98)	(4.16)	(3.54)		
Revenue-earnings combined momentum strategy: R5 × E5 – R1 × E1							**0.0081**	**(6.25)**
Panel E. Risk-adjusted returns sorted on revenue surprises (SURGE) and prior price								
		Prior price performance					*Risk-adjusted returns on portfolios sorted by price*	
		P1(Loser)	P2	P3	P4	P5(Winner)		
	R1(Low)	−0.0050	−0.0028	−0.0018	−0.0015	0.0002	0.0052	(2.22)
	R2	−0.0037	−0.0018	−0.0004	−0.0002	0.0015	0.0052	(2.26)
SURGE	R3	−0.0025	−0.0002	0.0009	0.0017	0.0040	0.0066	(2.76)
	R4	−0.0026	0.0002	0.0013	0.0025	0.0059	0.0085	(4.05)
	R5(High)	−0.0032	−0.0005	0.0029	0.0044	0.0086	0.0118	(4.98)
Risk-adjusted returns on portfolios sorted by price		0.0018	0.0023	0.0047	0.0059	0.0087		
		(1.38)	(2.49)	(5.80)	(6.86)	(7.49)		
Revenue-price combined momentum strategy: R5 × P5 – R1 × P1							**0.0097**	**(7.86)**
Panel F. Risk-adjusted returns sorted on earnings surprises (SUE) and prior price performance								
		Prior price performance					*Risk-adjusted returns on portfolios sorted by price*	
		P1(Loser)	P2	P3	P4	P5(Winner)		
	E1(Low)	−0.0051	−0.0024	−0.0012	−0.0010	0.0010	0.0062	(2.65)
	E2	−0.0039	−0.0014	−0.0005	0.0006	0.0027	0.0066	(2.79)
SUE	E3	−0.0033	−0.0011	0.0004	0.0013	0.0042	0.0075	(3.20)
	E4	−0.0025	−0.0004	0.0012	0.0020	0.0051	0.0076	(3.35)
	E5(High)	−0.0018	0.0003	0.0029	0.0041	0.0083	0.0096	(4.07)

(*continued*)

Table 81.7 (continued)

Risk-adjusted returns on portfolios sorted by earnings	0.0036	0.0027	0.0041	0.0051	0.0072		
	(3.25)	(3.64)	(5.61)	(7.07)	(7.05)		
Price-revenue combined momentum strategy: E5 × P5 – E1 × P1						**0.0133**	**(6.09)**

For each month, we form equal-weighted portfolios according to the breakpoints of two of three firm characteristics: a firm's revenue surprises (SURGE), its earnings surprises (SUE), and its prior 6-month stock performance. Panel A and Panel D present raw returns and risk-adjusted returns of the 25 portfolios independently sorted on SURGE and on SUE. The returns of a revenue-earnings combined momentum strategy are obtained by buying the portfolio of the best SURGE stocks and the stocks with the best SUE (SURGE = 5 and SUE = 5) and selling the portfolio of the poorest SURGE stocks and the stocks with the poorest SUE (SURGE = 1 and SUE = 1). Panel B and Panel E present raw returns and risk-adjusted returns of the 25 portfolios independently sorted on SURGE and on prior price performance. The returns of a revenue-price combined momentum strategy is obtained by buying stocks in the portfolio of the best SURGE and the highest price performance and selling stocks in the portfolio of the poorest SURGE and the lowest price performance. Panel C and Panel F present the raw returns and risk-adjusted of the 25 portfolios independently sorted on SUE and on prior price performance. The returns of a earnings-price combined momentum strategy is obtained by buying stocks in the portfolio of the best SUE and the highest price performance and selling stocks in the portfolio of the poorest SUE and the lowest price performance. We also present the arbitrage returns and risk-adjusted arbitrage returns of single sorted portfolios based on the quintiles of price performance, SUE or SURGE at the bottom (and on the right hand side) of each panel for the purpose of comparisons. Risk-adjusted return is the intercept of the Fama-French 3-factor regression where the dependent variable is the arbitrage return or the excess return which is the difference between the raw return and the risk-free rate

Investors may execute various long-short strategies with those 25 portfolios. Those listed in the farthest right column of Panel A indicate earnings momentum returns for stocks with a particular level of SURGE, while those listed in the last row are returns on revenue momentum for stocks with a given level of SUE.[12]

We now examine the cross-contingencies of momentum. The revenue momentum measure is 0.36 % per month in the high-SUE subsample E5 and 0.43 % per month in the low-SUE subsample E1. Meanwhile, the earnings momentum measure is 0.39 % per month in the high-SURGE subsample R5, and 0.49 % per month in the low-SURGE subsample R1. We do not observe significant variations in momentum returns across SUE or SURGE. Panel D shows similar patterns when returns to momentum portfolios are adjusted for size and B/M risk factors. All of the profits generated earnings momentum strategies or revenue momentum strategies remain significantly positive.

81.5.2.2 Two-Way Sorts on Revenue Surprises and Prior Returns

We apply similar sorting procedures based on the joint information of revenue surprises and prior price performance. The results for raw returns as shown in Panel

[12] Similar to Hong et al. (2000), one may characterize the former strategy as earnings momentum strategies that are "revenue-momentum-neutral" and the latter as revenue momentum strategies that are "earnings-momentum-neutral."

B of Table 81.7, generally exhibit a pattern similar to Panel A but with the following differences. Although the future returns still rise with SURGE among the average and winner stocks, they become insensitive to SURGE for loser stocks. A closer look at the return for portfolio $R1 \times P1$ down to the return for portfolio $R5 \times P1$ indicates that loser portfolio returns simply do not vary much with the level of SURGE.

Panel E lists risk-adjusted returns for the 5×5 portfolios sorted on prior returns and SURGE. A similar monotonic pattern, now in relation with SURGE as well as with prior returns, is observed for most of those abnormal returns. That is, stocks that have performed well in terms of revenue (firm fundamental information) and prior returns (firm market information) continue to outperform expectations and yield higher future returns, and vice versa.

As to the cross-contingencies of momentums, the results in Panel B indicate that the revenue momentum strategies executed with winner stocks yield higher returns than those executed with loser stocks. For example, the revenue momentum strategy executed with the most winning stocks yields a monthly return of 0.78 % ($R5 \times P5 - R1 \times P5$), while with the most losing stocks it yields only a monthly return of 0.01 % ($R5 \times P1 - R1 \times P1$). Likewise, the price momentum strategy executed with stocks with greater SURGE yields higher returns than with those with lower SURGE. For example, the price momentum strategy executed with the lowest SURGE stocks yields a monthly return of 0.34 % ($R1 \times P5 - R1 \times P1$), while with the highest SUE stocks it yields a monthly return as high as 1.08 % ($R5 \times P5 – R5 \times P1$). The difference of 0.74 percentage between R1 and R5 subsamples is statistically and economically significant, with price momentum profits more than 200 % higher in R5 than in R1.

These observations suggest that the revenue surprise information is least efficient among winner stocks, producing the greatest revenue drift for the next period, and that the prior return information is least efficient among stocks with the most positive SURGE producing the strongest return continuation. One noteworthy point is that revenue momentum is no longer profitable among loser stocks. Panel E shows similar patterns of momentum cross-contingencies when returns to momentum portfolios are adjusted for size and B/M risk factors.

The message for investment strategy is that prior returns are most helpful in distinguishing future returns among stocks with high SURGE, and the same is true for the implications of revenue surprises for stocks of high prior returns. On the other hand, when a stock is priced unfavorably by the market, the information of revenue surprises does not offer much help in predicting its future returns.

81.5.2.3 Two-Way Sorts on Earnings Surprises and Prior Returns

Panel C of Table 81.7 shows the raw returns for multivariate momentum strategies based on the joint information of earnings surprises and prior returns. Several findings are observed. First, as in the cases shown in Panels A and B, the next-period returns of the 25 two-way sorted portfolios increase monotonically with SUE as well as with prior returns. For example, when a firm has a highly positive earnings surprises (E5) while having had winning stock returns (P5), these two pieces of information together are likely to have particularly strong joint implications for firm value. Such condition leads to an average monthly return as high as 2.01 % in the next 6-month period, possibly attributable to even greater investor underreactions.

Panel F of Table 81.7 shows the risk-adjusted abnormal returns for each of the 5×5 double-sorted portfolios based on SUE and prior returns. The monotonicity we see in raw returns in Panel C persists for the risk-adjusted returns. The most positive abnormal returns are for the portfolio of high-SUE and high-prior-return stocks ($E5 \times P5$) while the most negative abnormal returns are for the portfolio of low-SUE and low-prior-return stocks ($E1 \times P1$).

Looking now at the cross-contingencies between earnings momentum and price momentum, the earnings momentum strategy executed with winner stocks yields higher returns (0.78 %) than that executed with loser stocks (0.30 %), and that the price momentum strategy executed with positive-SUE stocks yields higher returns (0.92 %) than that executed with negative-SUE stocks (0.45 %). Panel F shows risk-adjusted returns for these momentum strategies and reveals a similar pattern as in Panel C for raw returns. Results indicate that the market underreactions to price performance are contingent upon the accompanying earnings performance, and vice versa.

Can we reconcile our results on momentum cross-contingencies with the behavioral explanations for momentum returns? Barberis et al. (1998) observe that a conservatism bias might lead investors to underreact to information and then result in momentum profits. The conservatism bias, described by Edwards (1968), suggests that investors underweight new information in updating their prior beliefs. If we accept the conservatism bias explanation for momentum profits, one might interpret our results as follows.

Investors update their expectations of stock value using firm fundamental performance information as well as technical information, and their information updates are subject to conservatism biases. The evidence of momentum cross-contingencies suggests that the speed of adjustment to market performance information (historical price) is contingent upon the accompanying fundamental performance information (earnings and/or revenue), and vice versa. Our results in Panel B and Panel C of Table 81.7 suggest that stock prices suffer from a stronger conservatism bias from investors and thus delay more in their adjustment to firm fundamental performance information (earnings or revenue) when those stocks experience good news, instead of bad news, as to market performance (prior returns). This then leads to greater earnings or revenue momentum returns for winner stocks than for loser stocks. Similar scenario also leads to greater price momentum returns for high-SUE or high-SURGE stocks than for low-SUE or low-SURGE stocks.

This would mean that investors are subject to a conservatism bias that is asymmetric with respect to good news vis-à-vis bad news. That is, investors tend to be even more conservative in reacting to information on firm fundamental performance (market performance) for stocks issuing good news than those issuing bad news about their market performance (fundamental performance).

81.5.3 Combined Momentum Strategies

The negative results on dominance tests in Table 81.5 and Table 81.6 mean that each of the information variables, SURGE, SUE, and prior returns, at least to some extent, independently leads to abnormal returns. This then suggests that a combined momentum strategy using more than one of these information measures should offer improved momentum profits. While Chan et al. (1996), Piotroski (2000), Griffin et al. (2005), Mohanram (2005), Sagi and Seasholes (2007), Asness et al. (2013), and Asem (2009) have examined the profitability of combined momentum strategies based on other measures, to the best of our knowledge, we offer the first evidence on the profitability of combined momentum strategies using the three most accessible information on firm performance, i.e., prior returns, earnings surprises, and revenue surprises altogether.

81.5.3.1 Bivariate Combined Momentums

Table 81.8 compares and analyzes the combined momentum returns. Panel A shows raw and FF-3 factor adjusted returns to momentum strategies based on one-way, two-way, and three-way sorts. We start with bivariate combined momentums.

If we buy stocks with the highest SURGE and the highest SUE ($R5 \times E5$) while selling stocks with the lowest SURGE and the lowest SUE ($R1 \times E1$), such a *revenue-and-earnings combined momentum strategy* yields a monthly return as high as 0.81 %, which is higher than the univariate momentum return earned solely on the basis of revenue surprises (0.47 %) or earnings surprises (0.58 %) when using quintile portfolios. This result is also a consequence of our observation that the sorted portfolio returns increase monotonically with both SURGE and SUE.

Panel A of Table 81.8 also shows that investors earn an average monthly return of 1.09 % by buying stocks with the highest SURGE and the most winning prior returns ($R5 \times P5$) and selling stocks with the lowest SURGE and the most losing prior returns ($R1 \times P1$). This *revenue-and-price combined momentum strategy* again outperforms the simple revenue momentum (0.47 %) and the simple price momentum strategy (0.72 %). Similarly, an *earnings-and-price combined momentum strategy* offers an average monthly return of 1.18 %, which outperforms the univariate earnings momentum (0.58 %) and the price momentum strategy (0.72 %).

Note that the strategy using SURGE and SUE yields a return (0.81 %) poorer than that using SURGE and prior returns (1.09 %) or that using SUE and prior returns (1.18 %). This suggests that it is important to take advantage of market information (prior returns) as well as firm fundamental information (SURGE and SUE) when it comes to formulation of investment strategies.

81.5.3.2 Multivariate Combined Momentums

Next, we further sort stocks into quintiles independently and simultaneously based on SURGE, SUE, and prior price performance to obtain three-way sorted portfolios. A *revenue-earnings-price combined momentum* strategy is performed by buying the stocks with the most positive revenue surprises, the most positive

Table 81.8 Comparisons of assorted single and combined momentum strategies

Panel A. Summary of momentum returns from various single/multiple sorting criteria								
One-way sorts			Two-way sorts			Three-way sorts		
Momentum Strategy	Raw Return	Adj. Return	Momentum Strategy	Raw Return	Adj. Return	Momentum Strategy	Raw Return	Adj. Return
***Mom*(*R*)**	0.0047***	0.0063***	***Mom*(*R* + *E*)**	0.0081***	0.0097***	***Mom*(*R* + *E* + *P*)**	0.0144***	0.0168***
	(4.42)	(6.77)		(6.25)	(7.86)		(6.06)	(7.12)
***Mom*(*E*)**	0.0058***	0.0063***	***Mom*(*R* + *P*)**	0.0109***	0.0136***			
	(8.17)	(8.81)		(4.53)	(5.75)			
***Mom*(*P*)**	0.0072***	0.0087***	***Mom*(*E* + *P*)**	0.0118***	0.0133***			
	(3.36)	(4.01)		(6.25)	(6.09)			
Panel B. Contribution of momentum returns from single prior performance information								
Incremental return contribution of revenue momentum			Incremental return contribution of earnings momentum			Incremental return contribution of price momentum		
Diff. in momentum strategies	Return difference		Diff. in momentum strategies	Return difference		Diff. in momentum strategies	Return difference	
***Mom*(*R* + *P*) − *Mom*(*P*)**	0.0038***		***Mom*(*E* + *P*) − *Mom*(*P*)**	0.0048***		***Mom*(*E* + *P*) − *Mom*(*E*)**	0.0061***	
	(3.91)			(6.69)			(3.48)	

(*continued*)

Table 81.8 (continued)

***Mom*(*R* + *E*) − *Mom*(*E*)**	0.0023***	***Mom*(*R* + *E*) − *Mom*(*R*)**	0.0035***	***Mom*(*R* + *P*) − *Mom*(*R*)**	0.0063***
	(2.28)		(5.76)		(3.58)
***Mom*(*R* + *E* + *P*) − *Mom*(*P* + *E*)**	0.0024***	***Mom*(*R* + *E* + *P*) − *Mom*(*R* + *P*)**	0.0033***	***Mom*(*R* + *E* + *P*) − *Mom*(*R* + *E*)**	0.0062***
	(2.70)		(4.47)		(4.04)
Panel C. Contribution of momentum returns from multiple prior performance information					
Incremental return contribution of (revenue + earnings) momentum		Incremental return contribution of (revenue + price) momentum		Incremental return contribution of (earnings + price) momentum	
Diff. in momentum strategies	Return difference	Diff. in momentum strategies	Return difference	Diff. in momentum strategies	Return difference
***Mom*(*R* + *E* + *P*) − *Mom*(*P*)**	0.0072***	***Mom*(*R* + *E* + *P*) − *Mom*(*E*)**	0.0085***	***Mom*(*R* + *E* + *P*) − *Mom*(*R*)**	0.0096***
	(5.47)		(4.38)		(5.54)

This table presents the return contribution by considering additional sorting criterion, being revenue surprises, earnings surprises or prior returns. In the table, ***R***, ***E***, and ***P*** respectively refer to revenue momentum, earnings momentum, and price momentum strategy. Momentum strategies based on combined criteria are indicated with plus signs. For example, ***R* + *P*** denotes *revenue-price combined momentum strategy*, that is, $R5 \times P5 - R1 \times P1$. Panel A summarizes raw returns and risk-adjusted returns obtained from momentum strategies based on one-way sorts, two-way sorts, and three-way sorts. Risk-adjusted return is the intercept of the Fama-French 3-factor regression on raw return. Panel B lists the return contributions of each additional sorting criterion based on the return differences. The associated *t*-statistics are in parentheses. Panel C lists the incremental returns obtained by applying additional two sorting criteria. All returns are expressed as monthly returns.***, **, and * indicate statistical significance at 1 %, 5 %, and 10 %, respectively earnings surprises, and the highest prior returns ($R5 \times E5 \times P5$), and selling the stocks with the most negative revenue surprises, the most negative earnings surprises, and the lowest prior returns ($R1 \times E1 \times P1$). This leads to a monthly momentum return of 1.44 %, which provides the highest investment returns of all the paired momentum strategies discussed so far.

Panels B and C of Table 81.8 present the differences in portfolio performance, which indicate the incremental contribution to momentum portfolio returns from each additional sorting criterion. The results are straightforward. The joint consideration of each additional performance measure, whether it is revenue surprises, earnings surprises, or prior returns, helps improve the profits of momentum strategies significantly. The net contribution from price momentum is the greatest (0.62 %), followed by earnings momentum (0.33 %), and then revenue momentum (0.24 %). This result further supports the argument that revenue, earnings, and price all convey to some extent exclusive but unpriced information.

81.5.3.3 Dependent Sorts Versus Independent Sorts

With highly correlated sorting criteria, as indicated in Table 81.2, independent multiple sorts may result in portfolios with limited numbers of stocks and therefore insufficient diversification. This will then lead to results that might be confounded by factors other than the intended sorting features. More important, only dependent sorts provide a way to identify the precise conditional momentum returns.

Table 81.9 presents the returns and the associated *t*-statistics for two-way and three-way sorted combined momentum strategies using independent sorts and dependent sorts in different orders. For two-way sorted combined momentum strategies, dependent sorts are found to generate returns that are insignificantly different from those from independent sorts. For three-way sorted combined momentum strategies, however, the results are found to vary significantly with the sorting method. The three-way dependent sorts, in any order, yield investment strategies that significantly outperform those using independent sorts; independent sorts create an average monthly return of 1.44 %, while dependent sorts lead to an average monthly return ranging from 1.66 % to 1.89 %. Yet to take advantage of a more simplified presentation, we report results from only independent sorts in Tables 81.7 and 81.8. Note that the general conclusions we have drawn remain unchanged with dependent sorts.

81.6 Persistency and Seasonality

81.6.1 Persistence of Momentum Effects

We next examine the persistence of momentum effects driven by revenue surprises, earnings surprises, and prior price performance. Stock prices tend to adjust slowly to information, and abnormal returns will not continue once information is fully incorporated into prices. Following the argument of conservatism bias (see Edwards 1968; and Barberis et al. 1998), an examination of the persistence of

Table 81.9 Returns of combined momentum strategies – a comparison between dependent sorts and independent sorts

Momentum Strategies	Independent sorts	Dependent sorts					
		SURGE* | *SUE	***SUE* | *SURGE***				
***Mom*(*R* + *E*)**	0.0081***	0.0084***	0.0088***				
	(6.25)	(6.95)	(6.88)				
***Dep_sorts* – *Indep_sorts* (*t*-statistic only)**		(0.55)	(1.49)				
		P6* | *SURGE	***SURGE* | *P6***				
***Mom*(*R* + *P*)**	0.0109***	0.0104***	0.0106***				
	(4.53)	(4.66)	(5.19)				
***Dep_sorts* – *Indep_sorts* (*t*-statistic only)**		(−1.17)	(−0.57)				
		P6* | *SUE	***SUE* | *P6***				
***Mom*(*E* + *P*)**	0.0118***	0.0111***	0.0115***				
	(5.47)	(5.24)	(6.20)				
***Dep_sorts* – *Indep_sorts* (*t*-statistic only)**		(−1.76)	(−0.55)				
		P6*|*SURGE*|*SUE	***SURGE*|*P6*|*SUE***	***P6*|*SUE*|*SURGE***	***SUE*|*P6*|*SURGE***	***SURGE*|*SUE*|*P6***	***SUE*|*SURGE*|*P6***
***Mom* (*R* + *E* + *P*)**	0.0144***	0.0175***	0.0166***	0.0189***	0.0188***	0.0171***	0.0168***
	(6.06)	(4.16)	(4.12)	(4.29)	(4.45)	(4.47)	(4.36)
***Dep_sorts* – *Indep_sorts* (*t*-statistic only)**		(1.86)	(1.45)	(2.44)	(2.60)	(1.61)	(1.39)

This table presents returns and the associated *t*-statistics from two-way and three-way sorted combined momentum strategies, which are formed using independent sorts or dependent sorts. A momentum strategy formed on the basis of multiple criteria, which we call combined momentum strategy, is said to apply independent sorts if portfolios are independently sorted into quintiles according to their SURGE, SUE, and prior price performance, with the partition points being independent across these criteria. A combined momentum strategy is said to apply dependent sorts if portfolios are sorted into quintiles according to their SURGE, SUE, and prior price performance, with a particular sorting order. For example, a two-way sorted momentum strategy based on SURGE and SUE using dependent sorts could be formed by first sorting on SURGE then on SUE (*SUE*/*SURGE*) or first sorting on SUE then on SURGE (*SURGE*/*SUE*). We present here the returns of momentum strategies following all possible sequences of two-way dependent sorts and three-way dependent sorts.***, **, and * indicate statistical significance at 1 %, 5 %, and 10 %, respectively momentum returns will reveal the speed of adjustment in reaction to revenue surprises, earnings surprises, and prior returns. More interestingly, the variations of persistence in conditional momentums will demonstrate how one element of information (e.g., revenue surprises) affects the speed of adjustment to another (e.g., prior returns).

Table 81.10 presents the cumulative returns from revenue, earnings, and price momentum strategies. The formation period is kept at 6 months, and the cumulative returns are calculated up to 36 months after the event time. Panel A shows that the zero-investment portfolios built upon revenue surprises maintain their return momentum for 6 months. The buy-and-hold returns drop to insignificance 21 months after the portfolio formation. In Panel B, the profits of earnings momentum portfolios, although are not as high as on price momentum in the short term, demonstrate greater persistence than price momentum, with the cumulative returns continuing to drift upward for 25 months after portfolio formation. The cumulative returns still remain significant at 4.65 % 3 years after portfolio formation. Panel C shows that the profits to price momentum portfolio drift upward for 11 months after portfolio formation and start to reverse thereafter. The cumulative returns remain significant at 3.22 % on monthly terms 36 months after portfolio formation.

Figure 81.1 compares the cumulative returns to those three univariate momentum strategies. Price momentum generates the highest cumulative returns in the short term (for a 1 year holding period), while earnings momentum demonstrates the most persistent performance, as cumulative returns continue to grow up to 2 years after portfolio formation. On the other hand, the payoffs to revenue momentum seem to be neither as persistent nor as strong as the other two strategies.

Figure 81.2 presents the cumulative returns for momentum strategies conditional on alternative performance measures. Figure 81.2a, b present the cumulative returns of revenue momentum conditional on high-low SUEs and prior returns. They show that the revenue momentums remain short-lived, regardless of the level of SUE or the level of prior returns. The portfolio returns to a revenue momentum strategy with loser stocks not only quickly dissipate in the short term and actually reverse to negative returns starting 7 months after portfolio formation.

Figure 81.2c, d demonstrate the cumulative returns for earnings momentums conditional on high-low SURGE and prior returns. Figure 81.2c shows that the earnings momentum returns remain similar for the low-SURGE and the high-SURGE stocks during the first 20 months after portfolio formation. Such finding of momentum contingencies in fact conforms to our results in Panel A of Table 81.8. More interesting, as we hold the portfolio for over 20 months, the earnings momentum strategy with low-SURGE stocks starts deteriorating while the strategy with high-SURGE stocks still maintain significantly positive returns up to 36 months after the portfolio formation. Figure 81.2d, on the other hand, shows that earnings momentum effects are both greater and longer-lasting for winner stocks than for loser stocks. The caveat on investment strategy is that earnings momentum returns are higher and more longer-lived when applied over stocks with superior price history in the past 6 months.

Table 81.10 Cumulative returns from revenue, earnings, and price momentum strategies

Panel A. Revenue momentum				Panel B. Earnings momentum				Panel C. Price momentum			
t (month)	Negative SURGE (%)	Positive SURGE (%)	PMN (%)	*t* (month)	Negative SUE (%)	Positive SUE (%)	PMN (%)	*t* (month)	Loser (%)	Winner (%)	WMN (%)
1	0.68	1.69	1.02^{***}	1	0.66	1.83	1.17^{***}	1	1.12	1.48	0.36
2	1.50	3.26	1.75^{***}	2	1.53	3.47	1.94^{***}	2	1.96	3.19	1.23^{***}
3	2.45	4.66	2.21^{***}	3	2.50	4.95	2.44^{***}	3	2.79	4.76	1.97^{***}
4	3.57	6.06	2.49^{***}	4	3.59	6.42	2.83^{***}	4	3.69	6.42	2.74^{***}
5	4.78	7.49	2.71^{***}	5	4.76	7.92	3.17^{***}	5	4.67	8.10	3.4^{***}
6	6.13	8.87	2.75^{***}	6	5.97	9.40	3.43^{***}	6	5.66	9.86	4.21^{***}
7	7.49	10.21	2.72^{***}	7	7.19	10.80	3.61^{***}	7	6.64	11.62	4.99^{***}
8	8.92	11.51	2.59^{***}	8	8.51	12.14	3.63^{***}	8	7.76	13.21	5.45^{***}
9	10.40	12.82	2.42^{***}	9	9.88	13.50	3.62^{***}	9	9.00	14.77	5.78^{***}
10	11.93	14.02	2.09^{***}	10	11.27	14.75	3.49^{***}	10	10.22	16.18	5.95^{***}
11	13.44	15.19	1.76^{***}	11	12.66	16.00	3.34^{***}	11	11.52	17.54	6.02^{***}
12	14.95	16.39	1.44^{***}	12	14.05	17.33	3.28^{***}	12	12.91	18.80	5.89^{***}
13	16.26	17.57	1.31^{***}	13	15.28	18.69	3.41^{***}	13	14.31	19.91	5.60^{***}
14	17.59	18.78	1.19^{***}	14	16.49	20.05	3.57^{***}	14	15.73	21.04	5.31^{***}
15	18.86	20.01	1.15^{***}	15	17.66	21.42	3.76^{***}	15	17.13	22.19	5.06^{***}
16	20.23	21.33	1.09^{**}	16	18.95	22.89	3.94^{***}	16	18.64	23.43	4.79^{***}
17	21.61	22.67	1.07^{**}	17	20.26	24.41	4.15^{***}	17	20.15	24.72	4.57^{***}
18	22.96	24.03	1.07^{**}	18	21.56	25.94	4.37^{***}	18	21.59	26.09	4.50^{***}
19	24.40	25.38	0.98^{**}	19	22.90	27.44	4.54^{***}	19	22.91	27.71	4.79^{***}
20	25.94	26.79	0.85^{*}	20	24.34	28.95	4.62^{***}	20	24.33	29.29	4.96^{***}
21	27.45	28.13	0.68	21	25.77	30.41	4.64^{***}	21	25.79	30.89	5.10^{***}
22	28.91	29.48	0.57	22	27.21	31.88	4.67^{***}	22	27.23	32.38	5.14^{***}
23	30.37	30.91	0.54	23	28.67	33.39	4.72^{***}	23	28.74	33.90	5.17^{***}
24	31.83	32.38	0.55	24	30.18	34.90	4.72^{***}	24	30.29	35.41	5.12^{***}

25	33.24	33.79	0.54	25	31.62	36.36	4.74***	25	31.87	36.67	4.79***
26	34.68	35.19	0.51	26	33.11	37.80	4.69***	26	33.48	38.00	4.52***
27	36.08	36.57	0.49	27	34.54	39.22	4.68***	27	35.03	39.29	4.26***
28	37.53	37.98	0.45	28	36.01	40.69	4.67***	28	36.67	40.61	3.94***
29	39.06	39.41	0.35	29	37.52	42.18	4.66***	29	38.38	41.90	3.53***
30	40.58	40.85	0.26	30	39.04	43.62	4.58***	30	40.00	43.31	3.31***
31	42.13	42.38	0.25	31	40.54	45.14	4.60***	31	41.54	44.86	3.32***
32	43.77	43.93	0.16	32	42.08	46.66	4.59***	32	43.10	46.50	3.39***
33	45.38	45.44	0.06	33	43.60	48.14	4.53***	33	44.70	48.09	3.39***
34	46.96	46.95	−0.01	34	45.11	49.62	4.51***	34	46.35	49.61	3.27***
35	48.49	48.46	−0.03	35	46.59	51.20	4.60***	35	47.86	51.15	3.29***
36	50.06	49.97	−0.10	36	48.15	52.80	4.65***	36	49.46	52.68	3.22***

This table reports the cumulative returns of zero-cost momentum portfolio in each month following the formation period. *t* is the month after portfolio formation. Three different momentum strategies are tested. The sample period is from 1974 through 2009. Panel A reports the results from the revenue momentum strategy, where sample firms are grouped into five groups based on the measure SURGE during each formation month. The revenue momentum portfolios are formed by buying stocks with the most positive SURGE and selling stocks with the most negative SURGE. Listed are the cumulative portfolio returns for the portfolio with the most negative SURGE, the portfolio with the most positive SURGE, and the revenue momentum portfolio. Panel B reports the results from the earnings momentum strategy, where firms are grouped into five groups based on the measure SUE during each formation month. The earnings momentum portfolios are formed by buying stocks with the most positive SUE and selling stocks with the most negative SUE. Listed are the cumulative portfolio returns for the portfolio with the most negative SUE, the portfolio with the most positive SUE, and the earnings momentum portfolio. Panel C reports the results from the price momentum strategy. The price momentum portfolios are formed by buying Quintile 1 (winner) stocks and selling Quintile 5 (loser) stocks on the basis of previous 6 months returns. Listed are the cumulative portfolio returns for the loser portfolio, the winner portfolio, and the price momentum portfolio.***, **, and * indicate statistical significance at 1 %, 5 %, and 10 %, respectively

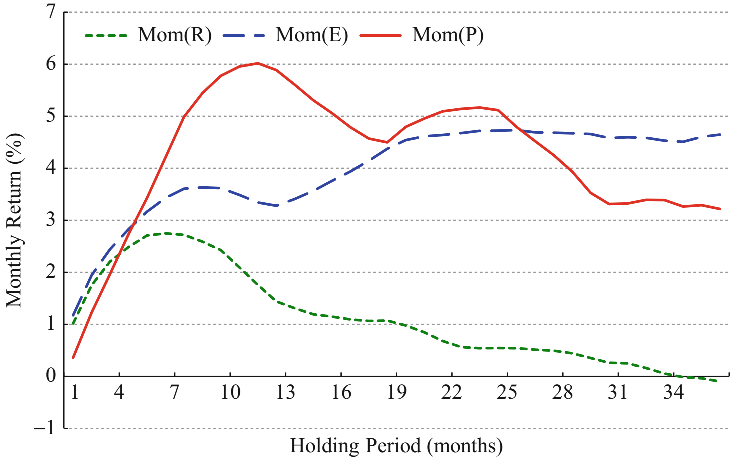


Fig. 81.1 Persistence of momentum effects. This figure shows the average cumulative returns of relative strength portfolios with respect to revenue surprises, earnings surprises, and prior price performance. The relative strength portfolio is buying stocks in highest quintile and selling stocks in lowest quintile on every formation date, and holding for 36 months. The cumulative returns are calculated by adding monthly returns from formation month t to month $t + i$

In Figure 81.2e, f, price momentum strategies yield higher and more persistent returns for stocks with positive SUE or SURGE than for stocks with negative SUE or SURGE. A comparison of Fig. 81.2e, f also finds that high-SURGE serves as a more effective driver than high-SUE for stocks to exhibit greater and more persistent price momentum.

These observations on momentum persistence provide further support for our claim on momentum cross-contingencies. We find that the persistence of a momentum, just like the magnitude of the momentum returns, depends on the accompanying condition of another firm information. Such cross-contingencies are again not as strong in the relation between revenue momentum and SUE or between earnings momentum and SURGE, as shown in Fig. 81.2a, c. Results suggest that investors update their expectations based on the joint information of revenue surprises, earnings surprises, and prior price performance, and the speed of adjustment to firm fundamental information (SURGE or SUE) depends on the prevailing content of firm market information (prior returns), and vice versa.

81.6.2 Seasonality

Jegadeesh and Titman (1993), Heston and Sadka (2008), Asness et al. (2013), Novy-Marx (2012), and Yao (2012) find that prior return winners outperform losers in all months except January, leading to positive profits for a price momentum

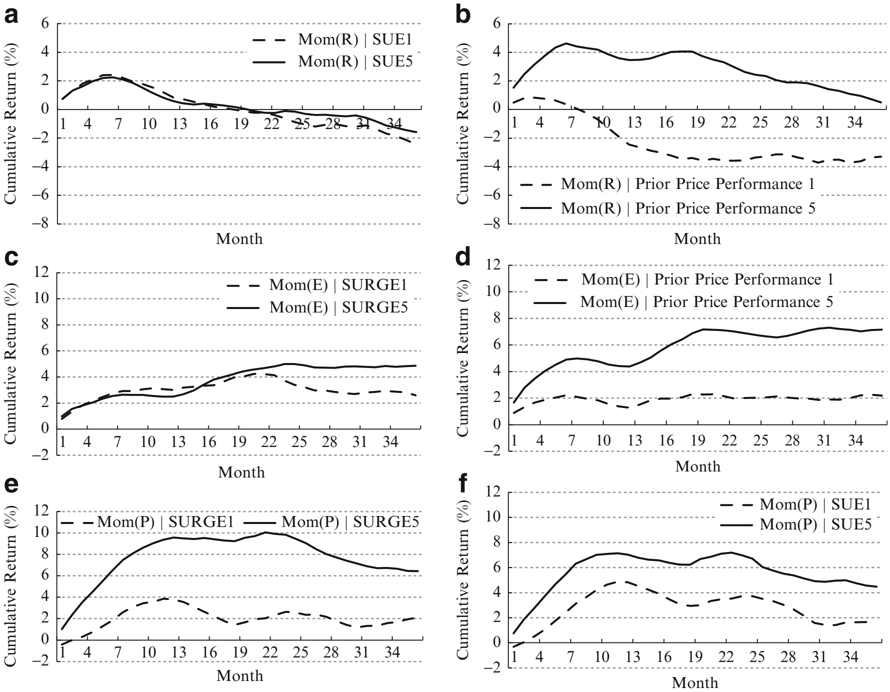


Fig. 81.2 **Cumulative returns of momentum effect conditional on performance measure**. These figures show the average cumulative returns of relative strength portfolio with respect to revenue surprises, earnings surprises, and prior price performance conditional on one another. The holding period is up to 36 months. The cumulative profits are calculated by adding monthly returns from formation month t to month $t + i$. (**a**) Cumulative returns of revenue momentum conditional on SUE. (**b**) Cumulative returns of revenue momentum conditional on prior price performance. (**c**) Cumulative returns of earnings momentum conditional on SURGE. (**d**) Cumulative returns of earnings momentum conditional on prior price performance

strategy in all months except January but negative profits for that strategy in January. Chordia and Shivakumar (2006) also find significant seasonality effects in returns to the earnings momentum strategy. Do a revenue momentum strategy and combined momentum strategies exhibit similar seasonalities?

Table 81.11 presents results for tests of seasonal patterns in returns to univariate momentum strategies and combined momentum strategies. For all types of momentum strategies, momentum profits in January are either negative or insignificantly different from zero. F-tests reject the hypothesis that the returns to momentum strategies are equal in January and non-January months. We therefore conclude that, as in finding elsewhere, there is seasonality in momentum strategies, and revenue surprises, earnings, surprises, and prior returns all yield significantly positive returns only in non-January months.

Table 81.11 Returns of momentum strategies in january and non-january months

Momentum Strategies	All months	Jan.	Feb.–Dec.	F-Statistic	p-Value
***Mom*(*R*)**	0.0047***	−0.0061	0.0057***	31.66	<0.01
	(4.42)	(−1.59)	(5.19)		
***Mom*(*E*)**	0.0058***	0.0026	0.0061***	6.12	0.01
	(8.17)	(0.72)	(8.67)		
***Mom*(*P*)**	0.0072***	−0.0134	0.0090***	28.28	<0.01
	(3.36)	(−1.32)	(4.25)		
***Mom*(*R* + *E*)**	0.0081***	−0.0062	0.0094***	37.42	<0.01
	(6.25)	(−1.09)	(7.22)		
***Mom*(*R* + *P*)**	0.0109***	−0.0164	0.0134***	40.05	<0.01
	(4.53)	(−1. 44)	(5.62)		
***Mom*(*E* + *P*)**	0.0118***	−0.0082	0.0136***	26.18	<0.01
	(5.47)	(−0.73)	(6.48)		
***Mom*(*R* + *E* + *P*)**	0.0144***	−0.0131	0.0169***	41.72	<0.01
	(6.06)	(−1.11)	(7.26)		

This table presents average monthly returns and the associated t-statistics for the returns obtained from single momentum strategies, two-way sorted combined momentum strategies, and two-way sorted combined momentum strategies for all calendar months, for January, and for non-January months. The F-statistics and p-values are computed under the hypothesis that the returns to momentum strategies are equal in January and non-January months. ***, **, and * indicate statistical significance at 1 %, 5 %, and 10 %, respectively

81.7 Conclusions

This study focuses on the three firm performance information that receive most attentions from investors – revenue, earnings and price. We attempt to understand how investors incorporate those information variables altogether in stock prices. Multivariate momentums are therefore used as a venue in the exploration. We provide new evidence that a revenue momentum strategy yields an average monthly return of 0.61 %, and remain significant after adjustment for market factor and FF-3 factors. Compared to the results of price momentum and earnings momentum, revenue momentum is less profitable and relatively short-lived.

Dominance tests show that none of the three momentum strategies generate returns that can be fully captured by the information driving an alternative strategy. This finding answers our first research question, and suggests that revenue surprises, earnings surprises, and prior returns each carry some unpriced information that is exclusive to itself. In particular, the information conveyed by revenue surprises and/or earnings surprises only makes a limited contribution to price momentum. The overall evidence indicates that while revenue serves as a base for a firm's earnings and stock valuation, revenue momentum neither drives nor rides earnings or price momentum.

Our second research question inquires how investors process the joint implications of multiple firm performance information. The results from double sorted portfolios find that next-period returns increase monotonically with each information variable, and the highest (lowest) abnormal return occurs for stocks receiving the best (worst) news in both variables. We further observe cross-contingencies of momentum profits in that momentum returns driven by fundamental performance information (SUE or SURGE) are positively associated with the accompanying market performance information (prior returns), and the reverse holds as well. For example, earnings/revenue momentum strategies with winner stocks yield higher returns than with loser stocks; a price momentum strategy with stocks with higher SURGE/SUE yields higher returns than with lower SURGE/SUE stocks. This pattern would mean that investors are subject to a conservatism bias that is asymmetric with respect to good news vis-à-vis bad news. The above findings are consistent with the claim that investors underestimate the joint implications of revenue surprises, earnings surprises, and prior returns, particularly when they point in the same direction. The speed of adjustment to firm fundamental information also depends on the accompanying market information, and vice versa.

The persistence of profitability also varies amongst the three momentums and exhibits inter-dependency. An earnings momentum strategy is found to present the strongest persistence, while the revenue momentum strategy is the shortest-lived among the three, except when the strategy is executed over price winner stocks. In general, the speed of adjustment to firm fundamental information also depends on the accompanying market information, and vice versa. Exploiting sources of momentums from three information variables altogether, a combined momentum strategy using independent sorts yields a monthly return of 1.44 %, amounting to an annual return as high as 17.28 %. The net contribution from prior return information is the greatest, followed by earnings surprises, and then revenue surprises.

Revenue, earnings, and historical prices are the most readily available firm performance information that investors use for stock evaluation. The pricing effect from investors' joint consideration of revenue, earnings, and prior returns is yet well explored in the finance literature. Our results are serving as useful guidance for asset managers identifying profitable investment strategies and for financial economists understanding the source of momentums in future research.

Appendix: A Measures of Earnings and Revenue Surprises

The literature provides a variety of measures to estimate earnings and revenue surprises. There are generally two approaches to building the measures; one is based on historical earnings/revenue data and the other on analysts' forecasts.

The empirical literature demonstrates consistent post-earnings-announcement drift, whichever method is used to measure the earnings surprises. For example, Foster et al. (1984) and Bernard and Thomas (1989) assume that the differences in quarterly earnings per share follow an AR(1) process, and find that firms with highly unexpected earnings outperform firms with poor unexpected earnings. Chan et al. (1996) analyze earnings momentum effects by applying three different earnings surprise measures built upon a seasonal random walk model, cumulative abnormal stock returns around the announcement date, and changes in analyst earnings forecasts. Jegadeesh and Livnat (2006a) use a seasonal random walk model with a drift and an analysts' forecast model to estimate earnings surprises, and find both approaches can capture the drift following earnings surprises.

Empirical research however finds inconsistent results as to whether revenues or expenses provide added information content over earnings, mostly thanks to the different measures being applied (e.g., see Hopwood and McKeown 1985; Swaminathan and Weintrop 1991; Ertimur et al. 2003; Rees and Sivamakrishnan 2001; Jegadeesh and Livnat 2006b). There are particular advantages and disadvantages when it comes to estimating expected earnings/revenues according to historical data or analyst forecast data. Considering that Compustat reports only restated accounting data, historical data on earnings/revenues might suffer a look-ahead bias to the extent that some input data are not available at the time we calculate earnings and revenue surprises. The analyst forecast approach has the advantage that it does not suffer from a potential look-ahead bias problem, and allows us to include in our sample young firms that do not have the accounting data required by the historical data approach. Its major disadvantage is that a sample will be limited to firms with analyst forecast data available.

Our study requires not only earnings forecast data but also revenue forecast data, which are not available from IBES until 1996, although even after 1996 many IBES sample firms still lack revenue forecasts. With such a restriction, the empirical results might be biased and lose their generality. Weighing the pros and cons, we elect to borrow the approach of Jegadeesh and Livnat (2006a, b) and measure earnings surprises and revenue surprises on the basis of historical earnings and revenues.

Specifically, we follow Jegadeesh and Livnat (2006b) and assume that quarterly earnings per share follow a seasonal random walk with a drift. We use the earnings per share in the same quarter of the previous year, instead of earnings per share in the previous quarter, to proxy for the earnings expectation; this approach takes into account the seasonality of earnings. We also accommodate a possible trend in earnings growth by including a drift term in the expected earnings. The drift term, $\delta^E_{i,t}$, is calculated from the average growth of previous eight quarters. Expected quarterly earnings per share for firm i and quarter t are estimated by

$$E\left(Q^E_{i,t}\right) = Q^E_{i,t-4} + \delta^E_{i,t} \tag{81.4}$$

and

$$\delta^E_{i,t} = \frac{\sum_{j=1}^{8}\left(Q^E_{i,t-j} - Q^E_{i,t-j-4}\right)}{8}. \tag{81.2}$$

The estimator for the standard deviation of quarterly earnings growth, $\sigma^E_{i,t}$, for computing earnings surprises is

$$\sigma^E_{i,t} = \frac{1}{7}\sqrt{\sum_{j=1}^{8}\left[Q^E_{i,t-j} - E\left(Q^E_{i,t-j}\right)\right]^2}. \tag{81.6}$$

We therefore define our measure of SUE for i in quarter t as Eq. 81.1 in the text:

$$SUE_{i,t} = \frac{Q^E_{i,t} - E\left(Q^E_{i,t}\right)}{\sigma^E_{i,t}}. \tag{81.1}$$

The same method is applied to measure revenue surprises. To deal with possible seasonal effects and trend effects in quarterly revenues, we again assume the quarterly revenue follows a seasonal random walk with a drift. That is, the expected quarterly revenue per share and the drift term are estimated as:

$$E\left(Q^R_{i,t}\right) = Q^R_{i,t-4} + \delta^R_{i,t} \tag{81.7}$$

and

$$\delta^R_{i,t} = \frac{\sum_{t=1}^{8}\left(Q^R_{i,t-j} - Q^R_{i,t-j-4}\right)}{8}. \tag{81.8}$$

For computing revenue surprises, the standard deviation of quarterly revenue growth is estimated by the year-to-year growth of revenue for the prior eight quarters:

$$\sigma^R_{i,t} = \frac{1}{7}\sqrt{\sum_{j=1}^{8}\left[Q^R_{i,t-j} - E\left(Q^R_{i,t-j}\right)\right]^2}. \tag{81.9}$$

Therefore the measure of revenue surprises is defined as Eq. 81.2 in the text:

$$SURGE_{i,t} = \frac{Q^R_{i,t} - E\left(Q^R_{i,t}\right)}{\sigma^R_{i,t}}. \tag{81.2}$$

References

Ahn, D. H., Conrad, J., & Dittmar, R. F. (2003). Risk adjustment and trading strategies. *Review of Financial Studies, 16*, 459–485.

Asem, E. (2009). Dividends and price momentum. *Journal of Banking & Finance, 33*, 486–494.

Asness, C. S., Moskowitz, T. J., & Pedersen, L. H. (2013). Value and momentum everywhere. *Journal of Finance, 68*, 929–985.

Bagnoli, M., Kallapur, S., Watts, & S. G. (2001). *Top line and bottom line forecasts: A comparison of internet firms during and after the bubble* (Working Paper). Purdue University.

Ball, R., & Brown, P. (1968). An empirical evaluation of accounting income numbers. *Journal of Accounting Research, 6*, 159–178.

Barberis, N., Shleifer, A., & Vishny, R. (1998). A model of investor sentiment. *Journal of Financial Economics, 49*, 307–343.

Bernard, V. L., & Thomas, J. K. (1989). Post-earnings-announcement drift: Delayed price response or risk premium? *Journal of Accounting Research, 27*, 1–36.

Bulkley, G., & Nawosah, V. (2009). Can the cross-sectional variation in expected stock returns explain momentum? *Journal of Financial and Quantitative Analysis, 44*, 777–794.

Chan, L. K., Jegadeesh, N., & Lakonishok, J. (1996). Momentum strategies. *Journal of Finance, 5l*, 1681–1713.

Chordia, T., & Shivakumar, L. (2002). Momentum, business cycle and time-varying expected return. *Journal of Finance, 57*, 985–1019.

Chordia, T., & Shivakumar, L. (2005). Inflation illusion and the post-earnings announcement drift. *Journal of Accounting Research, 43*, 521–556.

Chordia, T., & Shivakumar, L. (2006). Earnings and price momentum. *Journal of Financial Economics, 80*, 627–656.

Chui, A. C. W., Titman, S., & Wei, K. C. J. (2010). Individualism and momentum around the world. *Journal of Finance, 65*, 361–392.

Daniel, K., Hirshleifer, D., & Subrahmanyan, A. (1998). A theory of overconfidence, self-attribution, and security market under- and overreactions. *Journal of Finance, 53*, 1839–1885.

Edwards, W. (1968). Conservatism in human information processing. In B. Kleinmuntz (Ed.), *Formal representation of human judgment*. New York: Wiley.

Ertimur, Y., Livnat, J., & Martikainen, M. (2003). Differential market reactions to revenue and expense surprises. *Review of Accounting Studies, 8*, 185–211.

Fama, E. F. (1998). Market efficiency, long-term returns, and behavioral finance. *Journal of Financial Economics, 49*, 283–306.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance, 47*, 427–465.

Fama, E. F., & French, K. R. (1996). Multifactor explanations of asset pricing anomalies. *Journal of Finance, 51*, 55–84.

Foster, G., Olsen, C., & Shevlin, T. (1984). Earnings releases, anomalies, and the behavior of security returns. *The Accounting Review, 59*, 574–603.

George, T. J., & Hwang, C. Y. (2004). The 52-week high and momentum investing. *Journal of Finance, 59*, 2145–2176.

Ghosh, E., Gu, Z., & Jain, P. C. (2005). Sustained earnings and revenue growth, earnings quality, and earnings response coefficients. *Review of Accounting Studies, 10*, 33–57.

Griffin, J. M., Ji, X., & Martin, S. (2003). Momentum investing and business cycle risk: Evidence from pole to pole. *Journal of Finance, 58*, 2515–2547.

Griffin, J. M., Ji, X., & Martin, S. (2005). Global momentum strategies. *Journal of Portfolio Management, 31*, 23–39.

Grinblatt, M., & Moskowitz, E. J. (2004). Predicting stock price movements from past returns: The role of consistency and tax-loss selling. *Journal of Financial Economics, 71*, 541–579.

Grundy, B. D., & Martin, J. S. (2001). Understanding the nature of the risks and the source of the rewards to momentum investing. *Review of Financial Studies, 14*, 29–78.

Heston, S. L., & Sadka, R. (2008). Seasonality in the cross-section of stock returns. *Journal of Financial Economics, 87*, 418–445.

Hong, H., & Stein, J. C. (1999). A unified theory of underreaction, momentum trading and overreaction in asset markets. *Journal of Finance, 54*, 2143–2184.

Hong, H., Lim, T., & Stein, J. C. (2000). Bad news travels slowly: Size, analyst coverage, and the profitability of momentum strategies. *Journal of Finance, 55*, 265–295.

Hopwood, W., & McKeown, J. (1985). The incremental information content of interim expenses over interim sales. *Journal of Accounting Research, 23*, 161–174.

Jackson, A., & Johnson, T. (2006). Unifying underreaction anomalies. *Journal of Business, 79*, 75–114.

Jegadeesh, N., & Livnat, J. (2006a). Post-earnings-announcement drift: The role of revenue surprises. *Financial Analysts Journal, 62*, 22–34.

Jegadeesh, N., & Livnat, J. (2006b). Revenue surprises and stock returns. *Journal of Accounting and Economics, 41*, 147–171.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: Implications for stock market efficiency. *Journal of Finance, 48*, 65–91.

Jegadeesh, N., & Titman, S. (2001). Profitability of momentum strategies: An evaluation of alternative explanations. *Journal of Finance, 56*, 699–720.

Johnson, T. C. (2002). Rational momentum effects. *Journal of Finance, 57*, 585–608.

Keim, D. B. (1989). Trading patterns, bid-ask spreads and estimated stock returns: The cost of common stocks at calendar turning points. *Journal of Financial Economics, 25*, 75–97.

Korajczyk, R. A., & Sadka, R. (2004). Are momentum profits robust to trading costs? *Journal of Finance, 59*, 1039–1082.

Lee, C. M. C., & Swaminathan, B. (2000). Price momentum and trading volume. *Journal of Finance, 55*, 2017–2069.

Lee, C. F., & Zumwalt, J. K. (1981). Association between alternative accounting profitability measures and security returns. *Journal of Financial and Quantitative Analysis, 16*, 71–93.

Lesmond, D. A., Schill, M. J., & Zhou, C. (2004). The illusory nature of momentum profits. *Journal of Financial Economics, 71*, 349–380.

Li, X., Miffre, J., Brooks, C., & O'Sullivan, N. (2008). Momentum profits and time-varying unsystematic risk. *Journal of Banking & Finance, 32*, 541–558.

Liu, L., & Zhang, L. (2008). Momentum profits, factor pricing, and macroeconomic risk. *Review of Financial Studies, 21*, 2417–2448.

Mohanram, P. S. (2005). Separating winners from losers among low book-to-market stocks using financial statement analysis. *Review of Accounting Studies, 10*, 133–170.

Moskowitz, T. J., & Grinblatt, M. (1999). Do industries explain momentum? *Journal of Finance, 54*, 1249–1290.

Moskowitz, T. J., Ooi, Y. H., & Pedersen, L. H. (2012). Time series momentum. *Journal of Financial Economics, 104*, 228–250.

Novy-Marx, R. (2012). Is momentum really momentum? *Journal of Financial Economics, 103*, 429–453.

Odean, T. (1998). Are investors reluctant to realize their losses? *Journal of Finance, 53*, 1775–1798.

Piotroski, J. D. (2000). Value investing: The use of historical financial statement information to separate winners from losers. *Journal of Accounting Research, 38*, 1–41.

Raedy, J. S., Shane, P., & Yang, Y. (2006). Horizon-dependent underreatction in financial analysts' earnings forecasts. *Contemporary Accounting Research, 23*, 291–322.

Rees, L., & Sivaramakrishnan, K. (2001). *Valuation implications of revenue forecasts* (Working Paper). Texas A&M University.

Rouwenhorst, K. G. (1998). International momentum strategies. *Journal of Finance, 53*, 267–284.

Sagi, J. S., & Seasholes, M. S. (2007). Firm-specific attributes and the cross-section of momentum. *Journal of Financial Economics, 84*, 389–434.

Swaminathan, S., & Weintrop, J. (1991). The information content of earnings, revenues, and expenses. *Journal of Accounting Research, 29*, 418–427.

Verardo, M. (2009). Heterogeneous beliefs and momentum profits. *Journal of Financial and Quantitative Analysis, 44*, 795–822.

Wang, J., & Wu, Y. (2011). Risk adjustment and momentum sources. *Journal of Banking & Finance, 35*, 1427–1435.

Yao, Y. (2012). Momentum, contrarian, and the January seasonality. *Journal of Banking & Finance, 2757–2769*.

A VG-NGARCH Model for Impacts of Extreme Events on Stock Returns

82

Lie-Jane Kao, Li-Shya Chen, and Cheng-Few Lee

Contents

Abstract

This article compares two types of GARCH models, namely, the VG-NGARCH and the GARCH-jump model with autoregressive conditional jump intensity, i.e., the GARJI model, to make inferences on the log of stock returns when there

L.-J. Kao (✉)
Department of Finance and Banking, Kainan University, Taoyuan, ROC, Taiwan
e-mail: ljkao@mail.knu.edu.tw

L.-S. Chen
Department of Statistics, National Cheng-Chi University, Taipei City, Taiwan

C.-F. Lee
Department of Finance and Economics, Rutgers Business School, Rutgers, The State University of New Jersey, Piscataway, NJ, USA

Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
e-mail: cflee@business.rutgers.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_82,

are irregular substantial price fluctuations. The VG-NGARCH model imposes a nonlinear asymmetric structure on the conditional shape parameters in a variance-gamma process, which describes the arrival rates for news with different degrees of influence on price movements and provides an ex ante probability for the occurrence of large price movements. On the other hand, the GARJI model, a mixed GARCH-jump model proposed by Chan and Maheu (Journal of Business & Economic Statistics 20:377–389, 2002), adopts two independent autoregressive processes to model the variances corresponding to moderate and large price movements, respectively. An empirical study using daily stock prices of four major banks, namely, Bank of America, J.P. Morgan Chase, Citigroup, and Wells Fargo, from 2006 to 2009 is performed to compare the two models. The goodness of fit of the VG-NGARCH model vs. the GARJI model is demonstrated.

Keywords
VG-NGARCH model • GARCH-jump model • Autoregressive conditional jump intensity • GARJI model • Substantial price fluctuations • Shape parameter • Variance-gamma process • Ex ante probability • Daily stock price • Goodness of fit

82.1 Introduction

To model asset returns, the following two frequently observed circumstances must be recognized: the volatility clustering and the leverage effect (Nelson 1991; Campbell and Hentschel 1992; Engle and Ng 1993). The two phenomena have led to the development of the family of nonlinear asymmetric GARCH models in financial forecasting and derivatives pricing (Nelson 1991; Engle and Ng 1993; Glosten et al. 1993; Ding et al. 1993). Nevertheless, Gaussian distributed return innovations in conventional ARCH-/GARCH-type models are unable to capture irregular substantial price fluctuations resulting from extreme news reports, even when the heteroskedasticity in the conventional ARCH-/GARCH-type models has been taken care.

To account for both normal and large price movements, a mixed GARCH-jump model that combines a GARCH-type model with a Poisson jump process for the dynamics of log-returns was first proposed by Jorion (1988). Later, complicated mixed GARCH-jump models that consider jumps in both log-returns and volatilities were developed by Duffie et al. (2000), Pan (2002), Eraker et al. (2003), and Eraker (2004). The mixed GARCH-jump model with autoregressive jump intensity (GARJI), proposed by Chan and Maheu (2002), is a more advanced mixed GARCH-jump model, of which the conditional variance of asset returns is divided into two parts corresponding to moderate and large price movements resulting from normal and extreme news events, respectively. The dynamics of the two variances in a discrete-time setting are sketched by two conditionally independent autoregressive processes (Chan and Maheu 2002; Maheu and McCurdy 2004).

Instead of a jump-diffusion process in the mixed GARCH-jump model with a continuous sample path for the asset price dynamics, the VG-NGARCH model is a GARCH-type model that uses a variance-gamma (VG) process, a pure jump process having finite sum of absolute price movements during a defined time frame, to model the price dynamics to avoid the problem that the sum of absolute price movements during a finite time period is infinite. As pointed by Madan et al. (1998), the VG process is a purely jump Levy process of infinite activities characterizing a “high” arrival rate of jumps of different sizes and will adequately allow us to dispense with the need to consider the variant influences of news reports on the magnitude of price movements (Andersen 1996; Clark 1973; Ross 1989). With the VG process, the VG-NGARCH model captures the volatility clustering and the leverage effect by modeling the VG process’s shape parameter in a nonlinear asymmetric autoregressive process. For this reason, the VG-NGARCH model is more informative and parsimonious compared to the GARJI model. The specification of a VG process is given in Appendix 1.

The goodness of fit of the VG-NGARCH and the GARJI model to the log of stock price returns of four major banks listed in the S&P 500 are given. Since latent random business times are introduced into the VG framework, to find parameter estimates, Monte Carlo expectation-maximization (MCEM) algorithm together with the Metropolis algorithm are implemented. The two estimation approaches are given in Appendix 2.

The structure of this article continues as follows. In Sect. 82.2, the two GARCH-type models, namely, the VG-NGARCH and the GARJI models, are introduced. In Sect. 82.3, results of the empirical study and the performance of the two types of GARCH models are presented and compared. Finally, a concluding remark is made on Sect. 82.4.

82.2 Model Specifications

This section gives introduction and specifications for the GARJI model and the VG-NGARCH model.

82.2.1 GARJI Model

Chan and Maheu (2002) proposed a GARCH-jump model with autoregressive conditional jump intensity, i.e., the GARJI model, in which the conditional variance of return innovations is divided into two distinct modules that define smooth and steep fluctuations in price driven by normal and extreme news events, respectively. The GARJI model employs two conditionally independent autoregressive processes for the two components in a discrete-time economy in which the trading period

$[0, T]$ is partitioned into T subintervals $(0, 1], (1, 2], \ldots, (T-1, T]$. The dynamics of the log-return $Y_t = \ln(S_t/S_{t-1})$ are as follows:

$$Y_t = \mu + \varepsilon_t, \quad t = 1, \ldots, T. \tag{82.1}$$

Here the return innovation ε_t is partitioned into two independent components $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ corresponding to normal and unusual price movements, respectively. Let $\mathcal{F}_{t-1}$ be the information set available at time $t-1$. Conditional on $\mathcal{F}_{t-1}$, the innovation from normal price movement

$$\varepsilon_{1,t}|\mathcal{F}_{t-1} \sim N\left(0, \sigma_t^2\right) \tag{82.2}$$

is normally distributed with the conditional variance, $\sigma_t{}^2$, being parameterized by a GARCH function of the previous return innovation ε_{t-1} as

$$\sigma_t^2 = a_0 + a_1(\varepsilon_{t-1} - c)^2 + a_2\sigma_{t-1}^2. \tag{82.3}$$

The parameters employed for the GARCH function are based on the work of Chan and Maheu (2002) and Maheu and McCurdy (2004), albeit in a more simplified range that accommodates the asymmetric feedback from positive through negative news while allowing for the ex post evaluation of the expected number of jumps through the interval $(t-2, t-1]$ as a result of the information set $\mathcal{F}_{t-1}$ at time $t-1$. The second component, $\varepsilon_{2,t}$, represents the jump innovation and is the discrepancy between the total jump size and the expected total jump size of the n_t jumps during $(t-1, t]$, i.e.,

$$\varepsilon_{2,t} = \sum_{j=1}^{n_t} U_{t,j} - \theta\lambda_t,$$

where $U_{t,j}$ is the jth jump size being normally distributed with mean θ and standard deviation δ and n_t denotes the number of jumps distributed according to Poisson with an autoregressive conditional jump intensity (ARJI)

$$\lambda_t = \lambda_0 + \rho\lambda_{t-1} + \gamma\xi_{t-1}. \tag{82.4}$$

The intensity residual, $\xi_{t-1} = \mathrm{E}(n_{t-1}|\mathcal{F}_{t-1}) - \lambda_{t-1}$, is defined as the difference between the filter expected number of jumps given $\mathcal{F}_{t-1}$, $\mathrm{E}(n_{t-1}|\mathcal{F}_{t-1})$ and the previous intensity λ_{t-1}. The probabilities of jumps to fluctuate periodically and cluster with a persistence parameter of $0 < \rho < 1$ are afforded by specifying the conditional intensity. The conditional density of the log-return Y_t given the information set $\mathcal{F}_{t-1}$ is

$$f(Y_t|\mathcal{F}_{t-1}) = \sum_{n_t=0}^{\infty} f(Y_t|n_t, \mathcal{F}_{t-1}) \frac{e^{-\lambda_t}\lambda_t^{n_t}}{n_t!}, \tag{82.5}$$

where the conditional probability density $f(Y_t|n_t, \mathcal{F}_{t-1})$ is

$$\frac{1}{\sqrt{2\pi\left(\sigma_t^2 + n_t\delta^2\right)}}\exp\left(-\frac{(Y_t - \mu + \theta\lambda_t - \theta n_t)^2}{2\left(\sigma_t^2 + n_t\delta^2\right)}\right).$$

From Eqs. 82.1 to 82.5, it is clear that the main feature of GARJI model is the inclusion of both normal and extreme return innovations. Because these two types of innovations certainly affect future volatility differently. Nevertheless, it is impossible to identify the cutoff point between normal and extreme price movements by observing the log-returns. Consequently, only an ex post probability for the number of jumps, n_t, from the information set $\mathcal{F}_{t-1}$ at time t can be acquired, as n_t is non-observable. The ex post probability for n_t jumps given $\mathcal{F}_{t-1}$ is

$$P(n_t|\mathcal{F}_t) = \frac{f(Y_t|n_t, \mathcal{F}_{t-1})}{f(Y_t|\mathcal{F}_{t-1})} \times \frac{e^{-\lambda_t}\lambda_t^{n_t}}{n_t!}. \tag{82.6}$$

82.2.2 VG-NGARCH Model

To model stock price dynamics, Madan and Seneta (1990), Madan and Milne (1991), Madan et al. (1998), Carr et al. (2003), and Geman et al. (2001) considered the use of a VG process. In the following the specification of log-returns in terms of a VG process is given in a discrete-time setting. For $t = 1,\ldots, T$, the time-t log-return $Y_t = \ln(S_t/S_{t-1})$ can be formulated as

$$Y_t = m + \phi_t + \theta g_t + \varepsilon_t \tag{82.7}$$

where m denotes the mean of instantaneous return rate, g_t denotes a gamma-distributed random time change during the interval $(t - 1, t]$, and ϕ_t denotes a time-varying parameter. The specification of a VG process is given in Appendix 1. Because of the characteristics of the VG process, the return innovation ε_t is conditionally Gaussian distributed as

$$\varepsilon_t|\mathcal{F}_{t-1} \sim N\left(0, \sigma^2 g_t\right). \tag{82.8}$$

It is worth noting that the conditional variance of the innovation ε_t depends on g_t during the interval $(t - 1, t]$. To accommodate the volatility clustering effect, the random time change g_t is considered to be gamma-distributed with a time-varying shape parameter v_t, specifically

$$g_t|\mathcal{F}_{t-1} \sim \text{gamma}(v_t, 1). \tag{82.9}$$

It is further assumed that the shape parameter v_t follows a nonlinear asymmetric NGARCH (1,1) process that depends on the previous return innovation ε_{t-1} and shape parameter v_{t-1}, respectively. The relation among them is as follows:

$$v_t = a_0 + a_1\left(\varepsilon_{t-1} - c\sqrt{v_{t-1}}\right)^2 + a_2 v_{t-1}, t \geq 1 \tag{82.10}$$

where $c>0$. The time-varying parameter ϕ_t in Eq. 82.7 is defined to be

$$\phi_t = v_t \ln\left(1 - \theta - \frac{1}{2}\sigma^2\right). \tag{82.11}$$

The skewness and kurtosis of log-return at time t are functions of the drift parameter θ, volatility σ, and the first four moments of the shape parameter v_t, which depend on the NGARCH parameters $\boldsymbol{\alpha} = (a_0, a_1, a_2, c)$. The skewness and kurtosis functions are given in Appendix 3. According to the skewness and kurtosis functions, the sign of the skewness relies on the sign of the drift parameter θ. Moreover, if $a_0 > 0$ and $a_1(\sigma^2 + c^2) + a_2 < 1$, then shape parameter becomes stationary, and

$$v_\infty = \lim_{t\to\infty} E(v_{t+1}) = a_0\left[1 - \left(\sigma^2 + c^2\right)a_1 - a_2\right]^{-1}. \tag{82.12}$$

From Eq. 82.18, the proposal transition density $\boldsymbol{f}$ of the target distribution, i.e., the posterior distribution $p(\boldsymbol{g}|\boldsymbol{Y};\boldsymbol{\Theta})$, is chosen to be the distribution of T independent gamma random variables with shape parameters $v_1 - 0.5, \ldots, v_T - 0.5$, respectively, and scale parameter $1/\kappa$. Specifically,

$$\boldsymbol{f} \propto \prod_{t=1}^{T} \exp(-\kappa g_t + (v_t - 1.5)\log(g_t)).$$

At the lth iteration of the independent Metropolis chain algorithm, a random sample of time changes $\boldsymbol{g} = (g_1, \ldots, g_T)'$ is drawn from the proposed transition density $\boldsymbol{f}$. The random sample $\boldsymbol{g}$ is accepted and $\boldsymbol{g}^{(l)} = \boldsymbol{g}$ with probability

$$\min\left\{\exp\left[-\sum_{t=1}^{T}\left(\frac{\delta_t}{g_t} - \frac{\delta_t}{g_t^{(l-1)}}\right)\right], 1\right\}; \tag{82.13}$$

otherwise, $\boldsymbol{g}^{(l)} = \boldsymbol{g}^{(l-1)}$.

82.3 Empirical Study

Due to the extended periods of market instability experienced by banks as a consequence of the extreme news reports associated with the 2008 financial crisis, this study selects four big commercial banks listed in the S&P 500, namely,

Table 82.1 Correlations of daily log-returns

Bank	Bank of America	J.P. Morgan	Citigroup	Wells Fargo
Bank of America	1	0.81168	0.80231	0.84943
J.P. Morgan Chase		1	0.71070	0.84059
Citigroup			1	0.71889
Wells Fargo				1

Table 82.2 Summary for daily log-returns

Bank	Bank of America	J.P. Morgan	Citigroup	Wells Fargo
Days	754	754	754	754
Mean	−0.00138	−0.00019	−0.00246	0.00003
St.dev	0.03769	0.03269	0.04294	0.03153
Min	−0.30408	−0.19694	−0.3056	−0.21034
Max	0.2409	0.19368	0.45729	0.28371
Range	0.54498	0.39062	0.76289	0.49406
Skewness	−0.47452	0.02702	0.67837	0.09241
Kurtosis	16.62389	9.24964	27.40201	15.55658

the Bank of America (BAC), J.P. Morgan Chase (JPM), Citigroup (Citi), and Wells Fargo (WF), to study their stock price dynamics. Daily price data including intraday highs and lows, and adjusted close prices, are collected from January 3, 2006 to December 31, 2009. Table 82.1 lists the correlation coefficients among them. Since all correlation coefficients exceed 0.7, their price movements are highly correlated. Table 82.2 lists the mean, standard deviation, skewness, and kurtosis for log-returns of the four banks.

82.3.1 Estimation and Validation

Parameter estimates and log-likelihood of the VG-NGARCH model for each bank are given in Table 82.3. Besides model fitting, two testing procedures were conducted and the results are also contained in Table 82.3.

The first one with the null hypothesis, H_0: $a_1 = a_2 = c = 0$, helps us to determine whether the VG-NGARCH model can be reduced to a simpler model, i.e., the VG model by Madan et al. (1998). Based on the likelihood ratio test, the autoregressive shape dynamics are strongly favored for all banks over the VG model with constant shape parameter. The second testing procedure is the Ljung-Box test, with H_0 describing the randomness of residuals. To compute Ljung-Box Q statistic, the lag is set to be 25, and the large p-values for all banks, as shown in Table 82.3, indicate that there is no significant serial correlation remaining among the residuals for all banks after their log-returns were fitted by the VG-NGARCH model.

Table 82.3 VG-NGARCH model estimates and tests

Parameter	Bank of America	J.P. Morgan	Citigroup	Wells Fargo
θ	−0.0058	−0.0060	−0.0076	−0.0039
σ	0.0218	0.0205	0.0239	0.0178
v_1	2.7839	2.3371	2.9322	3.5000
a_0	0.1682	0.1082	0.1651	0.1619
a_1	0.4038	0.3929	0.4629	0.4529
a_2	0.5347	0.5408	0.4808	0.5008
c	0.0081	0.0005	0.0003	0.0083
Log-likelihood	1,951.7	1,857.4	1,841.6	1,945.0
H_0: $a_1 = a_2 = c = 0$				
Log-likelihood	1,684.6	1,688.1	1,636.7	1,763.4
Likelihood ratio test	526.2	345.2	409.8	386.8
p-value	0.0000	0.0000	0.0000	0.0000
H_0: Residuals are random				
Ljung-Box Q (lag = 25)	8.9911	6.5697	11.5924	8.2869
p-value	0.9986	0.9999	0.9929	0.9991

Table 82.4 GARJI model estimates and tests

Parameter	Bank of America	J.P. Morgan	Citigroup	Wells Fargo
λ_0	0.1165	0.1072	0.1164	0.0754
ρ	0.3923	0.4799	0.3053	0.3626
γ	0.4570	0.3579	0.1034	0.5886
θ	0.0102	0.0154	0.0099	0.0080
δ	0.0218	0.0341	0.0306	0.0546
σ_0	0.0312	0.0279	0.0405	0.0283
a_0	0.0001	0.0001	0.0001	0.0001
a_1	0.3568	0.4351	0.545	0.3634
a_2	0.5872	0.4997	0.3856	0.5761
c	0.0356	0.0645	0.0218	0.0909
Log-likelihood	1,887.1	1,843.7	1,764.1	1,918.0
H_0: Residuals are random				
Ljung-Box Q (lag = 25)	43.7338	30.5120	31.4383	43.7169
p-value	0.012	0.2061	0.1761	0.012

Parameter estimates and log-likelihood of GARJI for each bank are given in Tables 82.4. After fitting GARJI model for each bank, residuals are diagnosed by the Ljung-Box test, with the lag being 25. From Table 82.4, the p-values of the Ljung-Box Q statistic for all banks under GARJI model are much smaller than those under the VG-NGARCH model. To be more specific, p-values for Bank of America and Wells Fargo are too small to provide strong evidence that there is significant serial correlation among the residuals for Bank of America and Wells Fargo after fitting the GARJI model to their log-returns.

Table 82.5 Information criteria

	Bank of America	J.P. Morgan	Citigroup	Wells Fargo
VG-NGARCH model				
AIC	−5.1610	−4.9083	−4.8662	−5.0278
SC	−5.1181	−4.8654	−4.8233	−5.0056
HQ	−5.1445	−4.8918	−4.8497	−5.0106
GARJI model				
AIC	−4.9579	−4.8426	−4.6474	−4.9465
SC	4.8904	−4.7752	−4.5799	−4.8801
HQ	−4.9319	−4.8167	−4.6214	−4.9209

82.3.2 Model Selection Based on Information Criteria

Since the two models are not nested, their goodness of fit is measured by the following three criteria based on log of the maximum likelihood and the number of parameters: Akaike information criterion (AIC), Schwarz criterion (SC), and Hannan-Quinn criterion (HQ). The formulas for the three criteria are as follows:

$$\begin{aligned} \text{AIC} &= -2(\mathcal{L}/T) + 2(k/T) \\ \text{SC} &= -2(\mathcal{L}/T) + k\log(T)/T \\ \text{HQ} &= -2(\mathcal{L}/T) + 2k\log(\log(T))/T \end{aligned}$$

where $\mathcal{L}$ is log of the maximum likelihood, k is the number of parameters, and T is the sample size. The model minimizing these information criteria is preferred.

Comparing log-likelihoods and the three information criteria for the VG-NGARCH and GARJI models listed on Tables 82.3, 82.4, and 82.5, the VG-NGARCH model not only has higher log-likelihood values but also has smaller values on AIC, SC, and HQ for all banks, suggesting that it provides not only better fitting but also more parsimonious model specification for these bank data.

82.3.3 Evaluation of Volatility Forecasts

This subsection evaluates the performance of each model on variance forecasts through comparing the out-of-sample volatility forecasts of the VG-NGARCH model with those of the benchmark GARJI. To assess out-of-sample forecasts, a range-based estimate of ex post volatility was calculated in compliance with the method of Parkinson (1980) and Maheu and McCurdy (2004) as follows:

$$\text{Range}_t = \sqrt{\eta}\,\log\left(P_{t,h}/P_{t,l}\right),$$

where $P_{t,h}$ and $P_{t,l}$ represent the intraday high prices and low prices, respectively. The parameter η in the above formula is the calibration parameter to make the range

Table 82.6 Out-of-sample variance forecasts

	Bank of America	J.P. Morgan	Citigroup	Wells Fargo
VG-NGARCH				
β_0	0.0000	0.0023	0.0049	0.0098
β_1	1.1626	1.1666	1.2582	1.0759
R^2	0.6575	0.6407	0.6482	0.6148
F	963.8306	895.3023	924.9720	801.2075
p-value	0	0	0	0
GARJI				
β_0	−0.0392	−0.0221	−0.0333	−0.0206
β_1	1.3681	1.3284	1.4201	1.2472
R^2	0.6500	0.6317	0.6470	0.5727
F	932.1904	861.0316	919.9395	672.7951
p-value	0	0	0	0

estimate of the unconditional variance equal the unconditional variance of daily returns. To conduct out-of-sample analyses, all models were reestimated using observations from January 2, 2008 to December 31, 2009. These out-of-sample estimates were kept to derive the out-of-sample forecast for the date-t conditional variance given $\mathcal{F}_{t-1}$, denoted as $\widetilde{\mathrm{V}}\mathrm{ar}(Y_t|\mathcal{F}_{t-1})$. Following the approach of Maheu and McCurdy (2004), the range-based ex post volatilities were regressed on the out-of-sample forecasts of the conditional variances as

$$\mathrm{Range}_t = \beta_0 + \beta_1\sqrt{\widetilde{\mathrm{V}}\mathrm{ar}(Y_t|\mathcal{F}_{t-}1) + \mathrm{error}_t}. \tag{82.14}$$

The coefficient of determination, R^2, of the regression tells us the proportion of the total variation for the range-based volatilities explained by the out-of-sample conditional variance forecasts. Hence, the model with higher R^2 is considered to be superior in forecasting volatilities. The R^2s of the regression models given in Eq. 82.14 are displayed on Table 82.6 for all banks under the VG-NGARCH and GARJI models, respectively. Since R^2s under the VG-NGARCH model are all larger than those under the GARJI model for all banks, the VG-NGARCH outperforms the GARJI model on out-of-sample volatility forecasts.

82.3.4 Prediction of Large Price Movements or Jumps

In this subsection, the VG-NGARCH model is examined for its performance on predicting the probability of large price movements due to extreme events. Here, large price movement is defined to occur when the absolute log-return exceeds 0.05. From Eqs. 82.7, 82.8, and 82.9, the ex ante probability of large price movements is given by

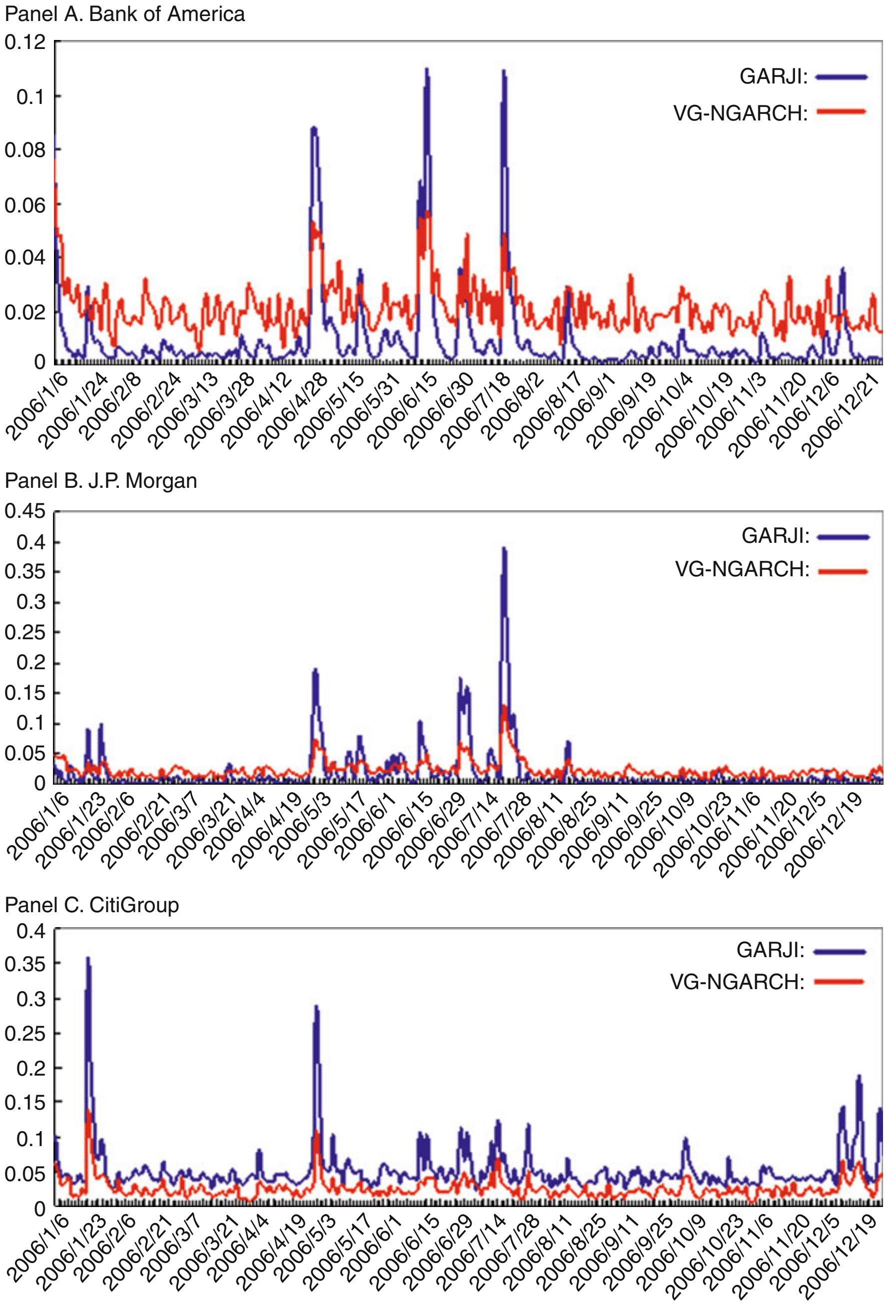


Fig. 82.1 (continued)

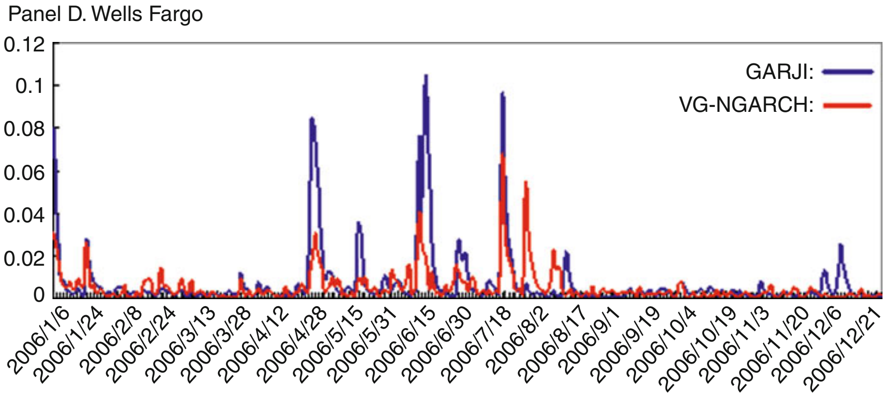


Fig. 82.1 Probabilities of large price movements or dumps during the year 2006

$$P(|Y_t| > 0.05|\mathcal{F}_{t-1}) = \int_0^\infty \left[1 - \Phi\left(\frac{0.05 - (m + \phi_t + \theta g_t)}{\sigma\sqrt{g_t}}\right)\right] h(g_t) dg_t + \int_0^\infty \Phi\left(\frac{-0.05 - (m + \phi_t + \theta g_t)}{\sigma\sqrt{g_t}}\right) h(g_t) dg_t \tag{82.15}$$

where $h(g_t)$ denotes the probability density function of a gamma distribution with shape and scale parameters v_t and 1, respectively.

For the GARJI model, the performance of prediction on jumps is based on the ex post probability of at least one jump occurring, which is expressed as

$$P(n_t \geq 1|\mathcal{F}_t) = 1 - P(n_t = 0|\mathcal{F}_t), \tag{82.16}$$

where, following Eq. 82.6,

$$P(n_t = 0|\mathcal{F}_t) = e^{-\lambda_t} \frac{f(Y_t|0, \mathcal{F}_{t-1})}{f(Y_t|\mathcal{F}_{t-1}).}$$

After the parameters in Eqs. 82.15 and 82.16 are replaced by their estimates, $P(|Y_t| > 0.05|\mathcal{F}_{t-1})$ and $P(|n_t \geq 1|\mathcal{F}_{t-1})$ are estimated for the whole study period for each bank. In order to compare their performances, the probabilities belonging to Years 2006 and 2008 are plotted on Figs. 82.1 and 82.2, respectively. From Fig. 82.1, the ex ante probabilities for large price movements under the VG-NGARCH model are smoother than those resulting from the GARJI model over the same period, where no noteworthy extreme events occurred during 2006.

Thus, the ex post probabilities of jumps from the GARJI model tend to over predict the chance of jumps when price movements are moderate.

On the other hand, as Year 2008 has been well recognized by the occurrences of financial turbulence and crisis, Fig. 82.2 does demonstrate that much higher

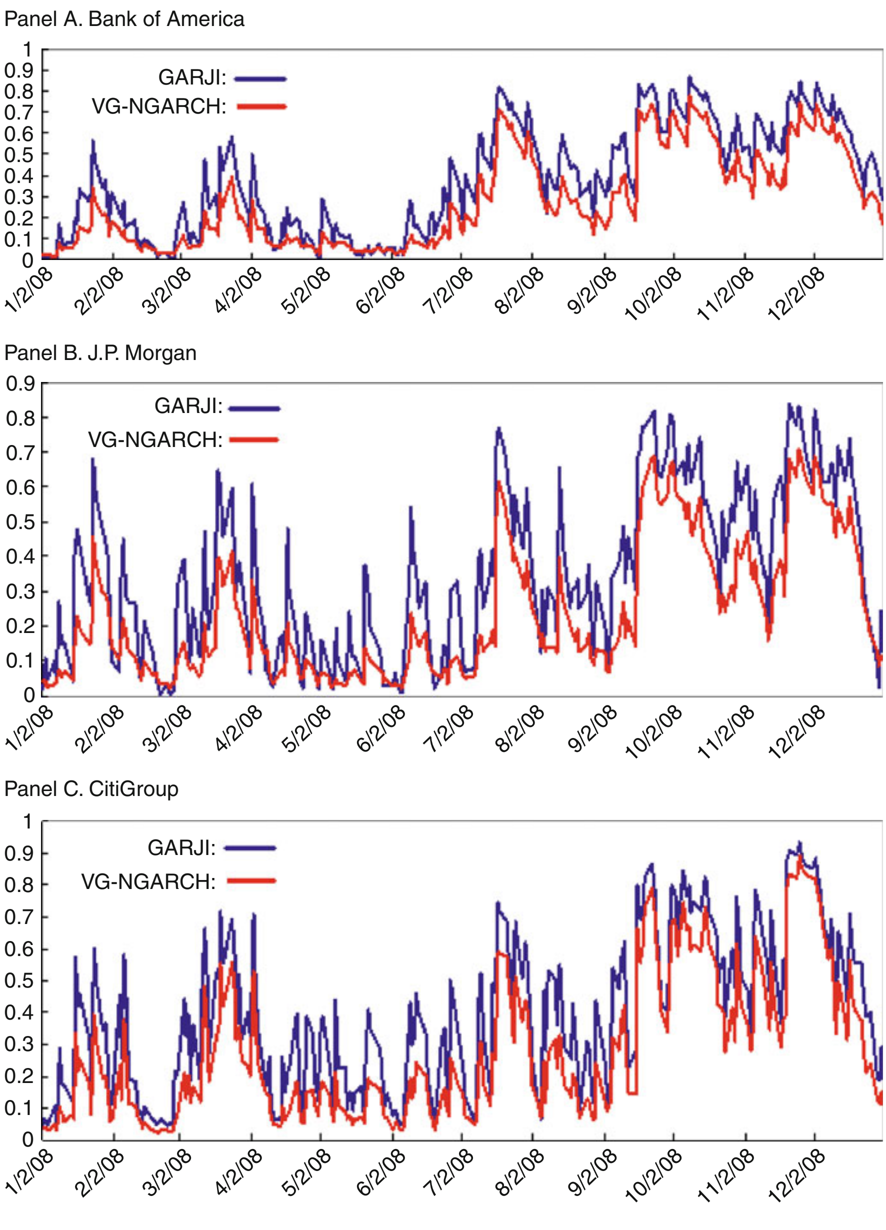


Fig. 82.2 (continued)

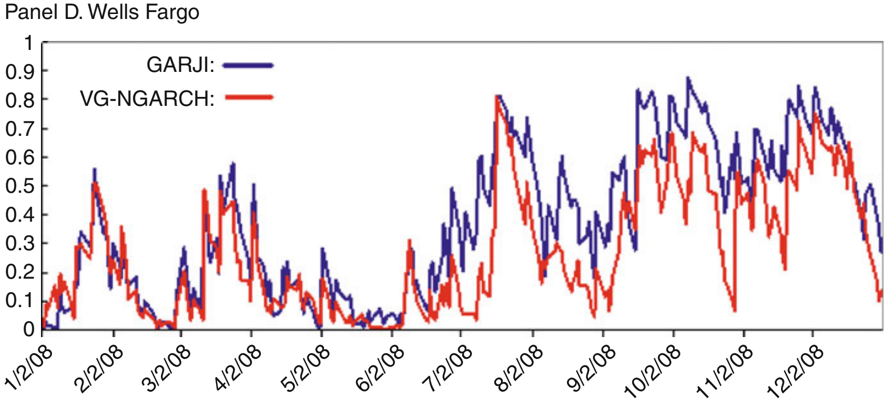


Fig. 82.2 Probabilities of large price movements or jumps during year 2008

probabilities of large price movements and jumps are predicted compared to those of Year 2006. Specifically, on September 16, 2008, the log-returns for Bank of America, J.P. Morgan, Citigroup, and Wells Fargo were −0.2398, −0.1066, −0.1642, and −0.1007, respectively. The corresponding ex ante probabilities of large price movements using the VG-NGARCH model were 0.7114, 0.4544, 0.6621, and 0.5368, respectively, while the ex post probabilities of jumps using the GARJI model were 0.8308, 0.6806, 0.7980, and 0.8348, respectively. Though both models show the ability to catch up large price movements or jumps, the VG-NGARCH model provides smoother and thus more reliable predictions than the GARJI model.

82.4 Conclusion

Differing from the GARJI model, for the VG-NGARCH model, based on a purely jump VG process, no cutoff point is required between normal and extreme price movements. In addition, instead of two independent autoregressive processes, a nonlinear asymmetric autoregressive process is used to model the shape parameter of the VG process. This makes the VG-NGARCH model more informative and parsimonious compared to the GARJI model. Furthermore, the empirical study demonstrates that through diagnosing the randomness of residuals, computing three information criteria (AIC, SC, and HQ), forecasting out-of-sample conditional volatility, and predicting the likelihood of large price movements or jumps, the VG-NGARCH model consistently outperforms the GARJI model. The superiority of the VG-NGARCH model relative to the benchmark GARJI model should improve the prediction ability of the occurrences of extreme events and hence is a better modeling approach to make financial management.

Appendix 1: Variance-Gamma Process

A VG process is a Brownian motion evaluated at a random business time modulated by a stochastic gamma process, to replace the role of Brownian motion. Specifically, at time s, a VG process X is given by

$$X(s) = \theta g(\mathrm{s}; 1, \gamma) + \sigma W(g(\mathrm{s}; 1, \gamma)), \tag{82.17}$$

where $g(\mathrm{s}; 1, \gamma)$ is the gamma process with unit mean rate and variance γ, W represents a standard Brownian motion, and θ and σ are the drift and volatility parameters, respectively. The extent of random time change $\Delta g = g(s;v,\gamma) - g(0;v,\gamma)$ is the increment of the gamma process during the interval (0, s]. Therefore, Δg follows gamma distribution with shape and scale parameters being vs and γ, respectively. Since the scale parameter γ can be transformed into one, it is set to one in our study.

Appendix 2: Parameter Estimation: Monte Carlo EM and Metropolis Algorithm

Method of maximum likelihood is adopted for the VG-NGARCH model. However, since the random time changes $g_1, \ldots, g_T$ are unobservable, the parameters of the VG-NGARCH model are unidentifiable. To resolve this problem, the mean of instantaneous return rate m is set to the mean of the log-returns $Y_1, \ldots, Y_T$, namely, $\hat{m} = \overline{Y}$, and the Monte Carlo EM (MCEM) algorithm (Wei and Tanner 1990; McCulloch 1997) is employed to estimate the parameters $\boldsymbol{\Theta} = (\theta, \sigma, v_1, \boldsymbol{\alpha})'$, where $\boldsymbol{\alpha} = (a_0, a_1, a_2, c)$ is the NGARCH parameter.

To perform the MCEM algorithm, at each iteration a set of K samples of the unobservable random time changes, $\boldsymbol{g}^{(1)}, \ldots, \boldsymbol{g}^{(K)}$, where $\boldsymbol{g}^{(l)} = (g_1{}^{(l)}, \ldots, g_T{}^{(l)})$, $1 \leqq l \leqq K$, are drawn from the posterior distribution $p(\boldsymbol{g}|\boldsymbol{Y};\boldsymbol{\Theta})$, which is

$$p(\boldsymbol{g}|\,\boldsymbol{Y}; \Theta) \propto \prod_{t=1}^{T} \exp\{-\kappa g_t - \delta_t/g_t + (v_t - 1.5)\log(g_t)\}, \tag{82.18}$$

where $v_t = a_0 + a_1\left(\varepsilon_{t-1} - c\sqrt{v_{t-1}}\right)^2 + a_2 v_{t-1}$ is the time-t shape parameter, and the coefficients κ and δ_t are

$$\kappa = \frac{\theta^2}{2\sigma^2} + 1 \ \text{ and } \ \delta_t = \frac{\left(Y_t - \overline{Y} - \phi_t\right)^2}{2\sigma^2}.$$

Since Eq. 82.18 is not proportional to any density function of well-known distributions, it is not possible to directly sample the time changes, $\boldsymbol{g}^{(1)}, \ldots, \boldsymbol{g}^{(K)}$, from the posterior distribution $p(\boldsymbol{g}|\boldsymbol{Y};\boldsymbol{\Theta})$ at each iteration of the EM algorithm. Consequently, the Metropolis chain strategy is carried out here

(Metropolis et al. 1953; Hastings 1970). In the independent Metropolis chain algorithm, a random outcome is sampled from its target distribution π by generating a chain of size L as follows: at the nth step of the chain, if the chain is at a point $X_n = \boldsymbol{x}$, a candidate value $\boldsymbol{y}$ is sampled from a proposal transition density $f(\boldsymbol{y})$ for the next location X_{n+1}. The candidate $X_{n+1} = \boldsymbol{y}$ is accepted with probability

$$p(x, y) = \min\left\{\frac{\pi(\boldsymbol{y})f(\boldsymbol{x})}{\pi(\boldsymbol{x})f(\boldsymbol{y})}, 1\right\}.$$

An independent uniform random variate U is generated; if $U < p(\boldsymbol{x}, \boldsymbol{y})$, then $X_{n+1} = \boldsymbol{y}$; otherwise the step is rejected and the chain remains at $X_{n+1} = \boldsymbol{x}$. After L such steps, where L is sufficiently large, a realization is obtained from the target distribution π.

Appendix 3: Skewness and Kurtosis of Log-returns

The unconditional skewness and kurtosis of log-return at time t are expressed in terms of the model parameters and the first four moments of the shape parameter v_t, which are

$$\text{Skewness}(Y_t) = \frac{\left(2\theta^3 + 3\theta\sigma^2\right)E(v_t) + 3(\tau+\theta)\left(\theta^2+\sigma^2\right)V(v_t) + (\tau+\theta)^3 E^3(v_t - E(v_t))}{\left[\left(\theta^2+\sigma^2\right)E(v_t) + (\tau+\theta)^2 V(v_t)\right]^{3/2}}; \tag{82.19}$$

$$\text{Kurtosis}(Y_t) = \frac{3\left(\theta^2+\sigma^2\right)^2 E\left(v_t^2\right) + \left(3\sigma^4 + 6\theta^4 + 12\theta^2\sigma^2\right)E(v_t) + Q}{\left[\left(\theta^2+\sigma^2\right)E(v_t) + (\tau+\theta)^2 V(v_t)\right]^2}, \tag{82.20}$$

where $\tau = \ln(1 - \theta - \sigma^2/2)$; and

$$\begin{aligned} Q = {} & 6(\tau+\theta)^2\left(\theta^2+\sigma^2\right)\left[E\left(v_t^3\right) + E^3(v_t) - 2E(v_t)E\left(v_t^2\right)\right] \\ & + 4(\tau+\theta)\left(2\theta^4 + 3\theta^2\sigma^2\right)V(v_t) + (\tau+\theta)^4 E^4(v_t - E(v_t)). \end{aligned}$$

References

Andersen, T. G. (1996). Return volatility and trading volume: An information flow interpretation of stochastic volatility. *Journal of Finance, 51*, 169–204.

Campbell, J. Y., & Hentschel, L. (1992). No news is good news: An asymmetric model of changing volatility in stock returns. *Journal of Financial Economics, 31*, 281–381.

Carr, P., Geman, H., Madan, D. B., & Yor, M. (2003). Stochastic volatility for Levy processes. *Mathematical Finance, 13*, 345–382.

Chan, W. H., & Maheu, J. M. (2002). Conditional jump dynamics in stock market returns. *Journal of Business & Economic Statistics, 20*, 377–389.

Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica, 41*, 135–155.

Ding, Z., Granger, C. W. J., & Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance, 1*, 83–106.

Duffie, D., Pan, J., & Singleton, K. (2000). Transform analysis and option pricing for affine jump-diffusions. *Econometrica, 68*, 1343–1376.

Engle, R. F., & Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *Journal of Finance, 48*, 1749–1778.

Eraker, B. (2004). Do stock prices and volatility jump? Reconciling evidence from spot and option prices. *Journal of Finance, 59*, 1367–1404.

Eraker, B., Johannes, M., & Polson, N. (2003). The impact of jumps in volatility and returns. *Journal of Finance, 53*, 1269–1300.

Geman, H., Madan, D. B., & Yor, M. (2001). Times changes for Lévy processes. *Mathematical Finance, 11*, 79–96.

Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between expected value and the volatility nominal excess return on stocks. *Journal of Finance, 48*, 1997–1801.

Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika, 57*, 97–109.

Jorion, P. (1988). On jump processes in the foreign exchange and stock markets. *The Review of Financial Studies, 1*, 427–445.

Madan, D. B., & Milne, F. (1991). Option pricing with VG martingale components. *Mathematical Finance, 1*(4), 39–55.

Madan, D. B., & Seneta, E. (1990). The variance gamma (V.G.) model for share market returns. *Journal of Business, 63*(4), 511–524.

Madan, D. B., Carr, P., & Chang, E. C. (1998). The variance gamma process and option pricing. *European Finance Review, 2*, 79–105.

Maheu, J. M., & McCurdy, T. H. (2004). News arrival, jump dynamics and volatility components for individual stock returns. *Journal of Finance, 59*(2), 755–793.

McCulloch, C. E. (1997). Maximum likelihood algorithms for generalized linear mixed models. *Journal of the American Statistical Association, 92*, 162–170.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equations of state calculations by fast computing machines. *Journal of Chemical Physics, 21*, 1087–1091.

Nelson, D. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica, 59*, 347–370.

Pan, J. (2002). The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of Financial Economics, 63*, 3–50.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business, 53*, 61–65.

Ross, S. A. (1989). Information and volatility: The no-arbitrage Martingale approach to timing and resolution irrelevancy. *Journal of Finance, 47*, 1–17.

Wei, G. C. G., & Tanner, M. A. (1990). A Monte Carlo implementation of the EM algorithm and the Poor Man's data augmentation algorithms. *Journal of the American Statistical Association, 85*, 699–704.

Risk-Averse Portfolio Optimization via Stochastic Dominance Constraints

83

Darinka Dentcheva and Andrzej Ruszczynski

Contents

Abstract

We consider the problem of constructing a portfolio of finitely many assets whose return rates are described by a discrete joint distribution. We present a new approach to portfolio selection based on stochastic dominance.

The portfolio return rate in the new model is required to stochastically dominate a random benchmark. We formulate optimality conditions and duality relations for these models and construct equivalent optimization models with

D. Dentcheva (✉)
Department of Mathematical Sciences, Stevens Institute of Technology, Hoboken, NJ, USA
e-mail: ddentche@stevens.edu

A. Ruszczynski
Department of Management Science and Information Systems, Rutgers, The State University of New Jersey, Piscataway, NJ, USA
e-mail: rusz@business.rutgers.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_83,

utility functions. Two different formulations of the stochastic dominance constraint, primal and inverse, lead to two dual problems which involve von Neumann–Morgenstern utility functions for the primal formulation and rank-dependent (or dual) utility functions for the inverse formulation. We also discuss the relations of our approach to value at risk and conditional value at risk. Numerical illustration is provided.

Keywords

Portfolio optimization • Stochastic dominance • Stochastic order • Risk • Expected utility • Duality • Rank-dependent utility • Yaari's dual utility • Value at risk • Conditional value at risk

83.1 Introduction

The problem of optimal portfolio selection is subject of major theoretical and computational studies in finance. A fundamental issue while dealing with uncertain outcomes is a theoretically sound approach to their comparison.

The theory of stochastic orders plays a fundamental role in economics (see Mosler and Scarsini 1991; Whitmore and Findlay 1978). These are relations which induce partial order in the space of real random variables in the following way. A random variable R dominates the random variable Y if $\mathbb{E}[u(R)] \geq \mathbb{E}[u(Y)]$ for all functions $u(\cdot)$ from certain set of functions, called the generator of the order. The concept of *stochastic dominance* is very popular and widely used in economics and finance because of its relation to models of risk-averse preferences (Fishburn 1964). It originated from the theory of majorization (Hardy et al. 1934; Marshall and Olkin 1979) for the discrete case and was later extended to general distributions (Quirk and Saposnik 1962; Hadar and Russell 1969; Rothschild and Stiglitz 1969). Stochastic dominance of second order is defined by the set of nondecreasing concave functions: a random variable R dominates another random variable Y in the second order if $\mathbb{E}[u(R)] \geq \mathbb{E}[u(Y)]$ for all nondecreasing concave functions $u(\cdot)$ for which these expected values are finite. Thus, no risk-averse decision maker will prefer a portfolio with return rate Y over a portfolio with return rate R.

A popular approach is the utility optimization approach. Von Neumann and Morgenstern in their book (von Neumann and Morgenstern 1944) developed the *expected utility theory*: for every rational decision maker there exists a utility function $u(\cdot)$ such that the decision maker prefers outcome R over outcome Y if and only if $\mathbb{E}[u(R)] \geq \mathbb{E}[u(Y)]$. This approach can be implemented also very efficiently; however, it is almost impossible to elicit the utility function of a decision maker explicitly. More difficulties arise when a group of decision makers with different utility functions who have to reach a consensus. Recently, the *dual utility theory* (or *rank-dependent expected utility theory*) has attracted much attention in economics. This approach was first presented in (Quiggin 1982) and later rediscovered in a special case in (Yaari 1987). From a different system of axioms, than those of von Neumann and Morgenstern, one derives that every decision maker has a certain *rank-dependent utility function* w: $[0, 1] \to \mathbb{R}$. Then a nonnegative outcome R is preferred over a nonnegative outcome Y, if and only if

$$-\int_0^1 w(p) dF_{(-1)}(R;p) \geq -\int_0^1 w(p) dF_{(-1)}(Y;p), \tag{83.1}$$

where $F_{(-1)}(R;\cdot)$ is the inverse distribution function of R. For a comprehensive treatment of the rank-dependent utility theory, we refer to (Quiggin 1993), and for its application in actuarial mathematics, see (Wang et al. 1997; Wang and Yong 1998).

Another classical approach, pioneered by (Markowitz 1952, 1959, 1987), is the mean–risk approach, which compares the portfolios with respect to two characteristics. One is the expected return rate (*the mean*) and another one is *the risk*, which is given by some scalar measure of the uncertainty of the portfolio return rate. The mean–risk approach recommends the selection of Pareto-efficient portfolios with respect to these two criteria. In a mean–risk portfolio model, we combine these criteria by specifying some parameter as a trade-off between them. As a parametric optimization problem, the mean–risk model can be solved numerically very efficiently, which makes this approach very attractive (Konno and Yamazaki 1991; Ruszczynski and Vanderbei 2003).

In this chapter we formulate a model for risk-averse portfolio optimization and demonstrate its relation to the expected utility approach and to rank-dependent utility approach. We optimize the portfolio performance under an additional constraint that the portfolio return rate stochastically dominates a benchmark return rate, for example, the return rate of an index. The model is based on our earlier publications (Dentcheva and Ruszczynski 2003a, b, c, 2004a, c) where we have introduced a new model for risk-averse optimization. This approach has a fundamental advantage over mean–risk models and utility function models. All data for our model are readily available. In mean–risk models the choice of the risk measure has an arbitrary character, and it is difficult to argue for one measure against another. Similarly, optimization of expected utility requires the form of the utility function to be specified. Our analysis, departing from the benchmark outcome, generates implied utility function of the decision maker. It is implicitly defined by the benchmark used and by the problem under consideration. We provide two problem formulations in which the stochastic dominance has a primal or inverse form: a Lorenz curve. The primal form has a dual problem in terms of expected utility functions, and the inverse form has a dual problem in terms of rank-dependent utility functions. In this way our model provides also a link between this two competing economic approaches. Duality relations with coherent measures of risk are explored in (Dentcheva and Ruszczynski 2008).

83.2 The Portfolio Problem

Let $R_1, R_2, \ldots, R_n$ be random return rates of assets $1, 2, \ldots, n$. We assume that $\mathbb{E}[|R_j|] < \infty$ for all $j = 1, \ldots, n$.

Our aim is to invest our capital in these assets in order to obtain some desirable characteristics of the total return rate on the investment. Denoting by $x_1, x_2, \ldots, x_n$ the fractions of the initial capital invested in assets $1, 2, \ldots, n$, we can easily derive the formula for the total return rate:

$$R(x) = R_1x_1 + R_2x_2 + \cdots + R_nx_n. \tag{83.2}$$

Clearly, the set of possible asset allocations can be defined as follows:

$$X = \left\{x \in \mathbb{R}^n : x_1 + x_2 + \cdots + x_n = 1, x_j \geq 0, j = 1, 2, \ldots, n\right\}.$$

In some applications one may introduce the possibility of *short positions*, i.e., allow some x_j's to become negative. Other restrictions may limit the exposure to particular assets or their groups, by imposing upper bounds on the x_j's or on their partial sums. One can also limit the absolute differences between the x_j's and some reference investments $\bar{x}_j$, which may represent the existing portfolio, etc. Our analysis does not depend on the detailed way this set is defined; we only use the fact that it is a convex polyhedron. All modifications discussed above define some convex polyhedral feasible sets and are, therefore, covered by our approach.

The main difficulty in formulating a meaningful portfolio optimization problem is the definition of the preference structure among feasible portfolios. If we use only the mean return rate $\mathbb{E}[R(x)]$, then the resulting optimization problem has a trivial and meaningless solution: invest everything in assets that have the maximum expected return rate. For these reasons the practice of portfolio optimization resorts usually to two approaches.

In the first approach we associate with portfolio x some dispersion measure $\rho(R(x))$ representing the variability of the return rate $R(x)$. In the classical Markowitz model (Markowitz 1952, 1959, 1987), the function $\rho(R(x))$ is the variance of the return rate,

$$\rho(R(x)) = \mathbb{V}[R(x)],$$

but many other measures are possible here as well.

The mean–risk portfolio optimization problem is formulated as follows:

$$\max_{x \in X} \mathbb{E}[R(x)] - \lambda\rho(R(x)). \tag{83.3}$$

Here, λ is a nonnegative parameter representing our desirable exchange rate of mean for risk. If $\lambda = 0$, the risk has no value, and the problem reduces to the problem of maximizing the mean. If $\lambda > 0$ we look for a compromise between the mean and the risk. Alternatively, one can minimize the risk function $\rho(x)$, while fixing the expected return rate $\mathbb{E}[R(x)]$ at some value m, and consider a family of problems parametrized by m. The reader is referred to the book (Elton et al. 2006) for the modern perspective on mean–risk analysis in portfolio theory.

The general question of constructing mean–risk models which are in harmony with the stochastic dominance relations has been the subject of the analysis of the recent papers (Ogryczak and Ruszczynski 1999, 2001, 2002). We have identified there several primal risk measures, most notably central semideviations, and dual risk measures, based on the Lorenz curve, which are consistent with the stochastic dominance relations.

The second approach is to select a certain *utility function* $u : \mathbb{R} \to \mathbb{R}$ and to formulate the following optimization problem

$$\max_{x \in X} \mathbb{E}[u(R(x))]. \tag{83.4}$$

It is usually required that the function $u(\cdot)$ is concave and nondecreasing, thus representing preferences of a risk-averse decision maker (Fishburn 1964, 1970).

Recently, a *dual* (*rank-dependent*) *utility* model attracts much attention. It is based on distorting the cumulative probability distribution of the random variable $R(x)$ rather than applying a nonlinear function $u(\cdot)$ to the realizations of $R(x)$. The corresponding problem has the following form

$$\max_{x \in X} \int_0^1 F_{(-1)}(R(x), p)\,dw(p). \tag{83.5}$$

Here $F_{(-1)}(R(x),p)$ is the p-quantile of the random variable $R(x)$, and $w(\cdot)$ is the *rank-dependent utility function*, which distorts the probability distribution. We discuss this in Sect. 83.3.2.

The challenge in both utility approaches is to select the appropriate utility function or rank-dependent utility function that represents our preferences and whose application leads to nontrivial and meaningful solutions of Eqs. 83.4 and 83.5.

In this chapter we propose an alternative approach, by introducing a comparison to a benchmark return rate into our optimization problem. The comparison is based on the stochastic dominance relation. More specifically, we consider only portfolios whose return rates stochastically dominates a certain benchmark return rate.

83.3 Stochastic Dominance

83.3.1 Direct Forms

In the stochastic dominance approach, random return rates are compared by a point-wise comparison of some performance functions constructed from their distribution functions. For a real random variable V, its first performance function is defined as the right-continuous cumulative distribution function of V:

$$F_1(V;\eta) = \mathbb{P}\{V \le \eta\} \quad \text{for } \eta \in \mathbb{R}.$$

A random return V is said (Lehmann 1955; Quirk and Saposnik 1962) to *stochastically dominate* another random return S in the first order, denoted $V \succeq_{(1)} S$, if

$$F_1(V;\eta) \leq F_2(S;\eta) \quad \text{for all } \eta \in \mathbb{R}.$$

We can say that V is "stochastically larger" than S, because it takes values lower than η with smaller (or equal) probabilities than S, no matter what the target η is.

The second performance function F_2 is given by areas below the distribution function F,

$$F_2(V;\eta) = \int_{-\infty}^{\eta} F_1(V;\xi)\,d\xi \quad \text{for } \eta \in \mathbb{R},$$

and defines the weak relation of the *second-order stochastic dominance* (SSD). That is, random return V stochastically dominates S in the second order, denoted $V \succeq_{(2)} S$, if

$$F_2(V;\eta) \leq F_2(S;\eta) \quad \text{for all } \eta \in \mathbb{R}$$

(see Hadar and Russell 1969; Rothschild and Stiglitz 1969).

We can express the function $F_2(V;\cdot)$ as the expected shortfall (see, e.g., Levy 2006; Ogryczak and Ruszczynski 1999): for each target value η, we have

$$F_2(V;\eta) = \mathbb{E}\big[(\eta - V)_+\big], \tag{83.6}$$

where $(\eta - V)_+ = \max(\eta - V, 0)$. The function $F_{(2)}(V;\cdot)$ is continuous, convex, nonnegative, and nondecreasing. It is well defined for all random variables V with finite expected value. Due to this representation, the second-order stochastic dominance relation $V \succeq_{(2)} S$ can be equivalently characterized by the system of inequalities on the expected shortfall below any target η:

$$\mathbb{E}\big[(\eta - V)_+\big] \leq \mathbb{E}\big[(\eta - S)_+\big] \quad \text{for all } \eta \in \mathbb{R}. \tag{83.7}$$

Also, we obtain an equivalent characterization in terms of the expected utility theory of von Neumann and Morgenstern (see, e.g., Hanoch and Levy 1969; Levy 2006; Müller and Stoyan 2002):

(i) For any two random variables V, S, the relation $V \succeq_{(1)} S$ holds true if and only if for all nondecreasing functions $u(\cdot)$ defined on $\mathbb{R}$, we have

$$\mathbb{E}[u(V)] \geq \mathbb{E}[u(S)]. \tag{83.8}$$

(ii) For any two random variables V, S with finite expectations, the relation $V \succeq_{(2)} S$ holds true if and only if Eq. 83.8 is satisfied for all nondecreasing concave functions $u(\cdot)$.

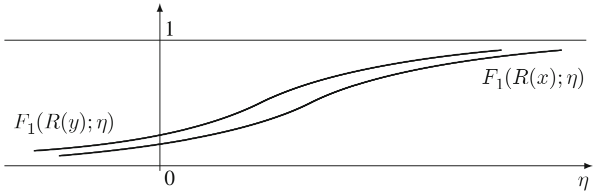


Fig. 83.1 First-order stochastic dominance $R(x) \succ_{(1)} R(y)$

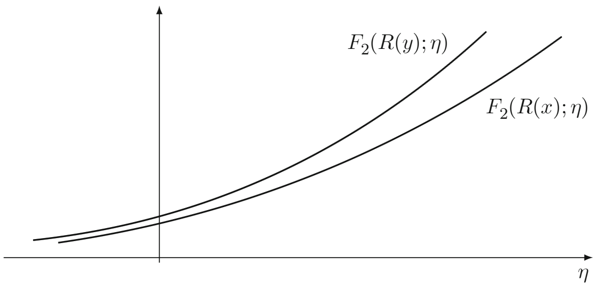


Fig. 83.2 Second-order dominance $R(x) \succ_{(2)} R(y)$

In the context of portfolio optimization, we consider stochastic dominance relations between random return rates defined by Eq. 83.2. Thus, we say that portfolio x *dominates* portfolio y *in the first order*, if

$$F_1(R(x);\eta) \leq F_1(R(y);\eta) \quad \text{for all } \eta \in \mathbb{R}.$$

This is illustrated in Fig. 83.1.

Similarly, we say that x *dominates* y *in the second-order* ($R(x) \succ_{(2)} R(y)$) if

$$F_2(R(x);\eta) \leq F_2(R(y);\eta) \quad \text{for all } \eta \in \mathbb{R}.$$

Recall that the individual return rates R_j have finite expected values, and thus the function $F_2(R(x);\cdot)$ is well defined. The second-order relation is illustrated in Fig. 83.2.

83.3.2 Inverse Forms

Let us consider the inverse model of stochastic dominance, frequently referred to as *Lorenz dominance*. For a real random variable V (e.g., a random return rate), we define the left-continuous inverse of the cumulative distribution function $F_1(V;\cdot)$ as follows:

$$F_{(-1)}(V;p) = \inf\{\eta : F_1(V;\eta) \geq p\} \quad \text{for} \quad 0 < p < 1.$$

Given $p \in (0.1)$, the number $q = q(V;p)$ is called a *p-quantile* of the random variable V if

$$\mathbb{P}\{V < q\} \leq p \leq \mathbb{P}\{V < q\}.$$

For $p \in (0.1)$ the set of p-quantiles is a closed interval, and $F_{(-1)}(V;p)$ represents its left end.

Directly from the definition of the first-order dominance, we see that

$$V \succ_{(1)} S \Leftrightarrow F_{(-1)}(V;p) \geq F_{(-1)}(S;p) \quad \text{for all} \quad 0 < p < 1. \tag{83.9}$$

The first-order dominance constraint can be interpreted as a continuum of probabilistic (chance) constraints, studied in stochastic optimization (see, Prekopa 2003; Dentcheva 2005).

Our analysis uses the *absolute Lorenz function* $F_{(-2)}(V;\cdot) : [0,1] \to \mathbb{R}$, introduced in (Lorenz 1905). It is defined as the cumulative quantile:

$$F_{(-2)}(V;p) = \int_0^p F_{(-1)}(V;t)dt \quad \text{for} \quad 0 < p < 1, \tag{83.10}$$

$F_{(-2)}(V;0) = 0$.

Similarly to $F_2(V;\cdot)$ the function $F_{(-2)}(V;\cdot)$ is well defined for any random variable V, which has a finite expected value. We notice that

$$F_{(-2)}(V;1) = \int_0^1 F_{(-1)}(V;t)dt = \mathbb{E}[V].$$

By construction, the Lorenz function is convex. Lorenz functions are commonly used for inequality ordering of positive random variables, relative to their (positive) expectations (see Arnold 1980; Gastwirth 1971; Muliere and Scarsini 1989). Such a Lorenz function, $p \mapsto F_{(-2)}(V;p)/\mathbb{E}[V]$, is convex and nondecreasing. The absolute Lorenz function, though, is not monotone when negative outcomes occur.

It is well known (see, e.g., Ogryczak and Ruszczynski 2002) that we may fully characterize the second-order dominance relation by using the function $F_{(-2)}(V;\cdot)$:

$$V \succ_{(2)} S \Leftrightarrow F_{(-2)}(V;p) \geq F_{(-2)}(S;p) \text{ for all } 0 \leq p \leq 1. \tag{83.11}$$

This characterization of stochastic dominance by Lorenz functions is widely used in economics and statistics.

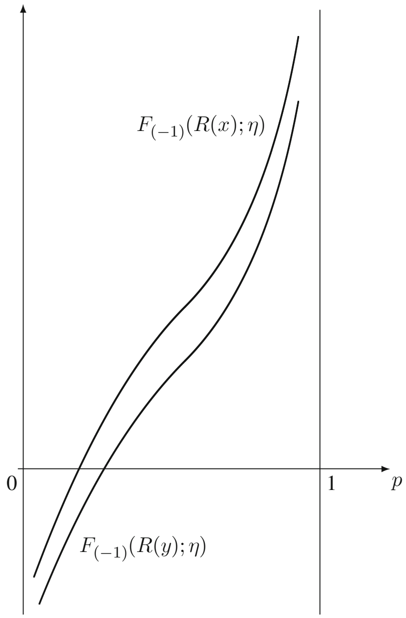
We now provide an equivalent characterization by rank-dependent utility functions. It is analogous to the characterization by expected utility functions.

In the chapter (Dentcheva and Ruszczynski 2006b) the following characterization has been shown:

(i) For any two random variables V, S, the relation $V \succ_{(1)} S$ holds true if and only if for all nondecreasing functions $w(\cdot)$ defined on [0,1], we have

$$\int_0^1 F_{(-1)}(V;p)dw(p) \geq \int_0^1 F_{(-1)}(S;p)dw(p). \tag{83.12}$$

Fig. 83.3 First-order stochastic dominance $R(x) \succ_{(1)} R(y)$ in the inverse form



(ii) For any two random variables V, S with finite expectations, the relation $V \succ_{(2)} S$ holds true if and only if Eq. 83.12 is satisfied for all nondecreasing concave functions $w(\cdot)$.

The functions $w(\cdot)$ appearing in this characterization are rank-dependent (dual) utility functions.

In the context of portfolio optimization, we consider stochastic dominance relations between random return rates defined by Eq. 83.2. Thus, we say that portfolio x *dominates* portfolio y *in the first order*, if

$$F_{(-1)}(R(x)); \mathrm{p}) \geq F_{(-1)}(R(y); p) \quad \text{for all } p \in (0, 1).$$

This is illustrated in Fig. 83.3.

Similarly, we say that x *dominates* y *in the second-order* ($R(x) \succ_{(2)} R(y)$), if

$$F_{(-2)}(R(x); p) \geq F_{(-2)}(R(y); p) \quad \text{for all } p \in [0, 1]. \tag{83.13}$$

Recall that the individual return rates R_j have finite expected values, and thus the function $F_{(-2)}(R(x); \cdot)$ is well defined. The second-order relation is illustrated in Fig. 83.4.

Second-order dominance $R(x) \succ_{(2)} R(y)$ in the inverse form.

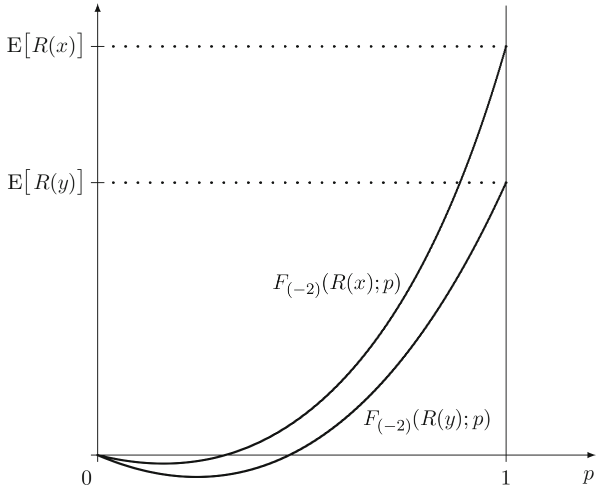


Fig. 83.4 Second order dominance in the inverse form

83.3.3 Relations to Value at Risk and Conditional Value at Risk

There are fundamental relations between the concepts of value at risk (VaR) and conditional value at risk (CVaR) and the stochastic dominance constraints. The VaR constraint in the portfolio context is formulated as follows. We define the loss rate $L(x) = - R(x)$. We specify the maximum fraction ω_p of the initial capital allowed for risk exposure at risk level $p \in (0,1)$, and we require that

$$\mathbb{P}\big[L(x) \leq \omega_p\big] \geq 1 - p.$$

Denoting by $\mathrm{VaR}_p(L(x))$ the left $(1 - p)$-quantile of the random variable $L(x)$, we can equivalently formulate the VaR constraint as

$$\mathrm{VaR}_p(L(x)) \leq \omega_p.$$

The first-order stochastic dominance relation between two portfolios is equivalent to the continuum of VaR constraints. Portfolio x dominates portfolio y in the first order, if

$$\mathrm{VaR}_p(L(x)) \leq \mathrm{VaR}_p(L(y)) \text{ for all } p \in (0, 1).$$

The CVaR at level p, roughly speaking, has the following form:

$$\mathrm{CVaR}_p(L(x)) = \mathbb{E}\big[L(x) | L(x) \geq \mathrm{VaR}_p(L(x))\big].$$

This formula is precise if $\mathrm{VaR}_p(L(x))$ is not an atom of the distribution of $L(x)$. More precisely we express it as follows:

$$\mathrm{CVaR}_p(L(x)) = \frac{1}{p}\int_0^p \mathrm{VaR}_t(L(x))dt.$$

We note that

$$\mathrm{CVaR}_p(L(x)) = -\frac{1}{p}F_{(-2)}(R(x), p). \tag{83.14}$$

Another description uses extremal properties of quantiles and equivalently represents CVaR as follows (see Rockafellar and Uryasev 2000):

$$\mathrm{CVaR}_p(L(x)) = \inf_\eta \left\{ \frac{1}{p}\mathbb{E}\left[(\eta - R(x))_+\right] - \eta \right\}. \tag{83.15}$$

A CVaR constraint on the portfolio x can be formulated as follows:

$$\mathrm{CVaR}_p(L(x)) \le \omega_p. \tag{83.16}$$

Using Eqs. 83.14 and 83.13, we conclude that the second-order stochastic dominance relation for two portfolios x and y is equivalent to the continuum of CVaR constraints:

$$R(x) \succ_{(2)} R(y) \Leftrightarrow \mathrm{CVaR}_p(L(x)) \le \mathrm{CVaR}_p(L(y)) \text{ for all } p \in (0, 1). \tag{83.17}$$

Assume that we compare the performance of a portfolio x with a random benchmark Y (e.g., an index return rate or another portfolio return rate) requiring $R(x) \succ_{(2)} Y$. Then the fraction ω_p of the initial capital allowed for risk exposure at level p is given by the benchmark Y:

$$\omega_p = \mathrm{CVaR}_p(-Y), \quad p \in (0, 1).$$

Assume that Y has a discrete distribution with realizations y_i, $i = 1, \ldots, m$. Then relation (83.7) is equivalent to

$$\mathbb{E}\left[(y_i - R(x))_+\right] \le \mathbb{E}\left[(y_i - Y)_+\right], \quad \mathrm{i} = 1, \ldots, m. \tag{83.18}$$

This result *does not* imply that the continuum of CVaR constraints (83.17) can be replaced by finitely many constraints of form

$$\mathrm{CVaR}_{pi}(R(x)) \ge \mathrm{CVaR}_{pi}(Y), \quad i = 1, \ldots, m.$$

with some fixed probabilities p_i, $i = 1, \ldots, m$. The reason is that we do not know at which probability levels the CVaR constraints have to be imposed.

83.4 The Dominance-Constrained Portfolio Problem

83.4.1 Direct Formulation

The starting point for our model is the assumption that a benchmark random return rate Y having a finite expected value is available. It may have the form of $Y = R(\bar{z})$, for some benchmark portfolio $\bar{z}$. It may be an index or our current portfolio. Our intention is to have the return rate of the new portfolio, $R(x)$, preferable over Y. Therefore, we introduce the following extension of the optimization problem (83.3):

$$\begin{array}{c}\max \mathbb{E}\left[R(x)\right] - \lambda\rho(R(x)) \\ \text{subject to}\end{array} \tag{83.19}$$

$$R(x) \succ_{(2)} Y, \tag{83.20}$$

$$x \in X. \tag{83.21}$$

Similarly to Eq. 83.3, we optimize a mean–risk objective function, but we introduce a constraint that the portfolio return dominates a benchmark. Even when $\lambda = 0$ and we maximize just the expected value of the return rate, our model will still lead to nontrivial solutions, due to the presence of the dominance constraint (83.20).

To increase flexibility of model (83.19–83.21), we may also allow a uniform shift of $R(x)$ by a constant c, as in the following model:

$$\begin{array}{l}\max \mathbb{E}[R(x)] - \lambda\rho(R(x)) - \delta c \\ \quad\text{subject to} \\ \quad R(x) + c \succ_{(2)} Y, \\ \quad x \in X.\end{array}$$

Here $\delta > 0$ can be interpreted a cost of the shift c. Observe that the shift c may also become negative, in which case we are rewarded for uniformity of dominating Y. The shift c may be interpreted as an additional cash added to the return, and δ is the interest to be paid when the loan is paid back.

To simplify the derivations, from now on we focus on the simplest formulation of the dominance-constrained problem:

$$\begin{array}{c}\max \mathbb{E}[R(x)] \\ \text{subject to}\end{array} \tag{83.22}$$

$$R(x) \succ_{(2)} Y, \tag{83.23}$$

$$x \in X. \tag{83.24}$$

We can observe the first advantage of our problem formulation: all data in it are readily available. Moreover, the set defined by Eq. 83.23 is convex (see Dentcheva and Ruszczynski 2003c, 2004a, b).

Let us assume now that Y has a discrete distribution with realizations y_i attained with probabilities π_i, $i = 1, \ldots, m$. We also assume that the return rates have a discrete joint distribution with realizations r_{jt}, $t = 1, \ldots, T$, $j = 1, \ldots, n$, attained with probabilities p_t, $t = 1, 2, \ldots, T$. Then the formulation of the stochastic dominance relation (83.23) resp. (83.18) simplifies even further. Introducing variables s_{it} representing the shortfall of $R(x)$ below y_i in realization t, $i = 1, \ldots, m$, $t = 1, \ldots, T$, we can formulate problem (83.22–83.24) as follows:

$$\max \sum_{t=1}^{T} pt \sum_{j=1}^{n} x_j r_{jt} \tag{83.25}$$

subject to

$$\sum_{j=1}^{n} x_j r_{jt} + s_{it} \geq y_i, \quad i = 1, \ldots, m, \quad t = 1, \ldots, T, \tag{83.26}$$

$$\sum_{t=1}^{T} p_t s_{it} \leq \sum_{k=1}^{m} \pi_k (y_i - y_k)_+, \quad i = 1, \ldots, m, \tag{83.27}$$

$$s_{it} \geq 0, \quad i = 1, \ldots, m, \quad t = 1, \ldots, T. \tag{83.28}$$

$$x \in X. \tag{83.29}$$

Indeed, or every feasible point x of Eqs. 83.22, 83.23, and 83.24, setting

$$s_{it} = \max\left(0, y_i - \sum_{j=1}^{n} x_j r_{jt}\right), \quad i = 1, \ldots, m, \quad t = 1, \ldots, T,$$

we obtain a feasible pair (x, s) for Eqs. 83.26, 83.27, 83.28, and 83.29. Conversely, for any feasible pair (x, s) for Eqs. 83.26, 83.27, 83.28, and 83.29, inequalities Eqs. 83.26 and 83.28 imply that

$$s_{it} \geq \max\left(0, y_i - \sum_{j=1}^{n} x_j r_{jt}\right), \quad i = 1, \ldots, m, \quad t = 1, \ldots, T.$$

Taking the expected value of both sides and using Eq. 83.27, we obtain

$$F_2(R(x); y_i) \leq F_2(Y; y_i), \quad i = 1, \ldots, m.$$

Therefore, problem (83.22–83.24) is equivalent to problem (83.25–83.29).

If the set X is a convex polyhedron, problem (83.25–83.29) becomes a large-scale linear programming problem. It may be solved by general-purpose linear programming solvers. However, the size of the problem increases dramatically with the number of assets n, their return realizations T, and benchmark realizations m, which makes it impractical for even moderate dimensions (in thousands). For the purpose of solving these problems, we developed a specialized decomposition method presented in (Dentcheva and Ruszczynski 2006a).

83.4.2 Inverse Formulation

Assume that the return rates have a joint discrete distribution realizations r_{jt}, $t = 1$, T, $j = 1,\ldots,n$, attained with probabilities p_t, $t = 1, 2,\ldots,T$. Moreover, we assume that *all probabilities* p_t are equal, that is, $p_t = 1/T$, $t = 1,\ldots,T$. This is the case of empirical distributions. Correspondingly, we assume that Y has a discrete distribution with $m = T$ equally probable realizations y_t, $t = 1,\ldots,T$. We use the symbol $R_{[t]}(x)$ to denote the ordered realizations of $R(x)$, that is,

$$R_{[1]}(x) \leq R_{[2]}(x) \leq \cdots \leq R_{[T]}(x).$$

Since $R(x)$ has a discrete distribution, the functions $F_2(R(x);\cdot)$ and $F_{(-2)}(R(x);\cdot)$ are piecewise linear. Owing to the fact that all probabilities p_t are equal, the break points of $F_{(-2)}(R(x);\cdot)$ occur at t/T, for $t = 0,1,\ldots,m$. The same applies to $F_{(-2)}(Y;\cdot)$. It follows from Eq. 83.13 that the stochastic dominance constraint (83.23) can be equivalently expressed as

$$F_{(-2)}\left(R(x);\frac{t}{T}\right) \geq F_{(-2)}\left(Y;\frac{t}{T}\right), \quad t = 1, \ldots, T.$$

Note that $F_{(-2)}(R(x);0) = F_{(-2)}(Y;0) = 0$. We have

$$F_{(-2)}(R(x);\frac{t}{T}) = \frac{1}{T}\sum_{k=1}^{t} R_{[k]}(x), \quad t = 1, \ldots, T.$$

Therefore problem (83.22–83.24) can be written with an equivalent inverse form of the dominance constraint:

$$\begin{array}{c} \max\ \mathbb{E}[R(x)] \\ \text{subject to} \end{array} \tag{83.30}$$

$$\sum_{k=1}^{t} R_{[k]}(x) \geq \sum_{k=1}^{t} y_{[k]}, \quad t = 1, \ldots, T, \tag{83.31}$$

$$x \in X. \tag{83.32}$$

It was shown in (Ogryczak and Ruszczynski 2002) that the function $x \mapsto \sum_{k=1}^{t} R_{[k]}(x)$ is concave and positively homogeneous. It is also polyhedral. Therefore, Eq. 83.31 are convex polyhedral constraints. If the set X is a convex polyhedron, problem (83.30–83.32) has an equivalent linear programming formulation.

All these transformations are possible due to the crucial assumption that the probabilities of all elementary events are equal. If they are not equal, the break points of the function $F_{(-2)}(R(x);\cdot)$ depend on x, and therefore inequality (83.13) cannot be reduced to finitely many convex inequalities. This is in contrast to the primal formulation, where the discreteness of Y alone was sufficient to reduce the stochastic dominance constraint to finitely many convex inequalities.

We have to observe that the quantile formulation (83.31) of stochastic dominance constraints is more involved than the primal formulation and requires more sophisticated computational methods. Using Eq. 83.31 directly would require employing nonsmooth optimization methods to solve problem (83.30–83.32). Equivalent formulation with linear constraints has very many constraints, because of the large number of pieces of the function $x \mapsto \sum_{k=1}^{t} R_{[k]}(x)$. Still, in (Dentcheva and Ruszczynski 2010) we have developed a highly efficient cutting plane method, which significantly outperforms direct approaches.

83.5 Optimality and Duality

83.5.1 Primal Form

From now on we assume that the probability distributions of the return rates are discrete with finitely many realizations r_{jt}, $t = 1, \ldots, T$, $j = 1, \ldots, n$, attained with probabilities p_t, $t = 1, 2, \ldots, T$. We also assume that there are finitely many ordered realizations of the benchmark outcome $Y : y_1 < y_2 < \cdots < y_m$. The probabilities of these realizations are denoted by π_i, $i = 1,\ldots,m$. We also assume that the set X is compact.

We define the set U of functions $u{:}\mathbb{R} \to \mathbb{R}$ satisfying the following conditions:

- $u(\cdot)$ is concave and nondecreasing.
- $u(\cdot)$ is piecewise linear with break points y_i, $i = 1, \ldots, m$.
- $u(t) = 0$ for all $t \geq y_m$.

It is evident that U is a convex cone.

Let us define the function $L{:}\mathbb{R}^n \times \mathrm{U} \to \mathbb{R}$ as follows:

$$L(x, u) = \mathbb{E}[R(x) + u(R(x)) - u(Y)]. \tag{83.33}$$

It will play for problem (83.22–83.24) a similar role to that of a Lagrangian. It is well defined, because for every $u \in \mathrm{U}$ and every $x \in \mathbb{R}^n$, the expected value $\mathbb{E}[u(R(x))]$ exists and is finite.

The following theorem has been proved in a more general version in (Dentcheva and Ruszczynski 2003c) (see also Dentcheva and Ruszczynski 2006a).

Theorem 1 *If $\hat{x}$ is an optimal solution of* Eqs. 83.22, 83.23, *and* 83.24, *then there exists a function $\hat{u} \in \mathrm{U}$ such that*

$$L(\hat{x}, \hat{u}) = \max_{x \in X} L(x\hat{u}) \tag{83.34}$$

$$\mathbb{E}[\hat{u}(R(\hat{x}))] = \mathbb{E}[\hat{u}(Y)]. \tag{83.35}$$

Conversely, if for some function $\hat{u} \in \mathrm{U}$ an optimal solution $\hat{x}$ of Eq. 83.34 satisfies Eqs. 83.23 and 83.35, then $\hat{x}$ is an optimal solution of Eqs. 83.22, 83.23, and 83.24.

We can also develop duality relations for our problem. With the function (83.33) we can associate the dual function

$$D(u) = \max_{x \in X} L(x, u).$$

We are allowed to write the maximization operation here, because the set X is compact and $L(\cdot,u)$ is continuous.

The dual problem has the following form:

$$\min_{u \in \mathcal{U}} D(u). \tag{83.36}$$

The set U is a closed convex cone and $D(\cdot)$ is a convex function, so Eq. 83.36 is a convex optimization problem.

Theorem 2 *Assume that* Eqs. 83.22, 83.23, *and* 83.24 *has an optimal solution. Then problem* (83.36) *has an optimal solution, and the optimal values of both problems coincide. Furthermore, the set of optimal solutions of* Eq. 83.36 *is the set of functions* $\hat{u} \in \mathcal{U}$ *satisfying* Eqs. 83.34 *and* 83.35 *for an optimal solution* $\hat{x}$ *of* Eqs. 83.22, 83.23, *and* 83.24.

Note that all constraints of our problem are linear or convex polyhedral, and therefore we do not need any constraint qualification conditions here.

The "Lagrange multiplier" u is directly related to the expected utility theory of von Neumann and Morgenstern. We have established earlier that the second-order stochastic dominance relation is equivalent to Eq. 83.8 for *all* utility functions in U. Our result shows that one of them, $\hat{u}(\cdot)$, assumes the role of a Lagrange multiplier associated with Eq. 83.23. A point $\hat{x}$ is a solution to Eqs. 83.22, 83.23, and 83.24 if there exists a utility function $\hat{u}(\cdot)$ such that $\hat{x}$ maximizes over X the objective function E[$R(x)$] augmented with this dual utility. We see that the optimization problem in Eq. 83.34 is equivalent to

$$\max_{x \in X} \mathbb{E}[v(R(x))], \tag{83.37}$$

where $v(\eta) = \eta + u(\eta)$. At the optimal solution the function $\hat{v}(\eta) = \eta + \hat{u}(\eta)$ is the implied utility function. It attaches higher penalty to smaller realizations of $R(x)$ (bigger realizations of $L(x)$). By maximizing $L(R(x),u)$ we look for x such that the left tail of the distribution of $R(x)$ is thin.

It is important to stress that the optimal function $\hat{u}(\cdot)$ is piecewise linear, with break points at the realizations $y_1, \ldots, y_m$ of the benchmark Y. Therefore, the dual problem has also an equivalent linear programming formulation.

83.5.2 Inverse Form

In addition to the assumption that all involved distributions are discrete, we also assume that all probabilities p_t are equal and that $m = T$.

We introduce the set W of concave nondecreasing functions w:[0,1] $\to \mathbb{R}$. It is evident that W is a convex cone.

Recall the identity

$$\mathbb{E}[R(x)] = \int_0^1 F_{(-1)}(R(x); p)dp.$$

Let us define the function $\Phi : X \times \mathcal{W} \to \mathbb{R}$ as follows:

$$\Phi(x, w) = \int_0^1 F_{(-1)}(R(x); p)dp + \int_0^1 F_{(-1)}(R(x); p)dw(p) - \int_0^1 F_{(-1)}(Y; p)dw(p). \tag{83.38}$$

It plays a role similar to that of a Lagrangian of Eqs. 83.30, 83.31, and 83.32.

Theorem 3 *If* $\hat{x}$ *is an optimal solution of* Eqs. 83.30, 83.31*, and* 83.32, *then there exists a function* $\hat{w} \in \mathcal{W}$ *such that*

$$\Phi(\hat{x}, \hat{w}) = \max_{x \in X} \Phi(x, \hat{w}) \tag{83.39}$$

$$\int_0^1 F_{(-1)}(R(\hat{x}); p)d\hat{w}(p) = \int_0^1 F_{(-1)}(Y; p)d\hat{w}(p). \tag{83.40}$$

Conversely, if for some function $\hat{w} \in \mathcal{W}$ an optimal solution $\hat{x}$ of Eq. 83.39 satisfies Eqs. 83.31 and 83.40, then $\hat{x}$ is an optimal solution of Eqs. 83.30, 83.31, and 83.32.

We can also develop a duality theory based on Lagrangian Eq. 83.38. For every function $w \in$ W the problem

$$\max_{x \in X} \Phi(x, w) \tag{83.41}$$

is a Lagrangian relaxation of problem (83.30–83.32). Its optimal value, $\psi(w)$, is always greater than or equal to the optimal value of Eqs. 83.30, 83.31, and 83.32.

We define the dual problem as

$$\min_{w \in \mathcal{W}} \psi(w). \tag{83.42}$$

The set W is a closed convex cone, and $\psi(\cdot)$ is a convex function, so problem (83.42) is a convex optimization problem. Duality relations in convex programming yield the following result.

Theorem 4 *Assume that problem* (83.30–83.32) *has an optimal solution. Then problem* (83.42) *has an optimal solution, and the optimal values of both problems coincide. Furthermore, the set of optimal solutions of* Eq. 83.42 *is the set of functions* $\hat{w} \in \mathcal{W}$ *satisfying* Eqs. 83.39 *and* 83.40 *for an optimal solution* $\hat{x}$ *of* Eqs. 83.30, 83.31, *and* 83.32.

The “Lagrange multiplier” w in this case is related to rank-dependent expected utility theory. We have established earlier that the second-order stochastic dominance relation is equivalent to Eq. 83.12 for *all* dual utility functions in W. Our result shows that one of them, $\hat{w}(\cdot)$, assumes the role of a Lagrange multiplier associated with Eq. 83.31. A point $\hat{x}$ is a solution to Eqs. 83.30, 83.31, and 83.32 if there exists a dual utility function $\hat{w}(\cdot)$ such that $\hat{x}$ maximizes over X the objective function E[$R(x)$] augmented with this dual utility. We can transform the Lagrangian Eq. 83.38 in the following way:

$$\begin{aligned}\Phi(X, w) &= \int_0^1 F_{(-1)}(R(x); p)dp + \int_0^1 F_{(-1)}(R(x); p)dw(p) - \int_0^1 F_{(-1)}(Y; p)dw(p)\\ &= \int_0^1 F_{(-1)}(R(x); p)dv(p) - \int_0^1 F_{(-1)}(Y; p)dw(p),\end{aligned}$$

where $v(p) = p + w(p)$. At the optimal solution the function $\hat{v}(p) = p + \hat{w}(p)$ is the quantile utility function implied by the benchmark Y. Since $\int_0^1 F_{(-1)}(Y; p)dw(p)$ is fixed, the problem at the right-hand side of Eq. 83.39 becomes a problem of maximizing the implied rank-dependent expected utility in X. It attaches higher weights to quantiles corresponding to smaller probabilities p. By maximizing $\Phi(R(x), w)$ we look for x such that the left tail of the distribution of $R(x)$ is thin.

Similarly to von Neumann–Morgenstern utility function, it is very difficult to elicit the dual utility function in advance. Our model derives it from a random benchmark.

The optimal function $\hat{w}(\cdot)$ is piecewise linear, with break points at $\frac{t}{T}$, $t = 1,\ldots, T$. Therefore, the dual problem has also an equivalent linear programming formulation. This property, though, is conditioned on the assumption of equal probabilities.

83.6 Numerical Illustration

We have tested our approach on a basket of 719 real-world assets, using 616 possible realizations of their joint return rates (Ruszczynski and Vanderbei 2003). Historical data on weekly return rates in the 12 years from Spring 1990 to Spring 2002 were used as equally likely realizations.

Implied utility functions corresponding to dominance constraints for four benchmark portfolios.

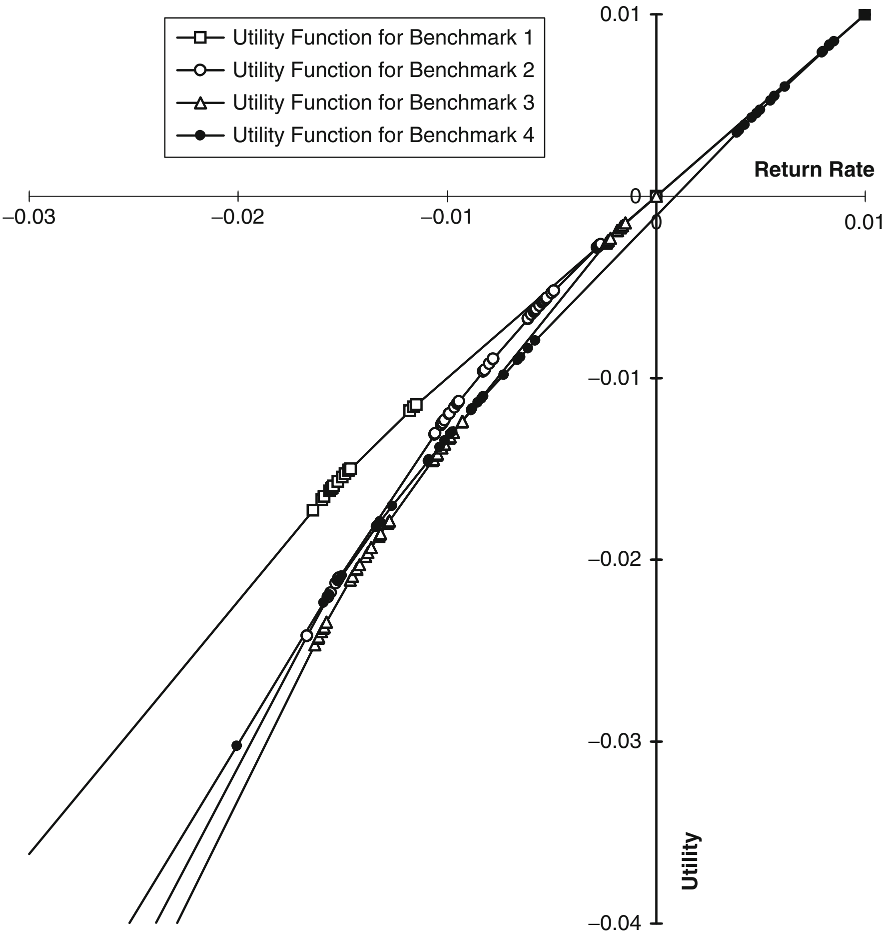


Fig. 83.5 Implied utility functions corresponding to dominance constraints for four benchmark portfolios

We have used four benchmark return rates Y. Each of them was constructed as a return rate of a certain index composed of our assets. Since we actually know the past return rates, for the purpose of comparison, we have selected equally weighted indexes composed of the N assets having the highest average return rates in this period. Benchmark 1 corresponds to $N = 26$, Benchmark 2 corresponds to $N = 54$, Benchmark 3 corresponds to $N = 82$, and Benchmark 4 corresponds to $N = 200$. Our problem was to maximize the expected return rate, under the condition that the return rate of the benchmark portfolio is dominated. Since the benchmark point was a return rate of a portfolio composed from the same basket, we have $m = T = 616$ in this case.

We have solved the problem by our method of minimizing the dual problem which was presented in (Dentcheva and Ruszczynski 2006a).

The implied utility functions from Eq. 83.37 obtained by solving the optimization problem (83.34) in the optimality conditions are illustrated in Fig. 83.5. We see that for Benchmark Portfolio 1, which contains only a small number of fast growing assets, the utility function is linear on almost the entire range of return rates. Only very negative return rates are penalized.

If the benchmark portfolio contains more assets and is therefore more diversified and less risky, in order to dominate it, we have to use a utility function which introduces penalty for a broader range of return rates and is steeper. For the broadly based index in Benchmark Portfolio 4, the optimal utility function is smoother and is nonlinear even for positive return rates. It is worth mentioning that all these utility functions, although nondecreasing and concave, have rather complicated shapes. It would be a very hard task to guess the utility function that should be used to obtain a solution which dominates our benchmark portfolio.

Obviously, the shape of the utility function is determined by the benchmark within the context of the optimization problem considered. If we change the optimization problem, the utility function will change.

Finally, we may remark that our model (83.22–83.24) can be used for testing the statistical hypothesis that the return rate Y of the benchmark portfolio is nondominated.

References

Arnold, B. C. (1980). *Majorization and the Lorenz order: A brief introduction* (Lecture Notes in Statistics, Vol. 43). Berlin: Springer.

Dentcheva, D. (2005). Optimization models with probabilistic constraints. In G. Calafiore & F. Dabbene (Eds.), *Probabilistic and randomized methods for design under uncertainty*. London: Springer.

Dentcheva, D., & Ruszczynski, A. (2003a). Optimization under linear stochastic dominance. *Comptes Rendus de l'Académie Bulgare des Sciences, 56*(6), 6–11.

Dentcheva, D., & Ruszczynski, A. (2003b). Optimization under nonlinear stochastic dominance. *Comptes Rendus de l'Académie Bulgare des Sciences, 56*(7), 19–25.

Dentcheva, D., & Ruszczynski, A. (2003c). Optimization with stochastic dominance constraints. *SIAM Journal on Optimization, 14*, 548–566.

Dentcheva, D., & Ruszczynski, A. (2004a). Optimality and duality theory for stochastic optimization problems with nonlinear dominance constraints. *Mathematical Programming, 99*, 329–350.

Dentcheva, D., & Ruszczynski, A. (2004b). Convexification of stochastic ordering constraints. *Comptes Rendus de l'Académie Bulgare des Sciences, 57*(3), 5–10.

Dentcheva, D., & Ruszczynski, A. (2004c). Semi-infinite probabilistic optimization: First order stochastic dominance constraints. *Optimization, 53*, 583–601.

Dentcheva, D., & Ruszczynski, A. (2006a). Portfolio optimization with stochastic dominance constraints. *Journal of Banking and Finance, 30*(2), 433–451.

Dentcheva, D., & Ruszczynski, A. (2006b). Inverse stochastic dominance constraints and rank dependent expected utility theory. *Mathematical Programming, 108*, 297–311.

Dentcheva, D., & Ruszczynski, A. (2008). Duality between coherent risk measures and stochastic dominance constraints in risk-averse optimization. *Pacific Journal of Optimization, 4*, 433–446.

Dentcheva, D., & Ruszczynski, A. (2010). Inverse cutting plane methods for optimization problems with second order stochastic dominance constraints. *Optimization, 59*, 323–338.

Elton, E. J., Gruber, M. J., Brown, S. J., & Goetzmann, W. N. (2006). *Modern portfolio theory and investment analysis*. New York: Wiley.

Fishburn, P. C. (1964). *Decision and value theory*. New York: Wiley.

Fishburn, P. C. (1970). *Utility theory for decision making*. New York: Wiley.

Gastwirth, J. L. (1971). A general definition of the Lorenz curve. *Econometrica, 39*, 1037–1039.

Hadar, J., & Russell, W. (1969). Rules for ordering uncertain prospects. *The American Economic Review, 59*, 25–34.

Hanoch, G., & Levy, H. (1969). The efficiency analysis of choices involving risk. *Review of Economic Studies, 36*, 335–346.

Hardy, G. H., Littlewood, J. E., & Polya, G. (1934). *Inequalities*. Cambridge, MA: Cambridge University Press.

Konno, H., & Yamazaki, H. (1991). Mean-absolute deviation portfolio optimization model and its application to Tokyo stock market. *Management Science, 37*, 519–531.

Lehmann, E. (1955). Ordered families of distributions. *Annals of Mathematical Statistics, 26*, 399–419.

Levy, H. (2006). *Stochastic dominance: Investment decision making under uncertainty* (2nd ed.). New York: Springer.

Lorenz, M. O. (1905). Methods of measuring concentration of wealth. *Journal of the American Statistical Association, 9*, 209–219.

Markowitz, H. M. (1952). Portfolio selection. *Journal of Finance, 7*, 77–91.

Markowitz, H. M. (1959). *Portfolio selection*. New York: Wiley.

Markowitz, H. M. (1987). *Mean-variance analysis in portfolio choice and capital markets*. Oxford: Blackwell.

Marshall, A. W., & Olkin, I. (1979). *Inequalities: Theory of majorization and its applications*. San Diego: Academic.

Mosler, K., & Scarsini, M. (Eds.). (1991). *Stochastic orders and decision under risk*. Hayward: Institute of Mathematical Statistics.

Muliere, P., & Scarsini, M. (1989). A note on stochastic dominance and inequality measures. *Journal of Economic Theory, 49*, 314–323.

Müller, A., & Stoyan, D. (2002). *Comparison methods for stochastic models and risks*. Chichester: Wiley.

Ogryczak, W., & Ruszczynski, A. (1999). From stochastic dominance to mean-risk models: Semideviations as risk measures. *European Journal of Operational Research, 116*, 33–50.

Ogryczak, W., & Ruszczynski, A. (2001). On consistency of stochastic dominance and mean-semideviation models. *Mathematical Programming, 89*, 217–232.

Ogryczak, W., & Ruszczynski, A. (2002). Dual stochastic dominance and related mean-risk models. *SIAM Journal on Optimization, 13*, 60–78.

Prekopa, A. (2003). Probabilistic programming. In A. Ruszczynski & A. Shapiro (Eds.), *Stochastic programming* (pp. 267–351). Amsterdam: Elsevier.

Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization, 3*, 225–243.

Quiggin, J. (1993). *Generalized expected utility theory – The rank-dependent expected utility model*. Dordrecht: Kluwer.

Quirk, J. P., & Saposnik, R. (1962). Admissibility and measurable utility functions. *Review of Economic Studies, 29*, 140–146.

Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk, 2*, 21–41.

Rothschild, M., & Stiglitz, J. E. (1969). Increasing risk: I. A definition. *Journal of Economic Theory, 2*, 225–243.

Ruszczynski, A., & Vanderbei, R. J. (2003). Frontiers of stochastically nondominated portfolios. *Econometrica, 71*, 1287–1297.

von Neumann, J., & Morgenstern, O. (1944). *Theory of games and economic behavior*. Princeton: Princeton University Press.

Wang, S. S., & Yong, V. R. (1998). Ordering risks: Expected utility versus Yaari's dual theory of risk. *Insurance Mathematics and Economics, 22*, 145–161.

Wang, S. S., Yong, V. R., & Panjer, H. H. (1997). Axiomatic characterization of insurance prices. *Insurance Mathematics and Economics, 21*, 173–183.

Whitmore, G. A., & Findlay, M. C. (Eds.). (1978). *Stochastic dominance: An approach to decision-making under risk*. Lexington: D.C. Heath.

Yaari, M. E. (1987). The dual theory of choice under risk. *Econometrica, 55*, 95–115.

Implementation Problems and Solutions in Stochastic Volatility Models of the Heston Type

84

Jia-Hau Guo and Mao-Wei Hung

Contents

Abstract

In Heston stochastic volatility framework, the main problem for implementing Heston semi-analytic formulae for European-style financial claims is the inverse Fourier integration. Without good implementation procedures, the numerical results obtained from Heston formulae cannot be robust, even for customarily used Heston parameters, as the time to maturity is increased. We compare three major approaches to solve the numerical instability problem inherent in the fundamental solution of the Heston model and show that the simple adjusted-formula method is much simpler than the rotation-corrected angle method of

J.-H. Guo (✉)
Institution of Finance, College of Management, National Chiao Tung University, Hsinchu, Taiwan
e-mail: jiahau@faculty.nctu.edu.tw

M.-W. Hung
College of Management, National Taiwan University, Taipei, Taiwan
e-mail: mwhung@ntu.edu.tw

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_84,

Kahl and Jäckel and also greatly superior to the direct integration method of Shaw if taking computing time into consideration.

In this chapter, we used the fundamental transform method proposed by Lewis to reduce the number of variables from two to one and separate the payoff function from the calculation of the Green function for option pricing. Furthermore, the simple adjusted formula is shown to be a robust option pricer as no complex discontinuities arise in this formulation even without the rotation-corrected angle.

Keywords

Heston • Stochastic volatility • Fourier inversion • Fundamental transform • Complex logarithm • Rotation-corrected angle • Simple adjusted formula • Green function

84.1 Introduction

Since the initial breakthrough by Heston (1993), the literature on asset pricing using stochastic volatility models has expanded dramatically over the last decade to capture the volatility smiles and skews appeared in market quotes. Within this class, the Heston stochastic volatility model famous for its closed-form characteristic function has spread widely in financial products. For example, Fang and Oosterlee (2011) used Heston model to price Bermudan and barrier options. However, the implementations of Heston formulae are not as straightforward as they may appear, and most numerical procedures are not reported in detail (see Lee 2005).

Recently, the robustness of Heston formula has become one of the main issues on option pricing (see the discussion of Albrecher et al. (2007)). The complex logarithm contained in the formula of the Heston model is the primary problem. As the claim of Lord and Kahl (2010), the characteristic function of Heston stochastic volatility model and all of its extensions can be discontinuous, leading to completely wrong option prices if the complex logarithm is restricted to its principal branch. This chapter compares three main approaches to this problem: rotation-corrected angle, direct integration, and simple adjusted formula.

It is a well-known fact that the logarithm of a complex variable $z = re^{i\theta}$ is multivalued, i.e., $\ln z = \ln |z| + i(\arg(z) + 2\pi n)$ where $\arg(z) \in [-\pi, \pi)$ and $n \in \mathrm{Z}$. If one restricts the logarithm to its principal branch by setting $n = 0$ (similar to most software packages, such as C++, Gauss, and Mathematica), it is necessarily discontinuous at the cut (see Fig. 84.1).

The Heston model is represented in Lewis' illustration, in which the type of financial claim is entirely decoupled from the calculation of the Green function. Different payoffs are then managed through elementary contour integration over functions and contours that depend on the payoff. In this way, one can see that the issue is fundamentally related to the Green function component of the solution. Once the implementation problems in the Green function component of the solution have been solved, the robustness of the formulae for all European-style financial claims in Heston model can be assured.

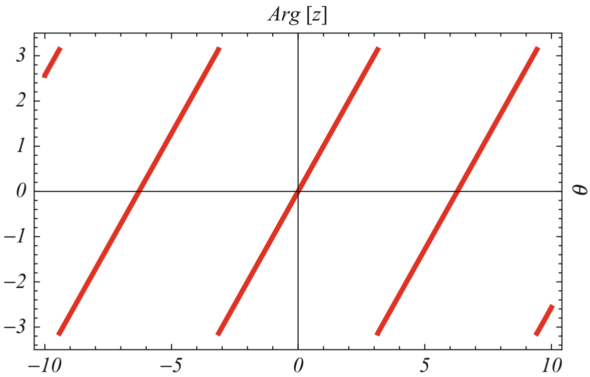


Fig. 84.1 Discontinuities occur at the cut by restricting the logarithm of a complex variable $z = e^{i\theta}$ to the principal branch

The rest of this chapter proceeds as follows: Sect. 84.2 gives the derivation of the transformed-based solution for Heston stochastic volatility model and introduces the discontinuity problem arising from the derived formula. Section 84.3 compares three main solutions to the discontinuity problem and gives some numerical examples to illustrate their usefulness. Section 84.4 summarizes the chapter.

84.2 The Transform-Based Solution for Heston Stochastic Volatility Model

Heston stochastic volatility model is based on the system of stochastic differential equations, which represent the dynamics of the stock price and variance processes under the risk-neutral measure

$$dS_t = rS_t dt + S_t\sqrt{V_t}dW_S(t) \tag{84.1}$$

$$dV_t = \kappa(\theta - V_t)dt + \sigma_V\sqrt{V_t}dW_V(t). \tag{84.2}$$

S_t and V_t denote the stock price and its variance at time t, respectively; r is the risk-free interest rate. The variance evolves according to a square-root process: θ is the long-run mean variance, κ is the speed of mean reversion, and σ_V is the parameter which controls the volatility of the variance process. W_S and W_V are two standard processes of Brownian motion having the correlation ρ. The Heston partial differential equation for a European-style claim $C(S,V,t)$ with expiration T is

$$\frac{\partial C}{\partial t} + \frac{VS^2}{2}\frac{\partial^2 C}{\partial S^2} + \sigma_V\rho SV\frac{\partial^2 C}{\partial S\partial V} + \frac{\sigma_V^2 V}{2}\frac{\partial^2 C}{\partial V^2} + rS\frac{\partial C}{\partial S} + \kappa(\theta - V)\frac{\partial C}{\partial V} - rC = 0. \tag{84.3}$$

The fundamental transform method proposed by Lewis (2000) can reduce Eq. 84.3 from two variables to one and entirely separate every (volatility independent) payoff function from the calculation of the Green function. After substituting the following, $\tau = T - t$, $x = \log(S) + r(T - t)$, $C(S,V,t) = W(x,V,\tau)e^{-r(T-t)}$ into Eq. 84.3, we have

$$\frac{\partial W}{\partial \tau} = \frac{1}{2}V\left(\frac{\partial^2 W}{\partial x^2} - \frac{\partial W}{\partial x}\right) + \rho\sigma_V V\frac{\partial^2 W}{\partial x \partial V} + \frac{1}{2}\sigma_V^2 V\frac{\partial^2 W}{\partial V^2} + \kappa(\theta - V)\frac{\partial W}{\partial V}. \quad (84.4)$$

Let $G(\phi,V,\tau)$ denote the Fourier transform of $W(x,V,\tau)$:

$$G(\phi, V, \tau) \equiv \int_{-\infty}^{\infty} e^{i\phi x} W(x, V, \tau)dx. \quad (84.5)$$

Given the transform $G(\phi,V,\tau)$, the inversion formula is

$$W(x, V, \tau) = \frac{1}{2\pi}\int_{i\mathrm{Im}[\phi]-\infty}^{i\mathrm{Im}[\phi]+\infty} e^{-i\phi x} G(\phi, V, \tau)d\phi. \quad (84.6)$$

so that differentiation w.r.t. x becomes multiplication by $-i\phi$ in the transform. By taking the τ derivative of both sides of Eq. 84.5 and then replacing $\partial W/\partial\tau$ inside the integral by the right-hand side of Eq. 84.4, we translate Eq. 84.4 into a PDE for $G(\phi,V,\tau)$:

$$\frac{\partial G}{\partial \tau} = \frac{1}{2}\sigma_V^2 V\frac{\partial^2 G}{\partial V^2} - \frac{1}{2}V\left(\phi^2 - i\phi\right)G + \left(\kappa(\theta - V) - i\phi\sigma_V\rho V\right)\frac{\partial G}{\partial V}. \quad (84.7)$$

Hence, a solution $G(\phi,V,\tau)$ to Eq. 84.7, which satisfies $G(\phi,V,0) = 1$, is called a fundamental transform. Given the fundamental transform, a solution for a particular payoff can be obtained by

$$C(S, V, t) = \frac{1}{2\pi}e^{-r(T-t)}\int_{i\mathrm{Im}[\phi]-\infty}^{i\mathrm{Im}[\phi]+\infty} e^{-i\phi x}\,\widetilde{W}(\phi, V, 0)G(\phi, V, \tau)d\phi, \quad (84.8)$$

where $\widetilde{W}(\phi, V, 0)$ is the Fourier transform of the payoff function at maturity.

We will deal with a few common types of payoff functions and see what restrictions are necessary for their Fourier transforms to exist.

84.2.1 Call Option

At maturity, the payoff of a vanilla call option with strike K is $\mathrm{Max}[S_T - K, 0]$ in terms of our original variables. In terms of the logarithmic variables, we have $W(x,V,0) = \mathrm{Max}[e^x - K, 0]$, so the Fourier transform of the payoff is of the form

$$\widetilde{W}(\phi, V, 0) = \int_{-\infty}^{\infty} e^{i\phi x} W(x, V, 0) dx = \int_{\log[K]}^{\infty} e^{i\phi x}(e^x - K) dx$$
$$= K^{1+i\phi} / \left(i\phi - \phi^2\right), \tag{84.9}$$

which does not exist unless $\mathrm{Im}[\phi] > 1$.

84.2.2 Put Option

The payoff of a vanilla put option with strike K is $\mathrm{Max}[K - S_T, 0]$. Its transformed payoff is also $K^{1+i\phi}/(i\phi - \phi^2)$, but the restriction is $\mathrm{Im}[\phi] < 0$.

84.2.3 Digital Call

The payoff of a digital call with strike K is $H[S_T - K]$ where H is a Heaviside function. Its transformed payoff is $-K^{i\phi}/(i\phi)$, subject to $\mathrm{Im}[\phi] > 0$.

84.2.4 Cash-Secured Put

The payoff of a cash-secured put with strike K is $\mathrm{Min}[S_T, K]$. Its transformed payoff is $K^{1+i\phi}/(\phi^2 - i\phi)$, subject to $0 < \mathrm{Im}[\phi] < 1$.

The fundamental solution of Eq. 84.7 is in the form

$$G(\phi, V, \tau) = e^{A(\tau,\phi)+B(\tau,\phi)V}. \tag{84.10}$$

After substituting Eq. 84.10 into Eq. 84.7, a pair of ordinary differential equations for $A(\tau,\phi)$ and $B(\tau,\phi)$ is obtained

$$\dot{A} = \theta\kappa B. \tag{84.11}$$

$$\dot{B} = \frac{1}{2}B^2\sigma_V^2 - B(\kappa + i\phi\sigma_V\rho) - \frac{1}{2}\left(\phi^2 - i\phi\right). \tag{84.12}$$

The solutions can be expressed by

$$B(\tau, \phi) = \frac{(\kappa + i\rho\sigma_V\phi + d(\phi))}{\sigma_V^2} \frac{\left(1 - e^{d(\phi)\tau}\right)}{\left(1 - g(\phi)e^{d(\phi)\tau}\right)} \tag{84.13}$$

$$A(\tau, \phi) = \frac{\kappa\theta}{\sigma_V^2}\left((\kappa + i\rho\sigma_V\phi + d(\phi))\tau - 2\log\left[\frac{g(\phi)e^{d(\phi)\tau} - 1}{g(\phi) - 1}\right]\right) \tag{84.14}$$

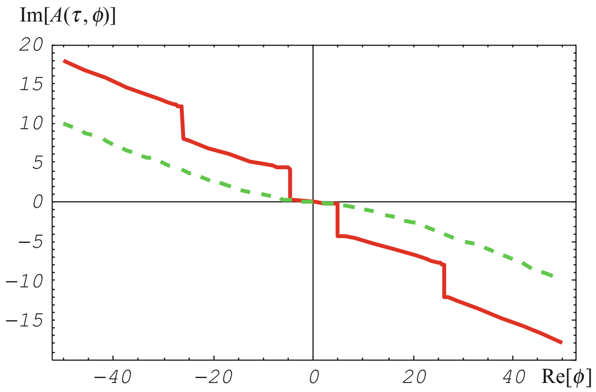


Fig. 84.2 The discontinuity occurs in the fundamental solution of the Heston model if the logarithm with complex arguments is restricted to the principal branch. Underlying: $dS_t = r S_t dt + \sqrt{V_t} S_t dW_S(t)$ with $S_0 = 100$ and $r = 0.0319$. Variance: $dV_t = \kappa(\theta - V_t)dt + \sigma_V \sqrt{V_t} dW_V(t)$ with $V_0 = 0.010201$, $\kappa = 6.21$, $\theta = 0.019$, $\sigma_V = 0.61$, and $\rho = -0.70$. Time to maturity: $T = 2.00$. The *red line* was obtained by evaluating $A(\tau,\phi)$ with the unfixed form given in Eq. 84.14. The *green dashed line* was obtained by evaluating $A(\tau,\phi)$ with the adjusted formula given in Eq. 84.31 and is the correct curve. The logarithmic function for both cases is restricted to using only the principal branch

using the auxiliary functions

$$g(\phi) = \frac{\kappa + i\rho\sigma_V\phi + d(\phi)}{\kappa + i\rho\sigma_V\phi - d(\phi)}, d(\phi) = \sqrt{\left(\phi^2 - i\phi\right)\sigma_V^2 + \left(\kappa + i\rho\sigma_V\phi\right)^2}. \quad (84.15)$$

If the complex, multivalued logarithm is restricted to the principal branch only, discontinuities are necessarily incurred at the cut of the complex logarithm along the integration path, resulting in an incorrect value for Heston formula. Figure 84.2 illustrates the discontinuity problem in the implementation of the fundamental solution. In this example, depicted in Fig. 84.2, $S_0 = 100$, $r = 0.0319$, $V_0 = 0.010201$, $\rho = -0.70$, $\kappa = 6.21$, $\theta = 0.019$, $\sigma_V = 0.61$, and $\text{Im}[\phi] = 2$.

Reasonable parameters in practice may incur the numerically induced discontinuity such that the correct treatment of the phase jump is very crucial. In fact, in examples with long maturity periods, discontinuities are certain to arise from the formula presented in Eq. 84.14 for $A(\tau,\phi)$, if the complex logarithm uses the principal branch only and $2\kappa\theta/\sigma_V^2$ is not an integer (see Fig. 84.3).

One may shift the problem from the complex logarithm to the evaluation of

$$G(\phi, V, \tau) = \left(\frac{g(\phi)e^{d(\phi)\tau} - 1}{g(\phi) - 1}\right)^{-2\alpha} e^{\alpha(\kappa + i\rho\sigma_V\phi + d(\phi))\tau + B(\tau,\phi)V}, \quad (84.16)$$

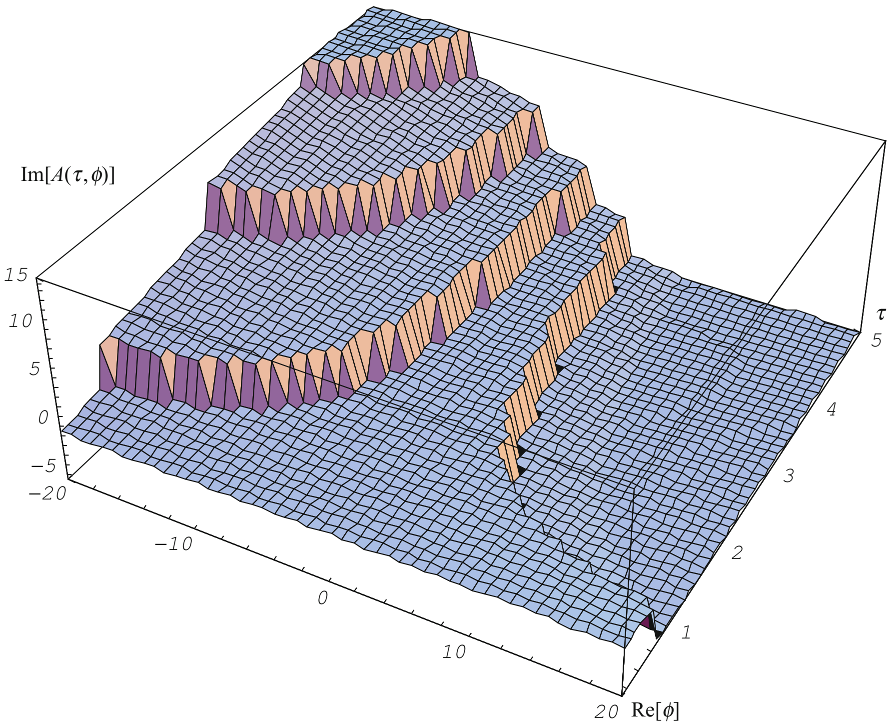


Fig. 84.3 Discontinuities arise quite naturally for customarily used Heston parameters, typically occurring in practice as time to maturity is increased. The other parameters are the same as those specified in Fig. 84.2

where $\alpha = \kappa\theta/\sigma_V^2$. However, this formula comes with the related branch-switching problem of the complex power function, and discontinuities do not diminish in its implementation. Note that taking a complex variable z to the power α gives

$$z^{\alpha} = r^{\alpha} e^{i\alpha\theta}. \tag{84.17}$$

After restricting $\arg(z) \in [-\pi, \pi)$, the complex plane is cut along the negative real axis. Whenever z crosses the negative real axis, the sign of its phase changes from π to $-\pi$. Therefore, the phase of z^{α} changes from $\alpha\pi$ to $-\alpha\pi$. This may lead to a jump because

$$e^{i\pi} = e^{-i\pi} \Rightarrow \begin{cases} e^{i\alpha\pi} \neq e^{-i\alpha\pi} & \text{if } \alpha \notin \mathrm{Z} \\ e^{i\alpha\pi} = e^{-i\alpha\pi} & \text{if } \alpha \in \mathrm{Z} \end{cases}. \tag{84.18}$$

To demonstrate this, Fig. 84.4 gives the scenario that $\alpha \in \mathrm{Z}$ and there is no jump at all. Figure 84.5 gives another scenario that $\alpha \notin \mathrm{Z}$ and the complex power function indeed incurs jumps.

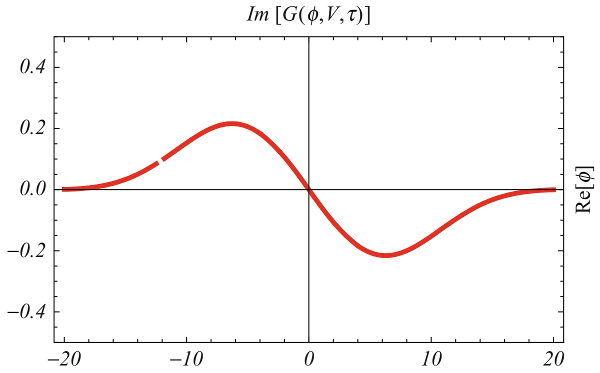


Fig. 84.4 If the power of a complex variable z^{α} is restricted to the principal branch, $\alpha \in \mathrm{Z}$ makes discontinuities diminish at the cut. The fundamental function $G(\phi,V,\tau)$ in Eq. 84.16 is evaluated with the same parameters specified in Fig. 84.2 except for $\kappa = 19.58421$. In this scenario, $\alpha = \kappa\theta/\sigma_V^2 = 1$

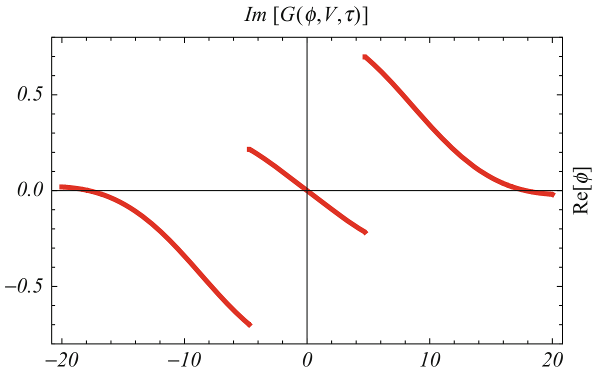


Fig. 84.5 If the power of a complex variable z^{α} is restricted to the principal branch, $\alpha \notin \mathrm{Z}$ makes discontinuities occur at the cut. The fundamental function $G(\phi,V,\tau)$ in Eq. 84.16 is evaluated with the same parameters specified in Fig. 84.2. In this scenario, $\alpha = \kappa\theta/\sigma_V^2 = 0.3170921$

84.3 Solutions to the Discontinuity Problem of Heston Formula

In the literature, various authors propose the idea of carefully keeping track of the branch by monitoring the complex logarithm function for each step along a discretized integral path to remedy phase jumps. As described in Kruse and Nögel (2005), if the imaginary value of the complex logarithm for one step differs from the previous one by more than 2π, the jump of 2π is added or subtracted to recover the continuity of phase. However, using this approach, the already complex integrals of Heston formula may become too complicated in practice. Therefore, simulation is also considered as a practical alternative for finding option prices (see Broadie and Kaya 2006).

To make matters worse, discontinuities arise quite naturally for customarily used Heston parameters simply as time to maturity is increased, thereby illustrating the importance of the correct treatment of phase jumps for Heston formula. Kahl and

Jäckel (2005) remedied these discontinuities using the rotation-corrected angle of the phase of a complex variable. Shaw (2006) dealt with this problem by replacing the call to the complex logarithm by direct integration of the differential equation. In addition, Guo and Hung (2007) also proposed a simple adjusted formula to solve this discontinuity problem. From a computational and convenience point of view, the solution of Guo and Hung (2007) can be implemented easily and is thereby suitable for practical application. These solutions are presented by the following statements.

84.3.1 Rotation-Corrected Angle

In order to guarantee the continuity of $A(\tau,\phi)$, a rotation-corrected term must be additionally calculated in advance. First, we introduce the notation

$$g(\phi) = r_g(\phi)e^{i\theta_g(\phi)}, \tag{84.19}$$

$$d(\phi) = a_d(\phi) + ib_d(\phi). \tag{84.20}$$

The next step is to have a closer look at the denominator of $(g(\phi)e^{d(\phi)\tau} - 1)/(g(\phi) - 1)$:

$$g(\phi) - 1 = r_g(\phi)e^{i\theta_g(\phi)} - 1 = r_g^*(\phi)e^{i\left(\chi_g^*(\phi)+2\pi m\right)} \tag{84.21}$$

where

$$m = \mathrm{int}\left[\frac{\theta_g(\phi) + \pi}{2\pi}\right], \tag{84.22}$$

$$\chi_g^*(\phi) = \arg(g(\phi) - 1), \tag{84.23}$$

$$r_g^*(\phi) = |g(\phi) - 1|, \tag{84.24}$$

and with int[•] denoting Gauss's integer brackets. The same calculation is done with the numerator:

$$\begin{aligned} g(\phi)e^{d(\phi)\tau} - 1 &= r_g(\phi)e^{i\theta_g(\phi)}e^{(a_d(\phi)+ib_d(\phi))\tau} - 1 = r_g(\phi)e^{a_d(\phi)\tau}e^{i\left(\theta_g(\phi)+b_d(\phi)\tau\right)} - 1 \\ &= r_{gd}^*(\phi)e^{i\left(\chi_{gd}^*(\phi)+2\pi n\right)} \end{aligned} \tag{84.25}$$

where

$$n = \mathrm{int}\left[\frac{\theta_g(\phi) + b_d(\phi)\tau + \pi}{2\pi}\right], \tag{84.26}$$

$$\chi^*_{gd}(\phi) = \arg\left(g(\phi)e^{d(\phi)\tau} - 1\right), \tag{84.27}$$

$$r^*_{gd}(\phi) = \left|g(\phi)e^{d(\phi)\tau} - 1\right|. \tag{84.28}$$

Hence, we can compute the logarithm of $(g(\phi)e^{d(\phi)\tau} - 1)/(g(\phi) - 1)$ quite simply as

$$\log\left[\frac{g(\phi)e^{d(\phi)\tau} - 1}{g(\phi) - 1}\right] = \log\left[r^*_{gd}(\phi)/r^*_g(\phi)\right] + i\left[\chi^*_{gd}(\phi) - \chi^*_g(\phi) + 2\pi(n - m)\right], \tag{84.29}$$

where $2\pi(n - m)$ is the rotation-corrected angle.

84.3.2 Direct Integration

Another way to avoid the branch cut difficulties arising from the choice of the branch of the complex logarithm is to perform direct numerical integration of $A(\tau,\phi)$ w.r.t. τ according to Eq. 84.11. Given $B(\tau,\phi)$, $A(\tau,\phi)$ can be obtained by

$$A(\tau, \phi) = \theta\kappa\int_0^\tau B(s, \phi)ds. \tag{84.30}$$

After replacing the call to the complex logarithm by direct integration of the differential equation, the complex logarithm cannot be a problem anymore, and the continuity of $A(\tau,\phi)$ is guaranteed.

84.3.3 Simple Adjusted Formula

Here, we briefly introduce the simple adjusted formula of Guo and Hung (2007) to the discontinuity problem in the implementation of Heston formula. The solution is to move $\exp[-d(\phi)\tau]$ into the logarithm of $A(\tau,\phi)$ by simply adjusting $A(\tau,\phi)$ as follows:

$$A(\tau, \phi) = \frac{\kappa\theta}{\sigma_V^2}\left((\kappa + i\rho\sigma_V\phi - d(\phi))\tau - 2\log\left[\frac{g(\phi)e^{d(\phi)\tau} - 1}{g(\phi) - 1}e^{-d(\phi)\tau}\right]\right). \tag{84.31}$$

The insight in formula (84.31) is that the subtraction of the number 1 from a complex variable, c, results simply in a shift parallel to the real axis. Because an imaginary component must be added to move a complex number across the negative real axis, the phases of $c - 1$ and c exist on the same phase interval.

Therefore, the logarithms of $c-1$ and c have the same rotation count number. It illuminates this simple solution to assure that the phase of $A(\tau,\phi)$ is continuous, without the necessity of a rotation-corrected term.

The logarithm presented in Eq. 84.31 is the only term possibly giving rise to discontinuity. Trivially,

$$\log\left[\frac{g(\phi)e^{d(\phi)\tau}-1}{g(\phi)-1}e^{-d(\phi)\tau}\right]=\log\left[\frac{g(\phi)}{g(\phi)-1}\right]+\log\left[\frac{g(\phi)e^{d(\phi)\tau}-1}{g(\phi)e^{d(\phi)\tau}}\right]. \tag{84.32}$$

Because the subtraction of 1 does not affect the rotation count of the phase of a complex variable, $g(\phi)$ has the same rotation count number (m) as $g(\phi)-1$, and $g(\phi)e^{d(\phi)\tau}-1$ also has the same rotation count number (n) as $g(\phi)e^{d(\phi)\tau}$. Hence, $\log[(g(\phi)e^{d(\phi)\tau}-1)/((g(\phi)-1)e^{d(\phi)\tau})]$ needs no rotation-corrected terms for all levels of Heston parameters because

$$\begin{aligned}\log\left[\frac{g(\phi)e^{d(\phi)\tau}-1}{(g(\phi)-1)e^{d(\phi)\tau}}\right] &= (\log[g(\phi)]-\log[g(\phi)-1])\\ &\quad+\left(\log\left[g(\phi)e^{d(\phi)\tau}-1\right]-\log\left[g(\phi)e^{d(\phi)\tau}\right]\right).\end{aligned} \tag{84.33}$$

Hence, the formula in Eq. 84.31, for $A(\tau,\phi)$, provides a simple solution to the discontinuity problem for Heston stochastic volatility model.

Compared to the rotation-corrected angle method, the simple adjusted-formula method needs no rotation-corrected terms in the already complex integral of Heston formula to recover its continuity for all levels of Heston parameters. Although the direct integration method neither needs the rotation-corrected terms to guarantee the continuity of Heston formula, it inevitably introduces the discretization bias into the evaluation of the Green function component of the solution. This bias may create another serious problem of computation. Many steps may be necessary to reduce the bias to an acceptable level, and, hence, more computational effort is needed to guarantee that the bias is small enough. As a consequence, the direct integration method requires more computing time than the simple adjusted-formula method to avoid the discontinuity problem arising from the complex logarithm.

Figure 84.6 illustrates a comparison of the computing time for applying the simple adjusted-formula and direct integration methods to evaluate a European call option. The direct integration method is more time-consuming than the simple adjusted-formula method. Moreover, the simple adjusted-formula method has an advantage in that its time consumption remains almost at the same level as the time to maturity increases. In contrast, the computing time via the direct integration method increases rapidly with an increase in time to maturity. These computational results were performed on a desktop PC with an Intel Pentium D 3.4 GHz processor and 1 GB of RAM, running Windows XP Professional. The codes were written using the Mathematica software.

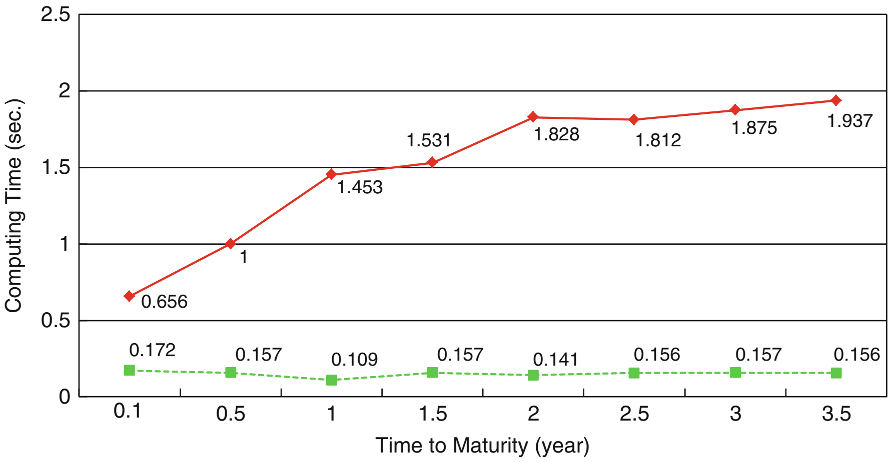


Fig. 84.6 Computing time comparison under Heston stochastic volatility model for a European call option: direct integration versus simple adjusted formula. The *red line* represents the computing time for evaluating a standard call option price using $A(\tau,\phi)$ via direct integration w.r.t. τ given in Eq. 84.30. The *green dashed line* represents the computing time for evaluating the same call option price using $A(\tau,\phi)$ via the simple adjusted formula given in Eq. 84.31. The other parameters are the same as those specified in Fig. 84.2

Table 84.1 Impact of the discontinuity problem on the evaluation of European call options in Heston model on stochastic volatility

		Fundamental solution of the Heston model	
T (year)	Monte Carlo simulation with the exact method (10,000 trials)	Adjusted formula using formula (84.31)	Unfixed formula using formula (84.14)
0.50	4.2658	4.2545	4.2555
1.00	6.7261	6.8061	6.4483
1.50	8.9510	8.9557	8.3286
2.00	10.9633	10.8830	9.7079
2.50	12.6100	12.6635	10.5542
3.00	14.2591	14.3366	10.9778

Here, $S_0 = 100.00$, $K = 100.00$, $r = 0.0319$, $V_0 = 0.010201$, $\rho = -0.70$, $\kappa = 6.21$, $\theta = 0.019$, and $\sigma_v = 0.61$. Note that the evaluation of the fundamental solution of the Heston model using formula (84.31) still yields values consistent with those of the Monte Carlo simulation for all time-to-maturity cases, although the complex logarithm is restricted to the principal branch

Table 84.1 is an illustration of the usefulness of the simple adjusted formula for evaluating European call options in Heston model using the complex logarithm restricted to the principal branch. The algorithm was verified using Monte Carlo simulation with the exact method proposed by Broadie and Kaya (2006) for the stochastic volatility process. Although the exact method of simulation has the advantage that its convergence rate is much faster than that of the conventional Euler discretization method, it is, of course, computationally more burdensome than the simple adjusted-formula method.

84.4 Summary

This chapter looks at the issue raised by branch cuts in the transform solutions for European-style financial claims in the Heston model. The multivalued nature of the complex logarithm and power functions results in numerical instability in the implementation of the fundamental transform. Compared to the work of Kahl and Jäckel (2005), neither the direct integration method of Shaw (2006) nor the simple adjusted formula of Guo and Hung (2007) requires rotation-corrected terms to assure the robustness of the evaluation of Heston formulae. After taking computing time into consideration, the evidence shows that the simple adjusted-formula method is greatly superior to the direct integration method.

References

Albrecher, H., Mayer, P., Schoutens, W., & Tistaert, J. (2007, January). *The little Heston trap. Wilmott Magazine*, 83–92.

Broadie, M., & Kaya, Ö. (2006). Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research, 54*(2), 217–231.

Fang, F., & Oosterlee, C. W. (2011). A Fourier-based valuation method for Bermudan and barrier options under Heston's model. *SIAM Journal on Financial Mathematics, 2*, 439–463.

Guo, J., & Hung, M. (2007). A note on the discontinuity problem in Heston's stochastic volatility model. *Applied Mathematical Finance, 14*(4), 339–345.

Heston, S. (1993). A closed form solution for options with stochastic volatility with application to bond and currency options. *Review of Financial Studies, 6*(2), 327–343.

Kahl, C., & Jäckel, P. (2005). Not-so-complex logarithm in the Heston model. *Wilmott Magazine, 19*, 94–103.

Kruse, S., & Nögel, U. (2005). On the pricing of forward starting options in Heston's model on stochastic volatility. *Finance and Stochastics, 9*, 223–250.

Lee, R. (2005). Option pricing by transform methods: Extensions, unification, and error control. *Journal of Computational Finance, 7*(3), 51–86.

Lewis, A. (2000). Option valuation under stochastic volatility. Finance Press.

Lord, R., & Kahl, C. (2010). Complex logarithms in Heston-like models. *Mathematical Finance, 20*(4), 671–694.

Shaw, W. (2006). Stochastic volatility: Models of Heston type, lecture notes. http://www.mth.kcl.ac.uk/~shaww/web_page/papers/StoVolLecture.pdf, http://www.mth.kcl.ac.uk/~shaww/web_page/papers/StoVolLecture.nb

Stochastic Change-Point Models of Asset Returns and Their Volatilities

85

Tze Leung Lai and Haipeng Xing

Contents

Abstract

We begin with an overview of frequentist and Bayesian approaches to incorporating change points in time series models of asset returns and their volatilities. It has been found in many empirical studies of stock returns and exchange rate that ignoring the possibilities of parameter changes yields time series models with long memory, such as unit-root nonstationarity and high volatility persistence. We therefore focus on the ARX-GARCH model and introduce two timescales, using the "short" timescale to define GARCH dynamics and the

T.L. Lai
Stanford University, Stanford, CA, USA
e-mail: lait@stanford.edu

H. Xing (✉)
SUNY at Stony Brook, Stony Brook, NY, USA
e-mail: xing@ams.sunysb.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_85,

"long" timescale to incorporate parameter jumps. This leads to a Bayesian change-point ARX-GARCH model, whose unknown parameters may undergo occasional changes at unspecified times and can be estimated by explicit recursive formulas when the hyperparameters of the Bayesian model are specified. We describe efficient estimators of the hyperparameters of the Bayesian model leading to empirical Bayes estimators of the piecewise constant parameters with relatively low computational complexity. We also show how the computationally tractable empirical Bayes approach can be applied to the frequentist problem of partitioning the time series into segments under sparsity assumptions on the change points.

Keywords

ARX-GARCH • Bounded complexity • Contemporaneous jumps • Change-point models • Empirical Bayes • Frequentist segmentation • Hidden Markov models • Hyperparameter estimation • Markov chain Monte Carlo • Recursive filters • Regression models • Stochastic volatility

85.1 Introduction

As we have noted in Lai and Xing (2010), portfolio theory, asset pricing, and hedging all involve volatilities, and volatility modeling is a cornerstone of empirical finance. Since the seminal works of Engle (1982) and Bollerslev (1986), generalized autoregressive conditionally heteroskedastic (GARCH) models have been widely used to model and forecast volatilities of financial time series. In many empirical studies of stock returns and exchange rates, estimation of the parameters v, a, and b in the GARCH (1,1) model

$$y_n = \sigma_n \epsilon_n, \quad \sigma_n^2 = (1 - a - b)v^2 + a y_{n-1}^2 + b\sigma_{n-1}^2$$

reveals high volatility persistence, with the maximum likelihood estimate of $a + b$ close to 1. To model such persistence, Engle and Bollerslev (1986) considered the "integrated" GARCH (IGARCH) models, and Baillie et al. (1996) introduced fractional integration in their FIGARCH models, with a slow hyperbolic rate of decay for the influence of the past innovations, to quantify the long memory of exchange rate volatilities. However, it has been pointed out that if the model parameters undergo occasional changes, then the fitted models that assume time-invariant parameters tend to exhibit long memory; see Diebold (1986), Perron (1989), Rappoport and Reichlin (1989), and Lamoureux and Lastrapes (1990). Therefore the long memory in the IGARCH and FIGARCH models may be due to the long timescale of the parameter changes that are ignored in these models.

The frequentist approach to incorporating change points in regression models assume that the change points are unknown parameters to be estimated from the data; see Quandt (1958, 1960), Andrews et al. (1996), Bai (1997, 1999), Bai et al. (1998), Bai and Perron (1998, 2003), Qu and Perron (2007); and the references therein. Finding these change points amounts to segmenting the data. Because of the computational complexity and analytic intractability, it is prohibitively difficult to extend these methods from relatively simple regression models to incorporate change points in a stochastic regression model with GARCH-type error variances. An alternative approach is Bayesian and assumes that the change points and the associated regimes are generated by some stochastic process so that the unknown regression parameters can be estimated from their posterior distribution via Bayes theorem; see Goldfeld and Quandt (1973) and Hamilton (1989, 1990). The Bayesian approach has been used to develop hidden Markov models (HMM) for asset returns, allowing contemporaneous jumps in their levels and volatilities. Markov chain Monte Carlo (MCMC) methods are used to estimate the hidden states and the unknown hyperparameters of the Bayesian method. Section 85.2 gives an overview of the literature, including the SVCJ models (stochastic volatility models with contemporaneous jumps in returns and volatilities) that have been an active area of research in finance and econometrics in the past decade.

In Sect. 85.3 we describe a much simpler class of Bayesian change-point models for asset returns and their volatility that we recently developed. This class of models yields explicit recursive filters and smoothers, thereby obviating the reliance on MCMC methods whose convergence properties and performance in change-point time series models have not been systematically studied because of their computational complexity. The optimal Bayes estimates in our change-point ARX-GARCH model involve unspecified hyperparameters, which can in principle be estimated by the EM algorithm. For regime-switching ARCH models, Cai (1994) have noted the "tremendous complication" of the normal equations of the EM algorithm, making it "extremely difficult" to implement for sample sizes exceeding 50. By using recursive representations of the summands of the log-likelihood function, we have a relatively simple algorithm to evaluate the log-likelihood function and estimate the hyperparameters of the change-point model. In Sect. 85.3, after introducing the change-point ARX-GARCH model, we first describe the associated filters for estimating the piecewise constant ARX parameter β_n and long-run volatility v_n based on the observations $y_1, \ldots, y_n$, assuming known hyperparameters that include the GARCH parameters associated with short-term volatility changes. Bounded complexity mixture (BCMIX) approximations are also developed for efficient computation of these filters and of the likelihood function to estimate the hyperparameters. We then use these closed-form recursions to develop BCMIX approximations to the Bayes estimates (smoothers) of β_t and v_t for $1 \leq t \leq n$. In contrast to our empirical Bayes approach that assumes a relatively simple stochastic model for change points, the frequentist approach which is often called "segmentation" assumes the change points and the pre- and post-change regression coefficients in regression models to be unknown parameters and uses maximum likelihood to estimate them and a model selection criterion to determine the number of change points. Section 85.3 also uses the relative simplicity of the EB smoothers to resolve difficulties in the frequentist segmentation problem. Section 85.4 gives some concluding remarks and discussion.

85.2 Overview of Bayesian Models with Jumps in Returns and Volatilities

85.2.1 Jumps in Regression Parameters, Error Variance, and ARCH Parameters

Since this chapter focuses on volatility modeling as in Lai and Xing (2010), we only review Bayesian change-point models that include changes in volatilities. Albert and Chib (1993) consider ARX models (autoregressive models with exogenous inputs) whose levels and error variances are subject to regime changes determined by a two-state Markov chain with unknown transition probabilities. Specifically, assuming the state space for the two-state Markov chain to be $\{0, 1\}$, with $s_t = 1$ representing that a change has occurred at time t, their model can be expressed as

$$y_t = \beta^T \mathbf{x}_t + \gamma s_t + \sum_{i=1}^{\kappa} \phi_i \left(y_{t-i} - \beta^T \mathbf{x}_{t-1} - \gamma s_{t-1} \right) + \left(\sigma^2 + \tau^2 s_t \right)^{1/2} \epsilon_t, \qquad (85.1)$$

in which ϵ_t are independent standard normal, σ^2, τ^2, γ, $\phi_1, \ldots, \phi_\kappa$ are unknown parameters such that $\phi(z) := 1 - \phi_1 z - \cdots - \phi_\kappa z^k$ has zeros outside the unit circle. Their abstract says, "The unobserved states, one for each time point, are treated as missing data and then analyzed via the simulation tool of Gibbs sampling. This method is expedient because the conditional posterior distribution of the parameters, given the states, and the conditional posterior distribution of the states, given the parameters, all have a form amenable to Monte Carlo sampling. The approach is straightforward and generates marginal posterior distribution for all parameters of interest. Posterior distributions of the states, future observations, and the residuals, averaged over the parameter space are also obtained." The unknown hyperparameters in their Bayesian model are estimated by maximum likelihood.

Instead of including ARX dynamics as in Eq. 85.1, Hamilton and Susmel (1994) consider regime-switching ARCH models

$$y_t = \alpha + \phi y_{t-1} + \tau(s_t) \sqrt{h_t} \epsilon_t, \qquad (85.2a)$$

$$h_t = a_0 + a_1 u_{t-1}^2 + \cdots + a_q u_{t-q}^2, \qquad (85.2b)$$

in which ϵ_t are independent standard normal, α, ϕ, $a_0, \ldots, a_q$, $\tau(1), \ldots, \tau(K)$ are unknown parameters, and s_t is an unobserved K-state Markov chain with unknown transition probabilities on state space $\{1, \ldots, K\}$. They use maximum likelihood to estimate the unknown parameters and also develop a complicated summation formula to compute the conditional density of $(s_t, s_{t-1}, \ldots, s_{t-q})$ given $y_1, \ldots, y_t$. They call (85.2) a *Markov-switching ARCH* model, denoted SWARCH(K, q), and have also extended it to include leverage effects by adding $\xi u_{t-1}^2 I_{\{u_{t-1} \leq 0\}}$ to the right-hand side of (85.2b), in which case the model is denoted SWARCH-L(K, q).

These Markov-switching ARCH (or ARCH-L) models can also use the standardized Student-t instead of standard normal errors as in the usual GARCH models; see Lai and Xing (2010, Sect. 93.2.3).

Instead of ARCH models, So et al. (1998) consider stochastic volatility models (Lai and Xing 2010, Sect. 93.2.7) and include regime-switching features in those models. Their Markov-switching stochastic volatility (MSSV) models are of the form

$$y_t = \sqrt{h_t}\epsilon_t, \quad \log h_t = \tau(s_t) + \phi \log h_{t-1} + \eta_t, \tag{85.3}$$

in which ϵ_t are independent standard normal, η_t are independent $N(0, \sigma^2)$ and ϕ, σ^2, $\tau(1), \ldots, \tau(K)$ are unknown parameters, and s_t is an unobserved K-state Markov chain with unknown transition probabilities on $\{1, \ldots, K\}$. As in Albert and Chib (1993), they use Gibbs sampling to estimate by Monte Carlo the posterior distribution of the states and also of the h_t after putting a prior distribution on the vector of unknown parameters and transition probabilities.

Assuming the number, but not the locations, of the change points to be known, Wang and Zivot (2000) consider a deterministically trending dynamic time series model in which multiple structural changes in level, trend, and error variance are modeled through specification of the prior distributions. Their model is a segmented heteroskedastic AR(k) model

$$y_t = \sum_{i=1}^{m+1} \left\{(\alpha_i + \beta_i t) I_{\{t_{i-1} \le t < t_i\}} + \sigma_i \epsilon_t\right\} + \sum_{j=1}^{k} \phi_j y_{t-j}, \tag{85.4}$$

with $t_0 = 1$, $t_{m+1} = T$, and m change points at unknown times $t_1, \ldots, t_m$ in the interval $[1, T]$. The ϵ_t in Eq. 85.4 are independent standard normal, and the α_i, β_i, σ_i, and ϕ_j are unknown parameters. Letting θ denote the vector consisting of these unknown parameters and $t_1, \ldots, t_m$, they put a prior distribution on θ and use Gibbs sampling to estimate its posterior distribution. Bayes factors, or posterior odds, or Schwarz's BIC can be used to select m, and they have found from simulation studies that Bayes factors and the BIC have satisfactory performance, while posterior odds are quite sensitive to the prior probabilities of models with different numbers of breaks.

McCulloch and Tsay (1993) consider AR(k) models with occasional level shifts

$$y_t = \mu_t + \sum_{i=1}^{k} \phi_i (y_{t-i} - \mu_{t-i}) + \eta_t, \tag{85.5}$$

in which η_t are independent $N(0, \sigma^2)$, or with occasional variance shifts

$$y_t = \mu + \sum_{i=1}^{k} \phi_i y_{t-i} + \sigma_t \epsilon_t, \tag{85.6}$$

in which ϵ_t are independent standard normal. The μ_t in Eq. 85.5 are piecewise constant so that $I_t := I_{\{\mu_t \neq \mu_{t-1}\}}$ are i.i.d. Bernoulli(p) with $P(I_t = 1) = p$. Similarly, in Eq. 85.6, $I_t := I_{\{\sigma_t \neq \sigma_{t-1}\}}$ are i.i.d. Bernoulli(p). To complete the prior specification of μ_t, McCulloch and Tsay (1993) write $\mu_t = \mu_{t-1} + z_t I_t$ and specify z_t to be a sequence of independent and identically distributed (i.i.d) normal random variables with variance τ^2 and independent of I_t. They also put conjugate priors on $\phi_1, \ldots, \phi_k$ and σ^2. Similarly, for Eq. 85.6, they write $\sigma_t = \sigma_{t-1} v_t^{I_t}$ and specify v_t^2 are i.i.d. inverted χ^2-random variables conditional on $(\mu, \phi_1, \ldots, \phi_k)$, which is assigned a conjugate prior distribution. Gibbs sampling is used to estimate the unknown parameters; after assigning "quite diffuse" prior distribution to the hyperparameters.

Lai et al. (2005) have introduced a more general change-point model than that of McCulloch and Tsay (1993). This change point has been reviewed in Sect. 93.3.3 of Lai and Xing (2010) and allows jumps in level, autoregressive parameters, and error variance, thereby incorporating both Eqs. 85.5 and 85.6. In addition, the model yields *explicit* recursive formulas for estimating the parameters and do not need Gibbs sampling or other Monte Carlo methods for its implementation. One may wonder why a more general model can be easier to implement. The key lies in the specification of the new parameter value at a change point. Specifically, McCulloch and Tsay (1993) assume $\mu_t = \mu_{t-1} + z_t I_t$ (or $\sigma_t = \sigma_{t-1} v_t^{I_t}$), while Lai et al. (2005) assume $\boldsymbol{\theta}_t = (1 - I_{t-1})\boldsymbol{\theta}_{t-1} + I_t \mathbf{z}_t$. Thus, the Bayesian model of Lai, Liu, and Xing samples a completely new value of the parameter vector at a change point while that of McCulloch and Tsay samples the *increment* $\mu_t - \mu_{t-1}$ or σ_t/σ_{t-1} at a change point. The great reduction in computational complexity by sampling a new parameter value over sampling an increment was first noticed by Yao (1984) in his simplification of the Bayesian model of Chernoff and Zacks (1964) for mean shifts in a Gaussian sequence of normal random variables with common known variance but piecewise constant means.

85.2.2 Stochastic Volatility Models with Jumps in Returns and Volatilities

To estimate the magnitude and assess the significance of stochastic volatility and jump risk premia in option pricing, contemporaneous jumps in prices and in stochastic volatilities have been incorporated in continuous-time models of asset price dynamics. Bakshi et al. (1997) have shown substantial improvement over including only jumps in prices, while Bates (2000) and others have found such improvement to be small. Further studies show unanimous support for jumps in prices but disagree on the importance of jumps in volatilities and even on how the volatility jumps should be modeled. Duffie et al. (2000) developed analytic methods for asset valuation and econometric analysis in the setting of affine jump-diffusion state processes. They pointed out that price jumps should depend on the size of volatility jumps via the jump intensities and introduced the following continuous-time stochastic volatility models that incorporate contemporaneous jumps in returns and volatilities for the asset price $S_t = e^{Y_t}$:

$$\begin{pmatrix} dY_t \\ dV_t \end{pmatrix} = \begin{pmatrix} \mu \\ \kappa(\theta - V_{t-})dt \end{pmatrix} + \sqrt{V}_{t-}\begin{pmatrix} 1 & 0 \\ \rho\sigma_v & \sqrt{1-\rho^2}\sigma_v \end{pmatrix} dW_t + \begin{pmatrix} \xi^y dN_t^y \\ \xi^v dN_t^v \end{pmatrix}, \tag{85.7}$$

where $V_{t-} = \lim_{s \uparrow t} V_s$, W_t is a standard 2-dimensional Brownian motion, N_t^y and N_t^v are Poisson processes with constant intensities λ_y and λ_v, and ξ^y and ξ^v are random jump sizes in returns and volatilities, respectively. Such specification covers most of the popular models used in portfolio allocation and option pricing. For example, the case $\lambda_y = \lambda_v = 0$ is equivalent to Heston's (1993) stochastic volatility model, and the case of $\lambda_v = 0$ and normal ξ^y is the same as Bates' (1996) model that has normally distributed jumps in returns. For the general model (85.7), Eraker et al. (2003) developed a simulation-based Bayes estimator, using Markov chain Monte Carlo (MCMC), of the jumps in returns and volatilities after discretizing Eq. 85.7 into

$$\begin{aligned} Y_{(t+1)\Delta} &= Y_{t\Delta} = \mu\Delta + \sqrt{V_{t\Delta}\Delta}\epsilon^y_{(t+1)\Delta} + \xi^y_{(t+1)\Delta}J^y_{(t+1)\Delta}, \\ V_{(t+1)\Delta} - V_{t\Delta} &= \kappa(\theta - V_{t\Delta})\Delta + \sigma_v\sqrt{V_{t\Delta}\Delta}\epsilon^v_{(t+1)\Delta} + \xi^v_{(t+1)\Delta}J^v_{(t+1)\Delta}, \end{aligned} \tag{85.8}$$

where $J^y_{(t+1)\Delta}$ (or $J^v_{(t+1)\Delta}$) is the indicator variable of a jump at $(t+1)\Delta$, $\epsilon^y_{(t+1)\Delta}$ and $\epsilon^v_{(t+1)\Delta}$ are standard normal random variables with correlation ρ, and Δ is the time-discretization interval. Letting $\Theta = (\theta_1, \ldots, \theta_k)$ denote the parameter vector and assuming $J^y = J^v$ (i.e., contemporaneous jumps), they note that the posterior density of Θ and the latent volatility, jump times, and jump sizes satisfy

$$f(\Theta, J, \xi^y, \xi^v, V|Y) \propto f(Y|\Theta, J, \xi^y, \xi^v, V) f(\Theta, J, \xi^y, , \xi^v, V)$$

and use the MCMC algorithm to generate samples by iteratively drawing from the following conditional posteriors till convergence to the stationary distribution:

$$\begin{aligned} &\textit{parameters}: f(\Theta_i|\Theta_{-i}, J, \xi^y, \xi^v, V, Y), \qquad i = 1, \ldots, k, \quad \text{with} \quad \Theta_{-i} = \Theta\backslash\{\theta_i\}, \\ &\textit{jump times}: f(J_{t\Delta} = 1|\Theta, \xi^y, \xi^v, V, Y), \qquad t = 1, \ldots, T \\ &\textit{jump sizes}: f\left(\xi^y_{t\Delta}|\Theta, J_{t\Delta} = 1, \xi^v, V, Y\right), f\left(\xi^v_{t\Delta}|\Theta, J_{t\Delta} = 1, \xi^v, V, Y\right), \quad t = 1, \ldots, T \\ &\textit{volatility}: f\left(V_{t\Delta}|V_{(t+1)\Delta}, V_{(t-1)\Delta}, \Theta, J, \xi^y, \xi^v, Y\right), \qquad t = 1, \ldots, T. \end{aligned}$$

Broadie et al. (2007) discussed the issue more thoroughly and found strong evidence in support of stochastic volatility and jumps in both price and volatility.

85.3 A Simple Stochastic Change-Point ARX-GARCH Model

The model (85.7), which is often referred to as SVCJ (stochastic volatility model with contemporaneous jumps in returns and volatilities), assumes Poisson arrivals of contemporaneous jumps in returns and volatilities with constant intensity. In discrete time, this is equivalent to $J_{t\Delta}$ being independent Bernoulli, which is assumed in the models of McCulloch and Tsay (1993) and Lai et al. (2005) discussed at the end of Sect. 85.2.1. What (85.8) adds to these models is the SV dynamics in the second equation of Eq. 85.8, in contrast to piecewise constant σ_t in Eq. 85.6 or in Lai et al. (2005). Because SV is a nonlinear state-space model, this makes the SVCJ model difficult to implement and also explains the need for MCMC algorithm. Replacing SV by a GARCH model achieves similar properties but obviates the need of filtering in a state-space model. To incorporate structural changes in the regression coefficients and the unconditional variance of the random disturbances in an autoregressive model with exogenous inputs (ARX), while allowing the conditional variances to follow a GARCH model, Lai and Xing (2013) consider the change-point ARX-GARCH model

$$y_t = \beta_t^T \mathbf{x}_t + v_t \sqrt{h_t} \epsilon_t, \tag{85.9}$$

in which the parameter vector β_t and the unconditional variance v_t^2 are piecewise constant, with jumps at times of structural change, the vector $\mathbf{x}_t$ consists of exogenous variables and the past observations $y_{t-1}, y_{t-2}, \ldots, y_{t-\kappa}$, and h_n represents short-term proportional fluctuations in variance generated by the GARCH model

$$h_n = \left(1 - \sum_{i=1}^{k} a_i - \sum_{l=1}^{k'} b_l\right) + \sum_{i=1}^{k} a_i w_{n-\mathrm{i}}^2 + \sum_{l=1}^{k'} b_l h_{n-l} \quad \text{with} \quad w_t = \sqrt{h_t}\epsilon_t. \tag{85.10}$$

The ϵ_t are assumed to be i.i.d. standard normal random variables such that ϵ_t are independent of $\mathbf{x}_t$ and the time-invariant GARCH parameters $a_1, \ldots, a_k$, $b_1, \ldots, b_{k'}$ are assumed to satisfy

$$a_i \geq 0, b_l \geq 0 \text{ and } \sum_{i=1}^{k} a_i + \sum_{l=1}^{k'} b_l \leq 1. \tag{85.11}$$

Letting $\tau_t = 1/(2v_t^2)$, we assume $\boldsymbol{\theta}_t = (\beta_t^T, \tau_t)^T$ to be piecewise constant and satisfy the following conditions:

(A1) For $t > t_0 = \max(k, k')$, the change times of $\boldsymbol{\theta}_t$ form a renewal process with i.i.d. inter-arrival times that are geometrically distributed with parameter p, or equivalently,

$$I_t := 1_{\{\boldsymbol{\theta}_t \neq \boldsymbol{\theta}_{t-1}\}} \text{ are i.i.d. } \text{Bernoulli}(p) \text{ with } P(I_t = 1) = p,$$

$I_{t_0} = 1$, and there is no change point prior to t_0.

(A2) $\boldsymbol{\theta}_t = (1 - I_t)\boldsymbol{\theta}_{t-1} + I_t(\mathbf{z}_t^T, \gamma_t)^T$, where $(\mathbf{z}_1^T, \gamma_1)^T$, $(\mathbf{z}_2^T, \gamma_2)^T$, ... are i.i.d. random vectors such that $\mathbf{z}_t|\gamma_t \sim$ Normal($\mathbf{z}$, $\mathbf{V}/(2\gamma_t)$), $\gamma_t \sim \chi_d^2/\rho$, where χ_d^2 denotes the chi-square distribution with d degrees of freedom.

(A3) The processes $\{I_t\}$, $\{\boldsymbol{\theta}_t\}$, and $\{(\mathbf{x}_t, \epsilon_t)\}$ are independent.

This model generalize the approach developed by Lai et al. (2005), who consider the special case of AR models with occasional jumps in regression parameters and error variances. The last paragraph of Sect. 93.3.3 of Lai and Xing (2010) has given a brief introduction of the model, and we provide more complete details here and also an updated account that includes some recent results.

85.3.1 Closed-Form Recursive Filters

Conditions (A1)–(A3) specify a Markov chain with unobserved states $(I_t, \boldsymbol{\theta}_t)$. The observations $(\mathbf{x}_t, y_t)$ are such that $(y_t - \beta_t^T \mathbf{x}_t)/v_t$ forms a GARCH process. This hidden Markov model (HMM) has hyperparameters p, $\mathbf{z}$, $\mathbf{V}$, ρ, d, $a_1, \ldots, a_k$, $b_1, \ldots, b_{k'}$. To estimate $\boldsymbol{\theta}_t$ assuming known hyperparameters, let $J_n = \max\{t \le n : I_n = 1\}$ and note that $n - J_n \ge k$ by

(A1). Define $\mathcal{Y}_n = (\mathbf{x}_1, y_1, \ldots, \mathbf{x}_n, y_n)$ and $\mathcal{Y}_{j,n} = \left(\mathbf{x}_j, y_j, \ldots, \mathbf{x}_n, y_n\right)$. The estimates $\hat{\beta}_n$ and $\hat{v}_n^2$ based on $\mathcal{Y}_n$ are weighted averages of $\hat{\beta}_{n,j}$ and $\hat{v}_{n,j}^2$ based on $\mathcal{Y}_{j,n}$, with the weights $p_{n,j}$ to be specified. The $\hat{\beta}_{n,j}$ and $\hat{v}_{n,j}^2$ can be computed recursively (with increasing n and fixed j). Initializing at $n = j - 1$ with $\hat{\beta}_{n,j} = \mathbf{z}$, $\hat{\mathbf{V}}_{n,j} = \mathbf{V}$, and $\hat{v}_{n,j}^2 = \rho/(2d)$, define for $n \ge j$

$$\hat{h}_{n,j} = \left(1 - \sum_{i=1}^{k} a_i - \sum_{l=1}^{k'} b_l\right) + \sum_{l=1}^{k'} b_l \hat{h}_{n-l,j} + \sum_{i=1}^{k} a_i \frac{\left(y_{n-i} - \hat{\beta}_{n-i,j}^T \mathbf{x}_{n-i}\right)^2}{\hat{v}_{n-i,j}^2} \tag{85.12a}$$

$$\mathbf{V}_{n,j} = \mathbf{V}_{n-1,j} - \left\{\mathbf{V}_{n-1,j}\mathbf{x}_n\mathbf{x}_n^T\mathbf{V}_{n-1,j} / \left(\hat{h}_{n,j} + \mathbf{x}_n^T\mathbf{V}_{n-1,j}\mathbf{x}_n\right)\right\}, \tag{85.12b}$$

$$\hat{\beta}_{n,j} = \hat{\beta}_{n-1,j} + \left\{\mathbf{V}_{n-1,j}\mathbf{x}_n\left(y_n - \hat{\beta}_{n-1,j}^T\mathbf{x}_n\right) / \left(\hat{h}_{n,j} + \mathbf{x}_n^T\mathbf{V}_{n-1,j}\mathbf{x}_n\right)\right\}, \tag{85.12c}$$

$$\hat{v}_{n,j}^2 = \frac{d+n-j-2}{d+n-j-1}\hat{v}_{n-1,j}^2 + \frac{1}{d+n-j-1} \cdot \frac{\left(y_n - \hat{\beta}_{n-1,j}^T\mathbf{x}_n\right)^2}{\hat{h}_{n,j} + \mathbf{x}_n^T\mathbf{V}_{n-1,j}\mathbf{x}_n}. \tag{85.12d}$$

The weights $p_{n,j}$ are given recursively by

$$p_{n,j} \propto p^*_{n,j} := \begin{cases} p\ f_{nn}/f_{00} & \text{if} \quad j = n, \\ (1-p)p_{n-1,j}f_{nj}/f_{n-1,j} & \text{if} \quad j \leq n-1 \end{cases} \tag{85.13}$$

where letting $\mathbf{z}_{n,j} = \mathbf{V}_{n,j}\left(\mathbf{V}^{-1}\mathbf{z} + \sum_{t=j}^{n} \mathbf{x}_t y_t/\hat{h}_{t,j}\right)$ and

$$\rho_{n,j} = \frac{1}{2}\rho + \mathbf{z}^T\mathbf{V}^{-1}\mathbf{z} - \mathbf{z}^T_{n,j}\mathbf{V}^{-1}_{n,j}\mathbf{z}_{n,j} + \sum_{t=j}^{n} y_t^2/\hat{h}_{t,j}, \tag{85.14}$$

$p_{n,j} = p^*_{n,j}/\sum_{j'=1}^{n} p^*_{n,j'}$, and the f_{nj} are given explicitly by

$$\begin{aligned} f_{nj} &= \left|\mathbf{V}_{n,j}\right|^{1/2}\Gamma((d+n-j+1)/2)\rho_{n,j}^{-(d+n-j+1)/2}, \\ f_{00} &= |\mathbf{V}|^{1/2}\Gamma(d/2)(\rho/2)^{-d/2}. \end{aligned} \tag{85.15}$$

These formulas are extensions of those for the special case $\mathbf{x}_t = (y_{t-1}, \ldots, y_{t-k})^T$ and $h_t \equiv 1$ considered by Lai et al. (2005). The extension from the case $h_t \equiv 1$ to more general known h_t basically amounts to extending ordinary least squares estimation (given the most recent change-point J_n) to generalized least squares estimation.

85.3.2 Estimation of Hyperparameters

The above Bayesian filter involves $\mathbf{z}$, $\mathbf{V}$, ρ, d, p, and $\boldsymbol{\eta} = (a_1, \ldots, a_k, b_1, \ldots, b_{k'})$. Note that $\mathbf{z}$ is the prior mean of β_t and $\rho/(d-2)$ is the prior mean of $2v^2$ at time t when parameter changes occur. As noted in Lai and Xing (2008), it is more convenient to represent the χ^2_d/ρ prior distribution for $(2v_t^2)^{-1}$ as a gamma $(d/2, \rho/2)$ distribution so that d does not need to be an integer. The recursions (85.12b) and (85.12c) are basically recursions for ridge regression which shrinks the generalized least squares estimate (using the weights $\hat{h}_{t,j}$) towards $\mathbf{z}$, with $\mathbf{V}^{-1}$ and $\sum_{t=j}^{n} \mathbf{x}_t\mathbf{x}_t^T$ being the matrix weights for the shrinkage target and the generalized least squares estimator, respectively. The shrinkage target $\mathbf{z}$ and its associated weight matrix $\mathbf{V}^{-1}$ are relevant when $n-j$ is small but become increasingly negligible with increasing $n-j$. We can estimate $\mathbf{z}$, $\mathbf{V}$, ρ, and d by applying the method of moments to the stationary distribution of the Markov chain $(I_t, \theta_t, \epsilon_t)$ that is partially observed via $(\mathbf{x}_t, y_t)$. Details are given in the next paragraph. With $\mathbf{z}$, $\mathbf{V}$, ρ, and d replaced by these estimates, we then estimate $\boldsymbol{\eta}$ and p by maximum likelihood, noting that the log-likelihood function ℓ_n based on $y_1, \ldots, y_n$ has the representation

$$\log\ \ell_n(\boldsymbol{\eta}, p) = \sum_{t=1}^{n} \log \left[\sum_{j=1}^{t} p^*_{t,j}(\boldsymbol{\eta}, p)\right], \tag{85.16}$$

where $p^*_{t,j}$ is given by Eq. 85.13.

Lai and Xing (2013) use the method-of-moments estimates of $\mathbf{z}$, $\mathbf{V}$, ρ, and d based on $(\mathbf{x}_t, y_t)$, $1 \le t \le n$. From (A2) and (A3), it follows that $E(\beta_t) = \mathbf{z}$, $\mathrm{Cov}(\beta_t) = (Ev_t^2)\mathbf{V}$, and $E(\mathbf{x}_t y_t) = \mathbf{x}_t \mathbf{x}_t^T \mathbf{z}$. From $n - L$ moving windows $\{(\mathbf{x}_t, y_t) : s \le t \le s + L\}$ of these data, compute the least squares estimates

$$\hat{\beta}^{(s)} = \left(\sum_{s \le t \le s+L} \mathbf{x}_t \mathbf{x}_t^T \right)^{-1} \sum_{s \le t \le s+L} \mathbf{x}_t y_t. \tag{85.17}$$

Each $\hat{\beta}^{(s)}$ is a method-of-moments estimate of $\mathbf{z}$, and so is $\hat{\beta} = (n - L)^{-1} \sum_{s=1}^{n-L} \hat{\beta}^{(s)}$. If an oracle would reveal the change times up to time n, then one would segment the time series accordingly and use the least squares estimate for each segment to estimate the regression parameter for that segment. The average of these least squares estimates over the segment would provide a method-of-moments estimate of $\mathbf{z}$. Similarly, the average squared residual in each segment is a method-of-moments estimate of $E(v_t^2) = \rho/[2(d - 2)]$, and so is the average of these values over the segments; see Engle and Mezrich (1996). In ignorance of the change points, the segments are replaced by moving windows of length $L + 1$ in Eq. 85.9 and estimate $\mathbf{z}$ by the average $\overline{\beta}$ of the $\hat{\beta}^{(s)}$. Likewise ρ and d can be estimated by equating the mean and variance of the inverted gamma distribution for v_t^2 to their sample counterparts for the average squared residuals:

$$\begin{aligned} \frac{\hat{\rho}}{\left[2\left(\hat{d} - 2\right)\right]} &= \overline{v} := (n - L)^{-1} \sum_{s=1}^{n-L} \left[\sum_{s \le t \le s+L} \left(y_t - \mathbf{x}_t^T \hat{\beta}^{(s)} \right)^2 / (L + 1) \right], \\ \frac{\hat{\rho}^2}{\left[\left(2\left(\hat{d} - 2\right)\right)^2 \left(\hat{d} - 4\right) \right]} &= (n - L)^{-1} \sum_{s=1}^{n-L} \left[\sum_{s \le t \le s+L} \left(y_t - \mathbf{x}_t^T \hat{\beta}^{(s)} \right)^2 / (L + 1) - \overline{v} \right]^2. \end{aligned} \tag{85.18}$$

Similarly Lai and Xing (2013) estimate $\mathbf{V}$ by

$$\hat{\mathbf{V}} = \left[2\left(\hat{d} - 2\right)/\hat{\rho}\right](n - L)^{-1} \sum_{s=1}^{n-L} \left(\hat{\beta}^{(s)} - \overline{\beta} \right) \left(\hat{\beta}^{(s)} - \overline{\beta} \right)^T. \tag{85.19}$$

85.3.3 Bounded Complexity Mixture Approximations

Although Eq. 85.13 provides closed-form recursions for updating the weights $p_{t,i}$, $1 \le i \le t$, the number of weights increases with t, resulting in rapidly increasing computational complexity and memory requirements for estimating θ_n as n increases. A natural idea to reduce the complexity and to facilitate the use of parallel algorithms for the recursions is to keep only a fixed number M of weights at every stage n (which is tantamount to setting the other weights to be 0).

Lai and Xing (2013) keep the most recent m weights $p_{n,i}$ (with $n - m < i \leq n$) and the largest $M - m$ of the remaining weights, where $1 \leq m < M$. Specifically, let $\mathcal{K}_{n-1}$ denote the set of indices i for which $p_{n-1,i}$ is kept at stage $n - 1$; thus $\mathcal{K}_{n-1} \supset \{n-1,, \ldots, n-m\}$. At stage n, define $p^*_{n,i}$ by Eq. 85.5 for $i \in \{n\} \cup \mathcal{K}_{n-1}$, and let i_n be the index not belonging to $\{n, n - 1, \ldots, n - m + 1\}$ such that

$$p^*_{n,i_n} = \min\left\{p^*_{n,i} : j \in \mathcal{K}_{n-1} \quad \text{and} \quad j \leq n - m\right\}, \tag{85.20}$$

choosing i_n to be the one farthest from n if the minimizing set in (85.12) has more than one element. Define $\mathcal{K}_n = \{n\} \cup (\mathcal{K}_{n-1} - \{i_n\})$, and let

$$p_{n,i} = \frac{p^*_{n,i}}{\sum_{j \in \mathcal{K}_n} p^*_{n,j}}. \tag{85.21}$$

Lai and Xing (2013) use these bounded complexity mixtures (BCMIX) not only to approximate the filters $(\beta_t, v_t)|\mathcal{Y}_t$ but also to approximate the likelihood function (85.8), in which we replace $\sum_{j=1}^{t}$ by $\sum_{j \in \mathcal{K}_t}$. They use a grid of the form $\{2^j/n : j_0 \leq j \leq j_1\}$, where $j_0 < 0 < j_1$ are integers, to search for the maximum $\hat{p}_n$ of $\ell_n(p; \hat{\eta}_n)$ over the grid. Letting $\lambda = (p, \boldsymbol{\eta}), \hat{\lambda}_n = (p_n, \hat{\eta}_n)$ is used to replace λ in the recursions (85.4) and (85.5). The update $\hat{\boldsymbol{\eta}}_n$ of the GARCH parameters after observing $(\mathbf{x}_n, y_n)$ uses simply a single iteration of the Newton-Raphson iteration procedure to maximize $\ell_n(\hat{p}_{n-1}, \boldsymbol{\eta})$ when $n \geq n_0$ and uses more iterations until convergence for small n. Therefore relatively fast updates of the hyperparameters estimates can be used to implement the adaptive BCMIX filters.

85.3.4 Sequential BCMIX Forecasts

The AR-GARCH model is often used to forecast future returns and their volatilities for portfolio optimization and risk management; see Sects. 6.4.1 and 12.2.3 of Lai and Xing (2008). Incorporating exogenous inputs and change points into the model improves the forecasts. For the change-point ARX-GARCH model, first assume that the hyperparameters are known. Since $\mathbf{x}_{n+1}$ consists of $y_n, \ldots, y_{n-\kappa+1}$ and other input variables up to time n and since time $n + 1$ has prior probability p of being a change point, the forecast $\hat{y}_{n+1|n}$ of y_{n+1} given $\mathcal{Y}_n$ is related to the filter $\hat{\beta}_n$ given in Sect. 85.1 by

$$\hat{y}_{n+1|n} = p\mathbf{z}^T\mathbf{x}_{n+1} + (1-p)\hat{\beta}_n^T\mathbf{x}_{n+1}, \tag{85.22}$$

noting that β_{n+1} is equal to β_n with probability $1 - p_{n,n}$. Similarly, the forecast $\hat{v}^2_{n+1|n}$ of v^2_{n+1} given $\mathcal{Y}_n$ is

$$\hat{v}^2_{n+1|n} = \frac{pp}{2(d-2)} + (1-p)\hat{v}^2_{n|n}, \tag{85.23}$$

assuming $d > 4$ in view of Eq. 85.10. Note that the conditional variance in the GARCH model involves $v_n^2 h_n$ rather than v_n^2. Lai and Xing (2013) can use Eq. 85.4a with n replaced by $n + 1$ to forecast h_{n+1}. In particular, for GARCH (1,1), they forecast h_{n+1} by $h_{n+1,j} = (1-a-b) + b\hat{h}_{n,j} + a\left(y_n - \hat{\beta}^T_{n,j}\mathrm{x}_n\right)^2/\hat{v}^2_{n,j}$ and use the weights $p_{n,j}$ in Eq. 85.13 to weight

$$\hat{v}^2_{n+1,j}\hat{h}_{n+1,j} = \left\{\frac{p\rho}{2(d-2)} + (1-p)\hat{v}^2_{n,j}\right\}\left\{(1-a-b) + b\hat{h}_{n,j} + \frac{a\left(y_n - \hat{\beta}^T_{n,j}\mathbf{x}_n\right)^2}{\hat{v}^2_{n,j}}\right\}. \tag{85.24}$$

85.3.5 BCMIX Smoothers

Lai and Xing (2013) begin by deriving the Bayes estimate (smoother) of $\boldsymbol{\theta}_t = (\beta_t^T, \tau_t)^T$ given $\mathcal{Y}_n$ for $1 \le t \le n$ in the "oracle" setting in which the h_t are specified exactly (by the oracle) so that there are explicit recursive representations of the posterior mean of $\boldsymbol{\theta}_t$ given $\mathcal{Y}_n$ for $1 \le t \le n$. To obtain the optimal smoother $E(\theta_t|\mathcal{Y}_n)$ for $1 \le t \le n$, they use Bayes theorem to combine the forward filter $\theta_t|\mathcal{Y}_t$ with the backward filter $\theta_t|\mathcal{Y}_{t+1,n}$. Because the h_t are assumed known in $\left(y_t - \beta_t^T\mathbf{x}_t\right)/\sqrt{h_t} = v_t\epsilon_t$ and the ϵ_t are i.i.d. standard normal, the backward filter has the same form as the forward filter. In fact, assumptions (A1)–(A3) define a reversible Markov chain of jump times and jump magnitudes, assuming $I_{n-t_0+1} = 1$ and no change points afterwards. Let π denote the density function of the stationary distribution. Letting $\widetilde{J}_{t+1} = \min\{s \ge t+1 : I_s = 1\}$ and $q_{t+1,j} = P\left(\widetilde{J}_{t+1} = j|\mathcal{Y}_{t+1,n}\right)$ for $j \ge t + 1$, one can reverse time and obtain a backward filter that is similar to the forward filter:

$$f\left(\boldsymbol{\theta}_t|\mathcal{Y}_{t+1,n}\right) = p\pi(\boldsymbol{\theta}_t) + (1-p) \sum_{j=t+1}^{n} q_{t+1,j} f\left(\boldsymbol{\theta}_{t+1}|\mathcal{Y}_{t+1,n}, \widetilde{J}_{t+1} = j\right),$$

in which

$$q_{t+1},j \propto q^*_{t+1,j} = \begin{cases} pf_{jj}/f_{00} & \text{if } j = t+1, \\ (1-p)q_{t+2,j} f_{t+1,j}/f_{t+2,j} & \text{if } j \ge t+2. \end{cases}$$

Application of Bayes theorem then yields

$$\begin{aligned} &f(\boldsymbol{\theta}_t|\mathcal{Y}_n) \propto f(\boldsymbol{\theta}_t|\mathcal{Y}_t) f\left(\boldsymbol{\theta}_t|\mathcal{Y}_{t+1,n}\right)/\pi(\boldsymbol{\theta}_t) \\ &\propto p\sum_{i=1}^{t} p_{i,t} f(\boldsymbol{\theta}_t|\mathcal{Y}_n, C_{it}) + (1-p) \sum_{1 \le i \le t < j \le n} p_{i,t} q_{t+1,j} \frac{f_{ij} f_{00}}{f_{it} f_{t+1,j}} f\left(\boldsymbol{\theta}_t|\mathcal{Y}_n, C_{ij}\right), \end{aligned}$$

where $C_{ij} = \{I_i = 1, I_{i+1} = \cdots = I_j = 0, I_{j+1} = 1\}$. Let $\boldsymbol{\tau}_n = (\tau_{j,i}, 1 \le i \le j \le n)$. The optimal smoother for β_t and v_t^2 are given by

$$\beta_t|\tau_n, \mathcal{Y}_n \sim \sum_{1 \le i \le t \le j \le n} \alpha_{ijt} N\left(\mathbf{z}_{j,i}, \mathbf{V}_{j,i}/\left(2\tau_{j,i}\right)\right), \quad \tau_{j,i}|\mathcal{Y}_n \sim \chi^2_{d+j-i+1}/\rho_{j,i}, \tag{85.25}$$

where $\mathbf{V}_{j,i}$, $\mathbf{z}_{j,i}$, and $\rho_{j,i}$ are same as those defined in Eqs. 85.4 and 85.6 and α_{ijt} can be determined recursively by

$$\begin{aligned} \alpha_{ijt} &= \alpha^*_{ijt}/A_t, \qquad A_t = \sum_{1 \le i \le t \le j \le n} \alpha^*_{ijt}, \\ \alpha^*_{ijt} &= \begin{cases} pp_{i,t}, & i \le t, j = t, \\ ap_{i,t} q_{t+1,j} f_{00} f_{ij}/f_{it} f_{t+1,j}, & i \le t, j > t. \end{cases} \end{aligned} \tag{85.26}$$

Moreover, the posterior probability of having a change point at time t is given by

$$P(I_{t+1} = 1|\mathcal{Y}_n) = \sum_{1 \le i \le t} P(C_{it}|\mathcal{Y}_n) = p/A_t.$$

The next step is to approximate α_{ijt} by $\hat{\alpha}_{ijt}$ that replaces the h_t, which is actually unknown, by the estimates $\hat{h}_{j,i}$ defined recursively for $j \ge i$ by Eq. 85.12a. As in Sect. 85.3.1, we assume known hyperparameters p and $\boldsymbol{\eta}$ for the time being. Using the BCMIX approximation to the forward and backward filters, we approximate the sum in Eq. 85.25 and that defining A_t in Eq. 85.26 by

$$\beta_t|\tau_n, \mathcal{Y}_n \sim \sum_{i \in \mathcal{K}_t, j \in \{t\} \cup \widetilde{\mathcal{K}}_{t+1}} \alpha_{ijt} N\left(\mathbf{z}_{j,i}, \mathbf{V}_{j,i}/\left(2\tau_{j,i}\right)\right), \tau_{j,i}|\mathcal{Y}_n \sim \chi^2_{d+j-i+1}/\rho_{j,i} \tag{85.27}$$

where $\mathcal{K}_t$ is the same as that in Sect. 85.3.3 for the forward filter and $\widetilde{\mathcal{K}}_{t+1}$ is the corresponding set for the backward filter. Assuming known $\boldsymbol{\eta}$ and p, the BCMIX estimates for β_t and v_t given $\mathcal{Y}_n$ are

$$\begin{aligned} \hat{\beta}_{t|n} &= \sum_{i \in \mathcal{K}_t, j \in \{t\} \cup \widetilde{\mathcal{K}}_{t+1}} \alpha_{ijt} \mathbf{z}_{j,i}, \qquad \hat{v}^2_{t|n} = \sum_{i \in \mathcal{K}_t, j \in \{t\} \cup \widetilde{\mathcal{K}}_{t+1}} \alpha_{ijt} \frac{\rho_{j,i}}{d + j - i - 1}, \\ \hat{\tau}_{t|n} &= \sum_{i \in \mathcal{K}_t, j\{t\} \cup \widetilde{\mathcal{K}}_{t+1}} \alpha_{ijt}(d + j - i + 1)/\left(2\rho_{j,i}\right). \end{aligned} \tag{85.28}$$

The conditional probability of a change point at time $t(\le n)$ given $\mathcal{Y}_n$ is estimated by

$$\hat{P}(I_{t+1} = 1|\mathcal{Y}_n) = p/A_t. \tag{85.29}$$

Without assuming p and $\boldsymbol{\eta}$ to be known, we can use the BCMIX approximation in the log-likelihood function (85.16) based on $\mathcal{Y}_n$ and evaluate its maximizer $(\hat{p}, \hat{\boldsymbol{\eta}})$. Replacing $(p, \boldsymbol{\eta})$ by $(\hat{p}, \hat{\boldsymbol{\eta}})$ in Eq. 85.28 yields the BCMIX empirical Bayes smoother.

85.3.6 Application to Segmentation

In principle, the frequentist approach to multiple change-point problems for regression models reviewed in Sect. 85.1 can be extended to ARX-GARCH models by maximizing the log-likelihood over the locations of the change points and the piecewise constant parameters when it is assumed that there are k change points. This optimization problem, however, is much more difficult than that for regression models and only constitutes an inner loop of an algorithm whose outer loop is another minimization, over k, of a suitably chosen model selection criterion to determine k. For computational and analytic tractability, the frequentist approach typically assumes that k is small relative to n and that adjacent change points are sufficiently far apart so that the segments are identifiable except for relatively small neighborhoods of change points; see Bai and Perron (1998). Lai and Xing (2011) formulate these assumptions for the piecewise constant parameter vectors $\boldsymbol{\theta}_t$ as follows:

(B1) The true change points occur at $t_1^{(n)} < \cdots < t_k^{(n)}$ such that $\liminf_{n \to \infty} n^{-1} (t_i^{(n)} - t_{i-1}^{(n)}) > 0$ for $1 \le i \le k+1$, with $t_0^{(n)} = 0$ and $t_{k+1}^{(n)} = n$.

(B2) There exists $\delta > 0$, which does not depend on n, such that $\min_{1 \le i \le k} \left\| \boldsymbol{\theta}_{t_i^{(n)}} - \boldsymbol{\theta}_{t_{i-1}^{(n)}} \right\| \ge \delta$ for all large n.

In the context of ARX-GARCH models, Lai and Xing (2013) also assume that the stochastic regressors satisfy the stability condition:

(B3) $\max_{1 \le t \le n} \|\mathbf{x}_t\|^2 / n \xrightarrow{P} 0$ and $\sum_{t=1}^{n} \mathbf{x}_t \mathbf{x}_t^T / n$ converge almost surely to a positive definite nonrandom matrix.

Under (B1)–(B3) and assuming that $m \sim |\log n|^{1+\epsilon}$ and $M - m = O(1)$ as $n \to \infty$, for some $\epsilon > 0$, Lai and Xing (2013) have shown that the BCMIX smoother $\hat{\boldsymbol{\theta}}_{t|n}$ satisfies

$$\max_{1 \le t \le n : \min_{1 \le i \le k} \left| t - t_i^{(n)} \right| \ge m} \left\| \hat{\boldsymbol{\theta}}_{t|n} - \boldsymbol{\theta}_t \right\| \to 0, \quad \text{as } n \to \infty \tag{85.30}$$

uniformly in $a_1/n \le p \le a_2/n$. They apply this result to estimate the change times $t_1^{(n)}, \ldots, t_k^{(n)}$ in (B1) as follows. Let

$$\Delta_t = \left\| \hat{\boldsymbol{\theta}}_{t+m} - \hat{\boldsymbol{\theta}}_{t-m} \right\|^2, \tag{85.31}$$

and let $\hat{\tau}_1$ be the maximizer of Δ_t over $m < t < n - m$. After $\hat{\tau}_1, \ldots, \hat{\tau}_{j-1}$ have been defined, define

$$\hat{\tau}_j = \arg \max_{t:m<t<n-m,\, \min_{1\le i\le j-1}|t-\hat{\tau}_i|\ge m} \Delta_t; \tag{85.32}$$

these estimates of the change times are unordered and do not depend on the number k of change points. Assuming that there are k change points, they order $\hat{\tau}_1, \ldots, \hat{\tau}_k$ as $\hat{t}_{(1),k} < \cdots < \hat{t}_{(k),k}$ to provide estimates of $t_1^{(n)} < \cdots < t_k^{(n)}$. Let $\boldsymbol{\theta}^{(j)}$ be the common value of $\boldsymbol{\theta}_t$ in the interval $t_{j-1}^{(n)} \le t < t_j^{(n)}$. Assuming $t_1^{(n)}, \ldots, t_k^{(n)}$ to be known, the parameter vectors $\boldsymbol{\theta}^{(1)}, \ldots, \boldsymbol{\theta}^{(k+1)}$ can be estimated by maximum likelihood. Replacing the $t_1^{(n)}, \ldots, t_k^{(n)}$ by the estimates $\hat{t}_{(1),k} < \cdots < \hat{t}_{(k),k}$ in these MLEs leads to the quasi-likelihood estimators $\widetilde{\boldsymbol{\theta}}^{(1)}, \ldots, \widetilde{\boldsymbol{\theta}}^{(k+1)}, \widetilde{\boldsymbol{\eta}}$, and the quasi-likelihood

$$\Lambda_n(k) = \sum_{j=1}^{k+1} \sum_{t=\hat{t}_{(j-1),k}}^{\hat{t}_{(j),k}-1} \log f\left(y_t; \boldsymbol{\theta}^{(j)}, \boldsymbol{\eta}\right) \tag{85.33}$$

in which $f(\cdot;\boldsymbol{\theta}, \boldsymbol{\eta})$ in the density function of y_t given the piecewise constant parameter values, noting that $\left(y_t - \beta_t^T \mathbf{x}_t\right)/\left(v_t\sqrt{h_t}\right)$ is standard normal. Assuming a known upper bound K on the number k of change points in (B1), Lai and Xing (2013) propose to estimate k by

$$\hat{k}_n = \arg \max_{1\le k\le K} \{\Lambda_n(k) - (k+1)C_n\},$$

where C_n is a penalty term that satisfies

$$C_n \to \infty \text{ and } C_n/n \to 0 \text{ as } n \to \infty; \tag{85.34}$$

making use of Theorem 1, they have shown that $\hat{k}_n \xrightarrow{P} k$, $\widetilde{\boldsymbol{\eta}} \xrightarrow{P} \boldsymbol{\eta}$, and $\widetilde{\boldsymbol{\theta}}^{(j)} \xrightarrow{P} \boldsymbol{\theta}^{(j)}$ for $1 \le j \le k + 1$. Their simulation studies show that this segmentation procedure performs well in frequentist and Bayesian scenarios.

85.4 Discussion

The idea of representing the GARCH model by $v_t\sqrt{h_t}\epsilon_t$, in which v_t^2 is the unconditional variance and h_t follows the GARCH dynamics in Eq. 85.10, has also been used by Engle and Rangel (2008) in their spline-GARCH model that uses a deterministic function of time and exogenous variables to model v_t by

$$\log v_t = \beta^T \mathbf{x}_t + \phi_0 t + \sum_{i=0}^{I} \phi_i (t - t_i)_+^2.$$

The change-point ARX-GARCH model uses a piecewise constant function to model v_t instead and relates the exogenous variables $\mathbf{x}_t$ to y_t via the regression model (85.9), allowing contemporaneous jumps in the regression coefficients and the unconditional variances. As in Lai et al. (2005) who consider the special case $h_t \equiv 1$ and $\mathbf{x}_t = (y_{t-1}, \ldots, y_{t-\kappa})^T$, the stochastic model (85.9) for jumps in (β_t, v_t) involves linear Bayes methods and conjugate priors, yielding BCMIX approximations to the Bayes estimate of (β_t, v_t^2). The BCMIX approximations also provide estimates of the hyperparameters in the Bayesian model with relatively low computational complexity, yielding empirical Bayes estimates of the piecewise constant parameters that are efficient from both computational and statistical viewpoints. Section 85.6.6 shows how the computationally attractive EB estimates can be used to address the challenging frequentist problem of segmentation.

The empirical study in Lai and Xing (2013) of weekly returns of SP500 index, from the week starting on January 2, 1990, to the week starting on August 24, 2009, shows that segmenting the data by the method in Sect. 85.3.6 can remove the spurious long memory in volatility exhibited by fitting the AR-GARCH model to the entire time series without incorporating possible parameter changes during the long period that has undergone several structural changes. The apparent long memory arises from the (long) timescale for parameter changes. The segments are more general than the "regimes" in regime-switching volatility models (which are HMMs) reviewed in Sect. 85.2.1, in which difficulties in estimating the hyperparameters are noted. To address these difficulties, Gray (1996, pp. 35–36) modifies the usual regime-switching GARCH model by aggregating the conditional variances from different regimes at each time step. In Lai and Xing's segmentation approach, the GARCH parameters are separately estimated for different segments. However, to determine the segments using the empirical Bayes estimates, the Bayesian model assumes changes only in the unconditional variance v_t^2 but not in the GARCH parameters $a_1, \ldots, a_k, b_1, \ldots, b_{k'}$. Not only does this model circumvent the computational difficulties of regime-switching GARCH (or even ARCH) models noted by Cai (1994) and Gray (1996), but it also captures the short-run dynamics of the conditional variance and the structural changes of the long-run volatility. Although not allowing the GARCH parameters to change over time may appear too restrictive, one can in fact estimate them and the other hyperparameter p from moving windows of current and past data, instead of from the entire past history as in Eq. 85.16, thereby implicitly allowing these hyperparameters to change slowly over time.

References

Albert, J. H., & Chib, S. (1993). Bayes inference via Gibbs sampling of autoregressive time series subject to Markov mean and variance shifts. *Journal of Business and Economic Statistics, 11*, 1–15.

Andrews, D. W. K., Lee, I., & Ploberger, W. (1996). Optimal changepoint tests for normal linear regression. *Journal of Econometrics, 70*, 9–36.

Bai, J. (1997). Estimation of a changepoint in multiple regression models. *Review of Economics and Statistics, 79*, 551–563.

Bai, J. (1999). Likelihood ratio tests for multiple structural changes. *Journal of Econometrics, 91*, 299–323.

Bai, J., & Perron, P. (1998). Testing for and estimation of multiple structural changes. *Econometrica, 66*, 817–858.

Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics, 18*, 1–22.

Bai, J., Lumsdaine, R., & Stock, J. H. (1998). Testing for and dating common breaks in multivariate time series. *The Review of Economic Studies, 65*, 395–432.

Baillie, R. T., Bollerslev, T., & Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics, 74*, 3–30.

Bakshi, G., Cao, C., & Chen, Z. (1997). Empirical performance of alternative option pricing models. *Journal of Finance, 52*, 2003–2049.

Bates, D. S. (1996). Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche Mark options. *Review of Financial Studies, 9*, 69–107.

Bates, D. S. (2000). Post-87 crash fears in the s&p 500 futures option market. *Journal of Econometrics, 94*, 181–238.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics, 31*, 307–327.

Broadie, M., Chernov, M., & Johannes, M. (2007). Model specification and risk premia: Evidence from futures options. *Journal of Finance, 62*, 1453–1490.

Cai, J. (1994). A Markov model of switching-regime ARCH. *Journal of Business and Economic Statistics, 12*, 309–316.

Chernoff, H., & Zacks, S. (1964). Estimating the current mean of a normal distribution which is subjected to change in time. *The Annals of Mathematical Statistics, 35*, 999–1018.

Diebold, F. X. (1986). Comment on modeling the persistence of conditional variance. *Econometric Reviews, 5*, 51–56.

Duffie, D., Pan, J., & Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica, 68*, 1343–1376.

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of variance of United Kingdom inflation. *Econometrica, 50*, 987–1008.

Engle, R. F., & Bollerslev, T. (1986). Modeling the persistence of conditional variances. *Econometric Reviews, 5*, 1–50.

Engle, R. F., & Mezrich, J. (1996). GARCH for groups. *Risk, 1996*, 36–40.

Engle, R. F., & Rangel, J. G. (2008). The spline-GARCH models for low-frequency volatility and its global macroeconomic causes. *Review of Financial Studies, 21*, 1187–1222.

Eraker, B., Johannes, M. S., & Polson, N. G. (2003). The impact of jumps in returns and volatility. *Journal of Finance, 53*, 1269–1300.

Goldfeld, S. M., & Quandt, R. E. (1973). A Markov model for switching regressions. *Journal of Econometrics, 1*, 3–16.

Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics, 42*, 27–62.

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica, 57*, 357–384.

Hamilton, J. D. (1990). Analysis of time series subject to changes in regimes. *Journal of Econometrics, 45*, 39–70.

Hamilton, J. D., & Susmel, R. (1994). A conditional heteroskedasticity and change in regime. *Journal of Econometrics, 64*, 307–333.

Heston, S. (1993). A closed-form solutions for options with stochastic volatility. *Review of Financial Studies, 6*, 327–343.

Lai, T. L., & Xing, H. (2008). *Statistical models and methods for financial markets*. New York: Springer.

Lai, T. L., & Xing, H. (2010). Time series modeling and forecasting of volatilities of asset returns. In C. F. Lee, A. C. Lee, & J. Lee (Eds.), *Handbook of quantitative finance and risk management* (pp. 1417–1426). New York: Springer.

Lai, T. L., & Xing, H. (2011). A simple Bayesian approach to multiple change-points. *Statistica Sinica, 21*, 539–569.

Lai, T. L., & Xing, H. (2013). Stochastic change-point ARX-GARCH models and their applications to econometric time series. *Statistica Sinica, 23*, 1573–1594.

Lai, T. L., Liu, H., & Xing, H. (2005). Autoregressive models with piecewise constant volatility and regression parameters. *Statistica Sinica, 15*, 279–301.

Lamoureux, C. G., & Lastrapes, W. D. (1990). Persistence in variance, structural change and the GARCH model. *Journal of Business and Economic Statistics, 8*, 225–234.

McCulloch, R. E., & Tsay, R. S. (1993). Bayesian inference and prediction for mean and variance shifts in autoregressive time series. *Journal of the American Statistical Association, 88*, 968–978.

Perron, P. (1989). The great crash, the oil price shock and the unit root hypothesis. *Econometrica, 57*, 1361–1401.

Qu, Z., & Perron, P. (2007). Estimating and testing structural changes in multivariate regressions. *Econometrica, 75*, 459–502.

Quandt, R. E. (1958). The estimation of the parameters of a linear regression system obeying two separate regimes. *Journal of the American Statistical Association, 53*, 873–880.

Quandt, R. E. (1960). Tests of the hypothesis that a linear regression system obeys two separate regimes. *Journal of the American Statistical Association, 55*, 324–330.

Rappoport, P., & Reichlin, L. (1989). Segmented trends and non-stationary time series. *The Economic Journal, 99*, 168–177.

So, M. K. P., Lam, K., & Li, W. K. (1998). A stochastic volatility model with Markov switching. *Journal of Business and Economic Statistics, 16*, 244–253.

Wang, J., & Zivot, E. (2000). A Bayesian time series model of multiple structural changes in level, trend, and variance. *Journal of Business and Economic Statistics, 18*, 374–386.

Yao, Y. C. (1984). Estimation of a noisy discrete-time step function: Bayes and empirical Bayes approaches. *The Annals of Statistics, 12*, 1434–1447.

Unspanned Stochastic Volatilities and Interest Rate Derivatives Pricing

86

Feng Zhao

Contents

Abstract

This paper first reviews the recent literature on the unspanned stochastic volatilities (USV) documented in the interest rate derivatives markets. The USV refers to the volatility factors implied in the interest rate derivatives prices that have little correlation with the yield curve factors. We then present the result in Li and Zhao (*Journal of Finance, 62*, 345–382, 2006) that a sophisticated DTSM without USV feature can have serious difficulties in hedging caps and cap straddles, even though they capture bond yields well. Furthermore, at-the-money straddle hedging errors are highly correlated with cap implied volatilities and can explain a large fraction of hedging errors of all caps and straddles across

F. Zhao
University of Texas at Dallas, Richardson, TX, USA
e-mail: feng.zhao@utdallas.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_86,

moneyness and maturities. These findings strongly suggest that the unmodified dynamic term structure models, assuming the same set of state variables for both bonds and derivatives, are seriously challenged in capturing the term structure volatilities. We also present a multifactor term structure model with stochastic volatility and jumps that yields a closed-form formula for cap prices from Jarrow et al. (*Journal of Finance, 61*, 341–378, 2007). The three-factor stochastic volatility model with Poisson jumps can price interest rate caps well across moneyness and maturity. Last we present the nonparametric estimation results from Li and Zhao (2009). Specifically, the forward densities depend significantly on the slope and volatility of LIBOR rates, and mortgage market activities have strong impacts on the shape of the forward densities. These results provide nonparametric evidence of unspanned stochastic volatility and suggest that the unspanned factors could be partly driven by activities in the mortgage markets. These findings reinforce the claim that term structure models need to accommodate the unspanned stochastic volatilities in pricing and hedging interest rate derivatives.

The econometric methods in this chapter include extended Kalman filtering, maximum likelihood estimation with latent variables, local polynomial method, and nonparametric density estimation.

Keywords

Term structure modeling • Interest rate volatility • Heath-Jarrow-Morton model • Nonparametric density estimation • Extended Kalman filtering

86.1 Introduction

Interest rate caps and swaptions are among the most widely traded interest rate derivatives in the world. According to the Bank for International Settlements, their combined notional values are more than ten trillion dollars in recent years, which are many times bigger than that of exchange-traded options. Because of the size of these markets, accurate and efficient pricing and hedging of caps and swaptions have enormous practical importance. Pricing interest rate derivatives are more demanding than pricing bonds in that the derivatives are more sensitive to the higher-order moments of the distributions for underlying and therefore the models need to be able to capture the interest rate volatilities as well as the interest rates themselves. Under the unified framework of the dynamic term structure models (hereafter DTSMs), a benchmark in the term structure literature, the same set of risk factors are used in pricing bonds and derivatives. Consequently, the set of risk factors can be indentified with the observations of bond yields or swap rates, while the inclusion of derivative prices can help in terms of the efficiency of the estimation but not essential. The practitioners, on the other hand, generally apply the Heath-Jarrow-Morton (HJM) type of models in pricing interest rate derivatives, in which the entire yield curve is taken as given, and sometimes factors independent of yield curve, such as stochastic volatilities and jumps, are added in a piece-meal approach. This divergence foreshadows one of the key issues of the fast-growing literature on LIBOR-based interest rate derivatives, the so-called unspanned stochastic volatility (USV) puzzle.[1]

Interest rate caps and swaptions are derivatives written on LIBOR and swap rates, and the traditional view is that their prices be determined by the same set of risk factors that determine LIBOR and swap rates. However, several recent studies have shown that there seem to be risk factors that affect the prices of caps and swaptions but are not spanned by the underlying LIBOR and swap rates. Heidari and Wu (2003) show that while the three common term structure factors (i.e., the level, slope, and curvature of the yield curve) can explain 99.5 % of the variations of bond yields, they explain less than 60 % of swaption implied volatilities. After including three additional volatility factors, the explanatory power is increased to over 97 %. Similarly, Collin-Dufresne and Goldstein (2002) show that there is a very weak correlation between changes in swap rates and returns on at-the-money (ATM) cap straddles: the R^2s of regressions of straddle returns on changes of swap rates are typically less than 20 %. Furthermore, one principal component explains 80 % of regression residuals of straddles with different maturities. As straddles are approximately delta neutral and mainly exposed to volatility risk, they refer to the factor that drives straddle returns but is not affected by the term structure factors as "unspanned stochastic volatility" (hereafter USV). Jagannathan et al. (2003) find that an affine three-factor model can fit the LIBOR/swap curve rather well. However, they identify significant shortcomings when confronting the model with data on caps and swaptions, thus concluding that derivatives must be used when evaluating term structure models. On the other hand, Fan et al. (2003) provide evidence against the existence of USV and show the swaptions can be hedged using bonds alone with an HJM model and the difference from the previous studies results from the nonlinear dependence of derivative prices on the yield curve factors. Li and Zhao (2006) show the yield curve factors extracted using a quadratic term structure model can hedge the bonds perfectly, but not the interest rate caps, and the unhedged components can systematically improve hedging performance across moneyness. They argue the difference is likely due to the fact that the interest caps are more sensitive to the volatility factors than the swaptions, and also the DTSMs are suitable to address the question whether the derivatives are redundant since the HJM type of models needs using both data sets for estimation. Overall, most studies suggest that interest rate derivatives are not redundant securities and cannot be hedged using bonds alone. In other words, bonds do not span interest rate derivatives. In the following table we regress weekly cap straddle returns at different moneyness and maturity on weekly changes in the three yield factors and obtain very similar results. In general, the R^2s in Table 86.1 are very small for straddles that are close to the money. For deep ITM and OTM straddles, the R^2s

[1]Another issue is the relative pricing between caps and swaptions. Although both caps and swaptions are derivatives on LIBOR rates, existing models calibrated to one set of prices tend to significantly misprice the other set of prices. For a more detailed review of the literature, see Dai and Singleton (2003).

Table 86.1 Regression analysis of USV in cap market

	Maturity												
Moneyness (K/F)	1.5	2	2.5	3	3.5	4	4.5	5	6	7	8	9	10
0.60	–	–	–	–	0.720	0.696	0.285	–	0.672	0.332	0.477	0.238	0.258
0.65	–	–	0.816	0.754	0.725	0.719	0.293	0.245	0.636	0.284	0.210	0.134	0.149
0.70	–	–	0.769	0.690	0.615	**0.521**	**0.258**	0.142	0.404	0.173	0.184	0.102	0.080
0.75	–	0.704	**0.677**	**0.654**	**0.596**	**0.491**	0.257	0.065	0.310	0.198	0.149	0.087	0.080
0.80	–	0.557	**0.624**	**0.577**	**0.418**	**0.481**	**0.174**	**0.090**	**0.225**	**0.128**	**0.112**	0.064	0.073
0.85	0.709	**0.400**	**0.538**	**0.519**	**0.329**	**0.319**	**0.133**	**0.093**	**0.164**	**0.088**	**0.068**	**0.041**	0.043
0.90	0.507	**0.242**	**0.364**	**0.322**	**0.208**	**0.223**	**0.078**	**0.060**	**0.094**	**0.056**	**0.035**	**0.027**	**0.066**
0.95	**0.268**	**0.076**	**0.256**	**0.223**	**0.124**	**0.108**	**0.041**	**0.031**	**0.042**	**0.026**	**0.031**	**0.024**	**0.053**
1.00	**0.300**	**0.061**	**0.209**	**0.160**	**0.083**	**0.031**	**0.044**	**0.021**	**0.025**	**0.014**	**0.046**	**0.024**	**0.035**
1.05	**0.473**	**0.104**	**0.249**	**0.169**	**0.137**	**0.063**	**0.076**	**0.024**	**0.044**	0.025	**0.086**	0.036	0.043
1.10	**0.567**	**0.218**	**0.366**	**0.256**	**0.263**	**0.171**	**0.147**	0.065	0.101	0.057	0.140	0.044	0.024
1.15	**0.667**	**0.374**	**0.497**	**0.390**	**0.423**	**0.306**	0.238	0.100	0.192	0.119	0.207	–	–
1.20	0.751	**0.543**	**0.608**	**0.510**	**0.552**	0.476	0.302	0.167	0.251	0.209	0.228	–	–
1.25	0.801	**0.658**	**0.691**	**0.613**	0.670	0.593	0.463	0.240	0.307	0.219	–	–	–
1.30	0.842	**0.739**	**0.775**	0.708	0.737	0.660	0.502	0.249	0.390	–	–	–	–
1.35	0.872	**0.784**	0.827	0.785	0.796	0.756	0.553	–	–	–	–	–	–
1.40	0.888	0.821	0.879	0.832	0.845	0.851	–	–	–	–	–	–	–

This table reports the R^2s of regressions of weekly returns of cap straddles across moneyness and maturity on weekly changes of the three yield factors. Due to changes in interest rates and strike prices, we do not have the same number of observations for each moneyness/maturity group. The bold entries represent moneyness/maturity groups that have less than 10 % of missing values, and the rest are the ones with 10–50 % of missing values increase significantly. This is consistent with the fact that the ATM straddles are more sensitive to the volatility risk than others away from money.

The existence of USV has profound implications for term structure modeling, especially on the existing multifactor dynamic term structure models, a widely popular term structure model followed by a huge literature in the last decade. One of the main reasons of the popularity of these models is their tractability: they provide closed-form solutions for the prices of not only zero-coupon bonds but also a wide range of interest rate derivatives (see, e.g., Duffie et al. 2000; Chacko and Das 2002; Leippold and Wu 2002). The closed-form formulae significantly reduce the computational burden of implementing these models and simplify their applications in practice. However, almost all existing DTSMs assume that derivatives are redundant and can be perfectly hedged using solely bonds. Hence, the presence of USV in the derivatives market implies that one fundamental assumption underlying all DTSMs does not hold, and these models need to be substantially extended to incorporate the unspanned factors before they can be applied to derivatives. However, as Collin-Dufresne and Goldstein (2002) show, it is rather difficult to introduce USV in traditional DTSMs: one must impose highly restrictive assumptions on model parameters to guarantee that certain factors that affect derivative prices do not affect bond prices. In other words, the ATSMs with USV are restricted version of the existing ATSMs. Some recent papers, for example, Bikbov and Chernov (2004), have tested the USV restrictions by comparing USV models to the nesting unrestricted ATSMs and rejected the USV restrictions when both models are fitted to both bonds and derivatives data. This approach, however, gives misleading conclusions. The USV for term structure models resembles the inclusion of stochastic volatility (SV) in the stock price process, where the natural comparison is between the Black-Scholes model and the SV model. Similarly, the USV model should be nesting the traditional DTSM without USV for statistical testing. Specifically, if the unrestricted three-factor affine model is a good fit for the term structure of interest rates, one should test whether adding one more factor, i.e., a four-factor affine model with USV, will help capture the derivatives data.

Some recent studies have also provided evidence in support of the existence of USV using bonds data alone. They show the yield curve volatilities backed out from a cross section of bond yields do not agree with the time-series filtered volatilities, via GARCH or high-frequency estimates from yields data. This challenges the traditional DTSMs even more since these models cannot be expected to capture the option implied volatilities if they cannot even match the realized yield curve volatilities. Specifically, Collin-Dufresne et al. (2004, CDGJ) show that the LIBOR volatility implied by an affine multifactor specification from the swap rate curve can be negatively correlated with the time series of volatility obtained from a standard GARCH approach. In response, they argue that an affine four-factor USV model delivers both realistic volatility estimates and a good cross-sectional fit. Andersen and Benzoni (2006), through the use of high-frequency data on bond yields, construct the model-free "realized yield volatility" measure by computing empirical quadratic yield variation for a cross section of fixed maturities. They find that the yield curve fails to span yield volatility, as the systematic volatility factors are largely unrelated to the cross section of yields. They claim that a broad class of affine diffusive, Gaussian-quadratic, and affine jump-diffusive models is incapable of accommodating the observed yield volatility dynamics. An important implication is that the bond markets per se are incomplete and yield volatility risk cannot be hedged by taking positions solely in the Treasury bond market. They also advocate using the empirical realized yield volatility measures more broadly as a basis for specification testing and (parametric) model selection within the term structure literature. Thompson (2008), on the LIBOR/swap data, argues when the affine models are estimated with the time-series filtered yield volatility, they can pass on his newly proposed specification test, but not with the cross-sectional backed-out volatility. From these studies on the yields data alone, there may exist an alternative explanation for the failure of DTSMs in effectively pricing derivatives in that the bonds small convexity makes bonds not sensitive enough to identify the volatilities from measurement errors. Therefore, efficient inference requires derivatives data as well.

It can be argued in the same fashion that identification of the unspanned factors can be most efficiently accomplished by adding derivatives data to the analysis. Duarte (2008) shows mortgage-backed security (MBS) hedging activity affects interest rate volatility and proposes a model that takes these effects as a measure for the stochastic volatility of the underlying term structure. However, it is unclear whether the realized volatility is indeed different from the implied volatility due to the MBS effects.

Li and Zhao (2009) provide one of the first nonparametric estimates of probability densities of LIBOR rates under forward martingale measures using caps with a wide range of strike prices and maturities.[2] The nonparametric estimates of LIBOR forward densities are conditional on the slope and volatility factors of LIBOR rates, while the level factor is automatically incorporated in existing methods.[3] They find that the forward densities depend significantly on the slope and volatility of LIBOR rates. For example, the forward densities become more dispersed (compact) when the slope of the term structure (the volatility of LIBOR rates) increases. Further analysis reveals a nonlinear relation between the forward densities and the volatility of LIBOR rates that depends on the slope of the term structure. With a flat (steep) term structure, higher volatility tends to lead to more dispersed (compact) forward densities. This result suggests that the speed of mean reversion of the volatility process depends on the slope of the term structure, a feature that has not been explicitly

[2]The nonparametric forward densities estimated using caps, which are among the simplest and most liquid OTC interest rate derivatives, allow consistent pricing of more exotic and/or less liquid OTC interest rate derivatives based on the forward measure approach. The nonparametric forward densities can reveal potential misspecifications of most existing term structure models, which rely on strong parametric assumptions to obtain closed-form formula for interest rate derivative prices.

[3]Andersen and Benzoni (2006) show the “curvature” factor is not significantly correlated with the yield volatility and it is true in this paper as well; therefore, the volatility effect here is not due to the “curvature” factor.

accounted for by existing term structure models. Additionally, this paper documents important impacts of mortgage market activities on the LIBOR forward densities even after controlling for both the slope and volatility factors. For example, the forward densities at intermediate maturities (3, 4, and 5 years) are more negatively skewed when refinance activities, measured by the Mortgage Bankers Association of America (MBAA) refinance index, are high. Demands for out-of-the-money (OTM) floors by investors in mortgage-backed securities (MBS) to hedge potential losses from prepayments could lead to more negatively skewed forward densities. The impacts of refinance activities are most significant at intermediate maturities because the durations of most MBS are around 5 years. The forward density at 2-year maturity is more rightly skewed when ARMs origination (measured by the MBAA ARMs index) is high. Since every ARM contains an interest rate cap that caps the mortgage rate at a certain level, demands for OTM caps from ARMs lenders to hedge their exposures to rising interest rate could lead to more rightly skewed forward densities. The impacts of ARMs are most significant at 2-year maturity because most ARMs get reset within the first 2 years. These empirical results have important implications for the unspanned stochastic volatility puzzle by providing nonparametric and model-independent evidence of USV. The impacts of mortgage activities on the forward densities further suggest that the unspanned factors could be partially driven by activities in the mortgage markets.

The next question naturally is how to best incorporate USV into a term structure model so it can price wide spectrum of interest rate derivatives effectively. In contrast to the approach of adding USV restrictions to DTSMs, it is relatively easy to introduce USV in the Heath et al. (1992) (hereafter, HJM) class of models, which include the LIBOR models of Brace et al. (1997) and Miltersen et al. (1997), the random field models of Goldstein (2000), and the string models of Santa-Clara and Sornette (2001). Indeed, any HJM model in which the forward rate curve has stochastic volatility and the volatility and yield shocks are not perfectly correlated exhibits USV. Therefore, in addition to the commonly known advantages of HJM models (such as perfectly fitting the initial yield curve), they offer the additional advantage of easily accommodating USV. Of course, the trade-off here is that in an HJM model, the yield curve is an input rather than a prediction of the model.

Recently, several HJM models with USV have been developed and applied to price caps and swaptions. Collin-Dufresne and Goldstein (2003) develop a random field model with stochastic volatility and correlation in forward rates. Applying the transform analysis of Duffie et al. (2000), they obtain closed-form formulae for a wide variety of interest rate derivatives. However, they do not calibrate their models to market prices of caps and swaptions. Han (2007) extends the model of LSS (2001) by introducing stochastic volatility and correlation in forward rates. Han (2007) shows that stochastic volatility and correlation are important for reconciling the mispricing between caps and swaptions. Trolle and Schwartz (2009) develop a multifactor term structure model with unspanned stochastic volatility factors and correlation between innovations to forward rates and their volatilities.

Jarrow et al. (2007) develop a multifactor HJM model with stochastic volatility and jumps in LIBOR forward rates. The LIBOR rates follow the affine jump diffusions (hereafter, AJDs) of Duffie et al. (2000), and a closed-form solution for cap prices is provided. Given a small number of factors can explain most of the variation of bond yields, they consider low-dimensional model specifications based on the first few (up to three) principal components of historical forward rates. Their model explicitly incorporates jumps in LIBOR rates, making it possible to differentiate the importance between stochastic volatility and jumps for pricing interest rate derivatives.

In section two of this review, we will discuss how the original DTSMs have difficulty in pricing and hedging interest rate derivatives, as shown in Li and Zhao (2006). In section three, we present the HJM model as in Jarrow et al. (2007). Finally, we will provide nonparametric evidence from Li and Zhao (2009) showing both the realized and implied yield volatilities cannot be spanned by the yield curve factors.

86.2 Term Structure Models with Spanned Stochastic Volatility

We begin with a two-factor spot rate model with stochastic volatility as in Longstaff and Schwartz (1992). Under the risk-neutral measure Q, the short rate r and its volatility V follow a two-dimensional square-root process:

$$dr_t = \kappa_r(\theta_r - r_t)dt + \sqrt{V_t}dW_{1t}^Q; \tag{86.1}$$

$$dV_t = \kappa_V(\theta_V - V_t)dt + \sigma\sqrt{V_t}dW_{2t}^Q. \tag{86.2}$$

The price of the zero-coupon bond $P(t, T)$ with maturity T can be solved through the fundamental PDE for bond pricing:

$$\frac{1}{2}V\left(P_{rr} + \sigma^2 Pvv\right) + P_r\kappa_r(\theta_r - r) + P_v\kappa_v(\theta_v - V) + P_t = rP \tag{86.3}$$

for $0 \le t \le T$ with the terminal condition $P(T, T) = 1$. The price $P(t, T)$ and yield $y(t, T)$ are functions of the state variables $\{r_t, V_t\}$:

$$P(t,T) = e^{A(T-t)+B(T-t)rt+C(T-t)V_t}, \tag{86.4}$$

$$y(t,T) = -\frac{\log(P(t,T))}{T-t} = -\frac{A(T-t)}{T-t} - \frac{B(T-t)}{T-t}r_t - \frac{C(T-t)}{T-t}V_t, \tag{86.5}$$

where A, B, and C are coefficients depending on maturity. It is clear here that the state variables are linear combinations of the yield curve factors such as level and slope. In this sense, the stochastic volatility is spanned by the bond yields. Both bonds and bond derivatives can be priced through the fundamental PDE and the same set of state variables enters into their prices. We should note here that the volatility process in this model serves two roles. First, it helps to price the cross section of bonds, making the model more flexible than one-factor models in generating various shapes of the yield curve. Second, it is the volatility process of the short rate; therefore, it can be inferred using the time series of the short rate alone. One essential question to address therefore is whether the volatility process inferred cross-sectionally fits the time-series properties stipulated by the model. The potential failure in the fit can be due to the model misspecification or the fact that the volatility process cannot be identified using the yield curve factors. For illustration, we discuss the example given in Casassus et al. (2005):

$$dr_t = \kappa_r(\theta_t - r_t)dt + \sqrt{V_t}dW_{1t}^{Q}, \tag{86.6}$$

$$d\theta_t = \kappa_r\left(\gamma_\theta(t) - 2\kappa_r\theta_t + \frac{V_t}{\kappa_r}\right)dt, \tag{86.7}$$

$$dV_t = \mu_V(V_t, t)dt + \sigma(V_t, t)dW_{2t}^{Q}, \tag{86.8}$$

where the long-run mean of the short rate θ_t has a pure drift process and the short rate volatility V_t follows a stochastic process with general drift and diffusion functions that only depend on the volatility itself. It can be shown that the zero-coupon bond price $P(t, T)$ depends only on the short rate and its long-run mean, not on the volatility, i.e.,

$$P(t,T) = e^{A(t,T)+B(T-t)r_t+C(T-t)\theta_t}. \tag{86.9}$$

It can also be shown that the price of a European call option on the zero-coupon bond, however, depends on the volatility V_t. For this example, the call option cannot be hedged by using bonds alone.

Therefore, it is an important exercise to test whether a sophisticated DTSM without the USV factor can be used to hedge the interest derivatives. Dai and Singleton (2003) review many of the current dynamic term structure models, and these models include the affine term structure models (ATSMs) of Duffie and Kan (1996) and the QTSMs of ADG (2002) and many others.[4]

In a typical dynamic term structure model, the economy is represented by the filtered probability space $(\Omega, \mathcal{F}, \{\mathcal{F}_t\}_{0\le t\le T}, P)$, where $\{\mathcal{F}_t\}_{0\le t\le T}$ is the augmented filtration generated by an N-dimensional standard Brownian motion, W, on this probability space. It is usually assumed that $\{\mathcal{F}_t\}_{0\le t\le T}$ satisfies the usual hypothesis (see Protter 1990).

[4]The affine models include the completely affine models of Dai and Singleton (2000), the essentially affine models of Duffee (2002), and the semi-affine models of Duarte (2004). Other DTSMs include the hybrid models of Ahn et al. (2003), the regime-switching models of Bansal and Zhou (2003), and models with macroeconomic jump effects, such as Piazzesi (2001), and many others.

The ATSMs rely on the following assumptions:

- The instantaneous interest rate r_t is an affine or quadratic function of the N-dimensional state variables X_t,

$$r(X_t) = \beta' X_t + \alpha, \tag{86.10}$$

- The state variables follow a multivariate affine process:

$$dX_t = K[\theta - X_t]dt + \Sigma S_t dW_t, \tag{86.11}$$

where S_t is a diagonal matrix with elements being the square root of an affine function of X_t. Hence, the conditional means and variances of the state variables are affine functions of the state variables.

- The market price of risk is a function of the state variables:

$$\zeta(X_t) = \eta_0 X_t S_t^- + \eta_1 S_t^-, \tag{86.12}$$

where S_t^- is a diagonal matrix with elements being the inverse of those in S_t wherever positive zero otherwise.

The zero-coupon bond with time to maturity τ can be priced by risk-neutral pricing:

$$\begin{aligned} D(t,\tau) &= E_t^Q\left[e - \int_t^{t+\tau} r(X_e)^{ds}.1\right] \\ &= e^{-A(\tau)-B(\tau)'Xt}. \end{aligned} \tag{86.13}$$

The functions of $A(\tau)$ and $B(\tau)$ satisfy a system of ordinary differential equations. The continuously compounding yield $y(t, \tau)$ follows

$$y(t,\tau) = \frac{1}{\tau}\left[A(\tau) + B(\tau)' X_t\right]. \tag{86.14}$$

The interest rate derivatives can be priced similarly via risk-neutral pricing. Without any restrictions on the model parameters, the loadings for the state variables, $B(\tau)$, are not zero in general. Hence, the state variables can always be backed out given enough number of yields, leaving the derivative prices been redundant in identifying the state variables.

The QTSMs rely on the following assumptions:

- The instantaneous interest rate r_t is an affine or quadratic function of the N-dimensional state variables X_t:

$$r(X_t) = X_t' \Psi X_t + \beta' X_t + \alpha \tag{86.15}$$

- The state variables follow a multivariate Gaussian process:

$$dX_t = [\mu + \xi X_t]dt + \Sigma dW_t, \tag{86.16}$$

- The market price of risk is an affine function of the state variables:

$$\varsigma(X_t) = \eta_0 + \eta_1 X_t. \tag{86.17}$$

Note that in the above equations ψ, ξ, $\sum$, and η_1 are N-by-N matrices; β, μ, and η_0 are vectors of length N; and α is a scalar. The quadratic relation between r_t and X_t has the desired property that r_t is guaranteed to be positive if ψ is positive semidefinite and $\alpha - \frac{1}{4}\beta'\Psi\beta \geq 0$. Although X_t follows a Gaussian process in Eq. 86.2, interest rate r_t exhibits conditional heteroskedasticity because of the quadratic relationship between r_t and X_t. As a result, the QTSMs are more flexible in modeling volatility clustering in bond yields and correlations among the state variables than the ATSMs.

Consequently, the yield to maturity, $y(t, \tau)$, is a quadratic function of the state variables:

$$y(t,\tau) = \frac{1}{\tau}\left[X_t' A(\tau) X_t + b(\tau)' X_t + c(\tau)\right]. \tag{86.18}$$

In contrast, in the ATSMs the yields are linear in the state variables, and therefore the correlations among the yields are solely determined by the correlations of the state variables. Although the state variables in the QTSMs follow multivariate Gaussian process, the quadratic form of the yields helps to model the time-varying volatility and correlation of bond yields. Leippold and Wu (2002) show that a large class of fixed-income securities can be priced in closed form in the QTSMs using the transform analysis of Duffie et al. (2000). The details of the derivation are in the appendix.

The first test for these models is to capture both the cross-sectional and time-series properties of bond yields, which has been reviewed in Dai and Singleton (2003). Even though the most sophisticated models can fit the cross section of bond prices very well and they can capture the time-series property of the first moment of the yield curve factors, they do not perform satisfactorily in capturing the second moment. The second test is to see whether these models can be used to price and hedge a cross section of interest rate derivatives. The task to performing the second task is made somewhat easier due to one major advantage of these DTSMs in that they provide closed-form solutions for a wide range of interest rate derivatives.

The empirical results shown below are from Li and Zhao (2006), in which they study the performance of QTSMs in pricing and hedging interest rate caps. Even though the study is based on QTSMs, the empirical findings are common to ATSMs as well.[5]

[5]In the empirical analysis of Li and Zhao (2006), the QTSMs are chosen for several reasons. First, since the nominal spot interest rate is a quadratic function of the state variables, it is guaranteed to be positive in the QTSMs. On the other hand, in the ATSMs, the spot rate, an affine function of the state variables, is guaranteed to be positive only when all the state variables follow square-root processes. Second, the QTSMs do not have the limitations facing the ATSMs in simultaneously fitting interest rate volatility and correlations among the state variables. That is, in the ATSMs, increasing the number of factors that follow square-root processes improves the modeling of volatility clustering in bond yields, but reduces the flexibility in modeling correlations among the state variables. Third, the QTSMs have the potential to capture observed nonlinearity in term structure data (see, e.g., Ahn and Gao 1999). Indeed, ADG (2002) and Brandt and Chapman (2002) show that the QTSMs can capture the conditional volatility of bond yields better than the ATSMs.

To price and hedge caps in the QTSMs, both model parameters and latent state variables need to be estimated. Due to the quadratic relationship between bond yields and the state variables, the state variables are not identified by the observed yields even in the univariate case in the QTSMs. Previous studies such as ADG (2002) have used the efficient method of moments (EMM) of Gallant and Tauchen (1998) to estimate the QTSMs. Li and Zhao (2006) use the extended Kalman filter (EKF) to estimate model parameters and extract the latent state variables in one step. The details of the implementation of the EKF are in the appendix.

The pricing analysis can reveal two sources of potential model misspecification. One is on the number of factors in the model as a missing factor usually causes large pricing errors. An analogy is using Black-Scholes model, while the stock price is generated from a stochastic volatility model. The other is on the assumption of the innovation process of each factor. If the innovation of the factor has a fat-tailed distribution, the convenient assumption of Gaussian distribution is going to deliver large pricing error as well. So from a pricing study, we cannot conclude one or the other or both cause large pricing errors. On the other hand, hedging analysis focuses on the changes of the prices, so even if the marginal distribution of the prices can be highly non-Gaussian, the conditional distribution for a small time step can still be reasonably approximated with Gaussian distribution. As the result, a deficiency in hedging, especially at high frequency, reveals more about the potential missing factors than the distribution assumption in a model.

In Li and Zhao (2006), QTSMs can capture yield curve dynamics extremely well. First, given the estimated model parameters and state variables, they compute the 1-day-ahead projection of yields based on the estimated model. Figure 86.1 shows that QTSM1 model projected yields are almost indistinguishable from the corresponding observed yields. Secondly, they examine the performance of the QTSMs in hedging zero-coupon bonds, assuming that the filtered state variables are traded and use them as hedging instruments. The delta-neutral hedge is conducted for zero-coupon bonds of six maturities on a daily basis. Hedging performance is measured by variance ratio, which is defined as the percentage of the variations of an unhedged position that can be reduced by hedging. The results on the hedging performance in Table 86.2 show that in most cases the variance ratios are higher than 95 %. This should not be surprising given the excellent fit of bond yields by the QTSMs.

If the LIBOR and swap market and the cap market are well integrated, the estimated three-factor QTSMs should be able to hedge caps well. Based on the estimated model parameters, the delta-neutral hedge of weekly changes of difference cap prices is conducted using filtered state variables as hedging instruments. It is also possible to use LIBOR zero-coupon bonds as hedging instruments by matching the hedge ratios of a difference cap with that of zero-coupon bonds. Daily rebalance – adjustment of the hedge ratios everyday given changes in market conditions – is implemented to improve hedging performance. Therefore, daily changes of a hedged position are the difference between daily changes of the unhedged position and the hedging portfolio. The latter equals to the sum of the products of a difference cap's hedge ratios with respect to the state variables and changes in the corresponding state variables. Weekly changes are just the accumulation over daily positions. The hedging

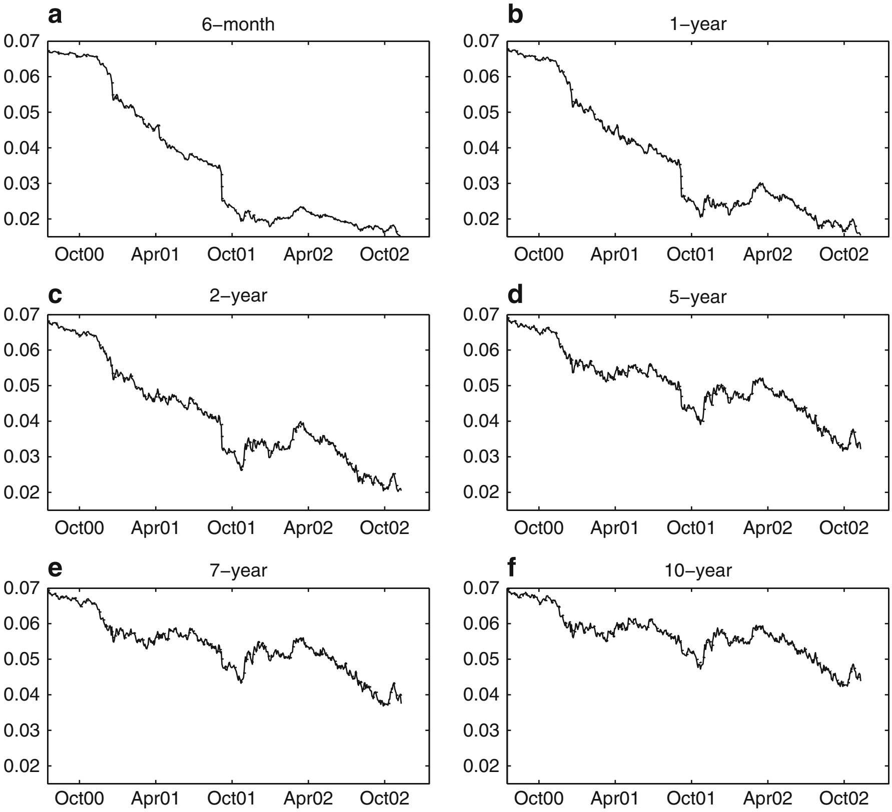


Fig. 86.1 The observed yields (*dot*) and the QTSM1 projected yields (*solid*)

effectiveness is measured by variance ratio, the percentage of the variations of an unhedged position that can be reduced by hedging. This measure is similar in spirit to R^2 in linear regression. The variance ratios of the three QTSMs in Table 86.3 show that all models have better hedging performance for ITM, short-term (maturities from 1.5 to 4 years) difference caps[6] than OTM, medium- and long-term difference caps (maturities longer than 4 years). There is a high percentage of variations in long-term and OTM difference cap prices that cannot be hedged. The maximal flexible model QTSM1 again has better hedging performance than the other two models. To control for the fact that the QTSMs may be misspecified, in Panel B of Table 86.3, the hedging errors of each moneyness/maturity group are further regressed on the changes of the three yield factors. While the three yield factors can explain some additional hedging errors, their incremental explanatory power is not very significant. Thus, even

[6]The difference cap is the difference of the caps with subsequent maturities and the same strike prices. Instead of having caplets ranging from as early as 6 months, the difference cap only has caplets of a small maturity region.

Table 86.2 The performance of QTSMs in modeling bond yields

	Maturity (year)					
	0.5	1	2	5	7	10
QTSM3	0.717	0.948	0.982	0.98	0.993	0.93
QTSM2	0.99	0.956	0.963	0.975	0.997	0.934
QTSM1	0.994	0.962	0.969	0.976	0.997	0.932

This table reports the performance of the three-factor QTSMs in capturing bond yields. Variance ratios of model-based hedging of zero-coupon bonds in QTSMs using filtered state variables as hedging instruments. Variance ratio measures the percentage of the variations of an unhedged position that can be reduced through hedging

excluding hedging errors that can be captured by the three yield factors, there is still a large fraction of difference cap prices that cannot be explained by the QTSMs. Table 86.4 reports the performance of the QTSMs in hedging cap straddles. The difference floor prices are computed from difference cap prices using the put-call parity and construct weekly straddle returns. As straddles are highly sensitive to volatility risk, both delta- and gamma-neutral hedges are needed. The variance ratios of QTSM1 are as low as the R^2s of linear regressions of straddle returns on the yield factors in Table 86.1, suggesting that neither approach can explain much variations of straddle returns. Collin-Dufresne and Goldstein (2002) show that 80 % of straddle regression residuals can be explained by one additional factor. Principal component analysis of ATM straddle hedging errors in Panel B of Table 86.4 shows that the first factor can explain about 60 % of the total variations of hedging errors. The second and third factor each explains about 10 % of hedging errors, and two additional factors combined can explain about another 10 % of hedging errors. The correlation matrix of the ATM straddle hedging errors across maturities in Panel C shows that the hedging errors of short-term (2, 2.5, 3, 3.5, and 4 years), medium-term (4.5 and 5 years), and long-term (8, 9, and 10 years) straddles are highly correlated within each group, suggesting that there could be multiple unspanned factors.

To further understand whether the unspanned factors are related to stochastic volatility, we study the relationship between ATM cap implied volatilities and straddle hedging errors. Principal component analysis in Panel A of Table 86.5 shows that the first component explains 85 % of the variations of cap implied volatilities. In Panel B, we regress straddle hedging errors on changes of the three yield factors and obtain R^2s that are close to zero. However, if we include the weekly changes of the first few principal components of cap implied volatilities, the R^2s increase significantly: for some maturities, the R^2s are above 90 %. Although the time series of implied volatilities are very persistent, their differences are not and we do not suffer from the well-known problem of spurious regression. In the extreme case in which we regress straddle hedging errors of each maturity on changes of the yield factors and cap implied volatilities with the same maturity, the R^2s in most cases are above 90 %. These results show that straddle returns are mainly affected by volatility risk but not term structure factors.

Table 86.3 The performance of QTSMs in hedging interest rate caps

	Maturity												
Moneyness (K/F)	1.5	2	2.5	3	3.5	4	4.5	5	6	7	8	9	10
Panel A. Variance ratio of QTSM1													
0.60	–	–	–	–	0.917	–	–	–	0.862	0.679	0.665	0.494	0.257
0.65	–	–	0.904	0.919	0.920	–	0.671	0.549	0.861	0.704	0.609	0.431	0.255
0.70	–	–	0.903	0.913	0.916	**0.862**	**0.666**	0.487	0.822	0.619	0.565	0.355	0.218
0.75	–	0.865	**0.884**	**0.911**	**0.902**	**0.852**	0.689	0.447	0.807	0.620	0.544	0.326	0.198
0.80	–	0.831	**0.890**	**0.900**	**0.876**	**0.864**	**0.670**	**0.504**	**0.785**	**0.594**	**0.537**	0.305	0.185
0.85	0.894	**0.818**	**0.880**	**0.893**	**0.869**	**0.833**	**0.649**	**0.531**	**0.773**	**0.590**	**0.516**	**0.296**	0.159
0.90	0.890	**0.810**	**0.853**	**0.872**	**0.851**	**0.832**	**0.631**	**0.514**	**0.748**	**0.577**	**0.491**	**0.314**	**0.171**
0.95	**0.888**	**0.779**	**0.832**	**0.855**	**0.847**	**0.833**	**0.596**	**0.481**	**0.716**	**0.578**	**0.481**	**0.303**	**0.182**
1.00	**0.875**	**0.677**	**0.803**	**0.824**	**0.826**	**0.815**	**0.575**	**0.456**	**0.695**	**0.533**	**0.476**	**0.287**	**0.164**
1.05	**0.856**	**0.619**	**0.767**	**0.799**	**0.797**	**0.805**	**0.536**	**0.424**	**0.671**	0.512	**0.492**	0.245	0.138
1.10	**0.851**	**0.575**	**0.737**	**0.779**	**0.763**	**0.773**	**0.523**	0.411	0.623	0.490	0.415	0.204	–
1.15	**0.789**	**0.529**	**0.692**	**0.755**	**0.724**	**0.722**	0.483	0.422	0.611	0.426	–	–	–
1.20	0.756	**0.489**	**0.645**	**0.692**	**0.654**	0.673	0.521	0.470	0.533	0.415	–	–	–
1.25	0.733	**0.438**	**0.603**	**0.645**	0.575	0.634	0.587	0.551	0.514	–	–	–	–
1.30	0.724	**0.393**	**0.534**	0.591	0.444	0.602	0.540	–	0.334	–	–	–	–
1.35	0.691	**0.324**	0.449	0.539	0.408	0.515	0.436	–	–	–	–	–	–
1.40	–	0.260	0.373	0.464	0.319	–	–	–	–	–	–	–	–

(continued)

Table 86.3 (continued)

	Maturity												
Moneyness (K/F)	1.5	2	2.5	3	3.5	4	4.5	5	6	7	8	9	10
Panel B. Variance ratio of QTSM1 combined with the changes of the three yield factors													
0.60	–	–	–	–	0.921	–	–	–	0.912	0.788	0.815	0.658	0.579
0.65	–	–	0.929	0.928	0.921	–	0.676	0.573	0.904	0.804	0.765	0.611	0.471
0.70	–	–	0.927	0.922	0.919	**0.872**	**0.675**	0.507	0.847	0.679	0.666	0.462	0.353
0.75	–	0.914	**0.908**	**0.922**	**0.903**	**0.861**	0.697	0.450	0.835	0.687	0.646	0.429	0.312
0.80	–	0.886	**0.915**	**0.912**	**0.882**	**0.873**	**0.675**	**0.510**	**0.811**	**0.653**	**0.639**	0.417	0.324
0.85	0.951	**0.870**	**0.905**	**0.899**	**0.872**	**0.839**	**0.654**	**0.543**	**0.802**	**0.647**	**0.610**	**0.397**	0.286
0.90	0.942	**0.853**	**0.876**	**0.882**	**0.855**	**0.837**	**0.633**	**0.524**	**0.776**	**0.632**	**0.573**	**0.412**	**0.310**
0.95	**0.935**	**0.825**	**0.860**	**0.864**	**0.853**	**0.837**	**0.597**	**0.488**	**0.738**	**0.623**	**0.554**	**0.395**	**0.306**
1.00	**0.923**	**0.746**	**0.841**	**0.836**	**0.839**	**0.820**	**0.578**	**0.462**	**0.709**	**0.566**	**0.541**	**0.361**	**0.269**
1.05	**0.906**	**0.694**	**0.816**	**0.816**	**0.819**	**0.814**	**0.539**	**0.428**	**0.679**	0.537	**0.545**	0.338	0.210
1.10	**0.895**	**0.659**	**0.799**	**0.804**	**0.794**	**0.786**	**0.530**	0.421	0.630	0.508	0.480	0.278	–
1.15	**0.857**	**0.624**	**0.770**	**0.789**	**0.773**	**0.742**	0.491	0.439	0.623	0.434	–	–	–
1.20	0.848	**0.611**	**0.741**	**0.746**	**0.729**	0.701	0.530	0.486	0.572	0.427	–	–	–
1.25	0.824	**0.577**	**0.712**	**0.716**	0.673	0.668	0.612	0.598	0.541	–	–	–	–
1.30	0.796	**0.560**	**0.680**	0.687	0.626	0.679	0.559	–	0.378	–	–	–	–
1.35	0.777	**0.511**	0.634	0.662	0.603	0.636	0.464	–	–	–	–	–	–
1.40	–	0.455	0.582	0.638	0.573	–	–	–	–	–	–	–	–

This table reports the performance of the three QTSMs in hedging difference caps. Hedging effectiveness is measured by variance ratio, the percentage of the variations of an unhedged position that can be reduced through hedging. The bold entries represent moneyness/maturity groups that have less than 10 % of missing values, and the rest are the ones with 10–50 % of missing values

Table 86.4 Hedging interest rate cap straddles

	Maturity												
(K/F)	1.5	2	2.5	3	3.5	4	4.5	5	6	7	8	9	10
Panel A. The performance of QTSM1 in hedging difference cap straddles measured by variance ratio													
0.60	–	–	–	–	0.711	–	–	–	0.709	0.329	0.596	0.362	0.250
0.65	–	–	0.794	0.756	0.711	–	0.270	0.215	0.672	0.427	0.495	0.283	0.185
0.70	–	–	0.776	0.723	0.674	**0.557**	0.250	0.152	0.473	0.206	0.187	0.096	0.074
0.75	–	0.690	**0.682**	**0.683**	**0.589**	**0.497**	0.254	0.078	0.399	0.211	0.148	0.072	0.063
0.80	–	0.560	**0.652**	**0.615**	**0.437**	**0.488**	**0.179**	**0.093**	**0.293**	**0.126**	**0.113**	0.053	0.070
0.85	0.727	**0.438**	**0.579**	**0.532**	**0.366**	**0.349**	**0.133**	**0.095**	**0.227**	**0.091**	**0.068**	**0.025**	0.037
0.90	0.558	**0.278**	**0.405**	**0.339**	**0.248**	**0.265**	**0.066**	**0.049**	**0.138**	**0.052**	**0.032**	**0.016**	**0.060**
0.95	**0.416**	**0.127**	**0.287**	**0.236**	**0.207**	**0.215**	**0.022**	**0.017**	**0.069**	**0.018**	**0.034**	**0.012**	**0.043**
1.00	**0.364**	**0.081**	**0.210**	**0.142**	**0.149**	**0.142**	**0.024**	**0.006**	**0.045**	**0.006**	**0.047**	**0.009**	**0.002**
1.05	**0.471**	**0.111**	**0.237**	**0.133**	**0.190**	**0.187**	**0.058**	**0.010**	**0.078**	0.035	**0.091**	0.022	0.035
1.10	**0.622**	**0.212**	**0.368**	**0.226**	**0.314**	**0.283**	**0.146**	0.065	0.133	0.054	0.091	0.018	–
1.15	**0.727**	**0.357**	**0.508**	**0.378**	**0.472**	**0.399**	0.235	0.162	0.252	0.107	–	–	–
1.20	0.788	**0.527**	**0.633**	**0.515**	**0.593**	0.481	0.368	0.201	0.343	0.256	–	–	–
1.25	0.831	**0.640**	**0.721**	**0.636**	0.662	0.600	0.464	0.296	0.408	–	–	–	–
1.30	0.851	**0.727**	**0.808**	0.728	0.781	0.729	0.525	–	0.454	–	–	–	–
1.35	0.876	**0.778**	0.852	0.802	0.833	0.804	0.551	–	–	–	–	–	–
1.40	–	0.817	0.894	0.863	0.880	–	–	–	–	–	–	–	–

(continued)

Table 86.4 (continued)

Principal component													
Panel B. Percentage of variance of ATM straddle hedging errors explained by the principal components													
1	2	3	4	5	6								
59.3 %	12.4 %	9.4 %	6.7 %	4.0 %	2.8 %								
	Maturity												
Maturity	1.5	2	2.5	3	3.5	4	4.5	5	6	7	8	9	10
Panel C. Correlation matrix of ATM straddle hedging errors across maturity													
1.5	1.00	–	–	–	–	–	–	–	–	–	–	–	–
2	0.38	1.00	–	–	–	–	–	–	–	–	–	–	–
2.5	0.28	0.66	1.00	–	–	–	–	–	–	–	–	–	–
3	0.03	0.33	0.73	1.00	–	–	–	–	–	–	–	–	–
3.5	0.27	0.52	0.63	0.59	1.00	–	–	–	–	–	–	–	–
4	0.13	0.44	0.37	0.37	0.77	1.00	–	–	–	–	–	–	–
4.5	0.20	0.21	−0.04	−0.08	−0.05	−0.06	1.00	–	–	–	–	–	–
5	0.10	0.11	−0.12	−0.13	−0.16	−0.15	0.96	1.00	–	–	–	–	–
6	0.21	0.16	0.19	0.13	0.25	0.05	0.27	0.23	1.00	–	–	–	–
7	0.30	0.34	0.33	0.35	0.46	0.38	0.28	0.22	0.08	1.00	–	–	–
8	0.10	0.12	0.30	0.30	0.25	0.11	0.36	0.34	0.29	0.29	1.00	–	–
9	0.14	0.11	0.25	0.29	0.26	0.12	0.39	0.37	0.32	0.38	0.83	1.00	–
10	0.08	−0.01	0.17	0.14	0.12	0.01	0.32	0.35	0.26	0.28	0.77	0.86	1.00

Table 86.5 Straddle hedging errors and cap implied volatilities

Principal component					
Panel A. Percentage of variance of ATM cap implied volatilities explained by the principal components					
1	2	3	4	5	6
85.73 %	7.91 %	1.85 %	1.54 %	0.72 %	0.67 %

Maturity												
Panel B. R^2s of the regressions of ATM straddle hedging errors on changes of the three yield factors (row one); changes of the three yield factors and the first four principal components of the ATM cap implied volatilities (row two); and changes of the three yield factors and maturity-wise ATM cap implied volatility (row three)												
1.5	2	2.5	3	3.5	4	4.5	5	6	7	8	9	10
0.10	0.06	0.02	0.01	0.01	0.04	0.00	0.00	0.01	0.01	0.00	0.01	0.04
0.29	0.49	0.54	0.43	0.63	0.47	0.95	0.96	0.21	0.70	0.68	0.89	0.96
0.68	0.70	0.81	0.87	0.85	0.90	0.95	0.98	0.95	0.98	0.97	0.98	0.99

This table reports the relation between straddle hedging errors and ATM cap implied volatilities

As ATM straddles are mainly exposed to volatility risk, their hedging errors can serve as a proxy of the USV. Panels A and B of Table 86.6 report the R^2s of regressions of hedging errors of difference caps and cap straddles across moneyness and maturity on changes of the three yield factors and the first five principal components of straddle hedging errors. In contrast to the regressions in Panel D of Table 86.6, which only include the three yield factors, the additional factors from straddle hedging errors significantly improve the R^2s of the regressions: for most moneyness/maturity groups, the R^2s are above 90 %. Interestingly for long-term caps, the R^2s of ATM and OTM caps are actually higher than that of ITM caps. Therefore, a combination of the yield factors and the USV factors can explain cap prices across moneyness and maturity very well.

While the above analysis is mainly based on the QTSMs, the evidence on USV is so compelling that the results should be robust to potential model misspecification. The fact that the QTSMs provide excellent fit of bond yields but can explain only a small percentage of the variations of ATM straddle returns is a strong indication that the models miss some risk factors that are important for the cap market. While we estimate the QTSMs using only bond prices, we could also include cap prices in model estimation. We do not choose the second approach for several reasons. First, the current approach is consistent with the main objective of our study: use risk factors extracted from the swap market to explain cap prices. Second, it is not clear that modifications of model parameters without changing the fundamental structure of the model could remedy the poor cross-sectional hedging performance of the QTSMs. In fact, if the QTSMs indeed miss some important factors, then no matter how they are estimated (using bonds or derivatives data), they are unlikely to have good hedging performance. Finally, Jagannathan et al. (2003) do not find significant differences between parameters of ATSMs estimated using LIBOR/swap rates and cap/swaption prices. The existence of USV strongly suggests that existing DTSMs need to relax their fundamental assumption that derivatives are redundant securities

Table 86.6 ATM straddle hedging error as a proxy of systematic USV

	Maturity												
Moneyness(K/F)	1.5	2	2.5	3	3.5	4	4.5	5	6	7	8	9	10
Panel A: R^2s of regressions of cap hedging errors													
0.60	–	–	–	–	0.945	–	–	–	0.948	0.880	0.884	0.786	0.880
0.65	–	–	0.938	0.949	0.954	–	0.947	0.952	0.960	0.928	0.871	0.807	0.838
0.70	–	–	0.934	0.944	**0.943**	**0.911**	**0.934**	0.936	0.940	0.885	0.839	0.791	0.776
0.75	–	0.934	**0.926**	**0.945**	**0.943**	**0.910**	0.936	0.919	0.950	0.899	**0.862**	0.814	0.791
0.80	–	0.917	**0.934**	**0.938**	**0.935**	**0.909**	**0.950**	**0.946**	**0.951**	**0.898**	**0.862**	0.821	0.840
0.85	0.958	**0.909**	**0.927**	**0.928**	**0.928**	**0.889**	**0.956**	**0.959**	**0.959**	**0.906**	**0.861**	**0.818**	0.843
0.90	0.949	**0.900**	**0.908**	**0.922**	**0.924**	**0.896**	**0.961**	**0.969**	**0.969**	**0.920**	**0.871**	**0.856**	**0.871**
0.95	**0.943**	**0.886**	**0.905**	**0.918**	**0.936**	**0.906**	**0.966**	**0.976**	**0.980**	**0.967**	**0.889**	**0.882**	**0.893**
1.00	**0.932**	**0.859**	**0.905**	**0.909**	**0.939**	**0.902**	**0.988**	**0.989**	**0.984**	**0.973**	**0.910**	**0.894**	**0.907**
1.05	**0.919**	**0.821**	**0.897**	**0.902**	**0.937**	**0.897**	**0.986**	**0.985**	**0.980**	0.969	**0.908**	0.917	0.885
1.10	**0.913**	**0.793**	**0.890**	**0.894**	**0.928**	**0.880**	**0.979**	0.976	0.974	0.967	0.913	0.921	–
1.15	**0.879**	**0.763**	**0.871**	**0.880**	**0.915**	**0.860**	0.970	0.968	0.966	0.963	–	–	–
1.20	0.881	**0.749**	**0.844**	**0.848**	**0.894**	0.846	0.966	0.963	0.954	0.957	–	–	–
1.25	0.870	**0.742**	**0.818**	**0.817**	0.870	0.819	0.945	0.943	0.941	–	–	–	–
1.30	0.861	**0.702**	**0.802**	0.808	0.836	0.802	0.920	–	0.908	–	–	–	–
1.35	0.855	**0.661**	0.764	0.774	0.801	0.758	0.884	–	–	–	–	–	–
1.40	–	0.640	0.725	0.743	0.761	0.536	–	–	–	–	–	–	–

Panel B: R^2s of regressions of cap straddle hedging errors													
0.60	–	–	–	–	0.851	–	–	–	0.839	0.589	0.724	0.712	0.744
0.65	–	–	0.883	0.916	0.876	–	0.891	0.902	0.860	0.704	0.789	0.768	0.785
0.70	–	–	0.874	0.897	0.825	**0.775**	0.854	0.886	0.800	0.621	0.688	0.749	0.730
0.75	–	0.869	**0.844**	**0.845**	**0.833**	**0.760**	0.850	0.853	0.812	0.655	0.740	0.800	0.761
0.80	–	0.812	**0.839**	**0.828**	**0.800**	**0.734**	**0.872**	**0.886**	**0.751**	**0.607**	**0.740**	0.805	0.805
0.85	0.915	**0.772**	**0.833**	**0.798**	**0.772**	**0.685**	**0.898**	**0.910**	**0.738**	**0.604**	**0.744**	**0.821**	0.807
0.90	0.883	**0.718**	**0.801**	**0.791**	**0.763**	**0.690**	**0.919**	**0.930**	**0.721**	**0.622**	**0.777**	**0.859**	**0.840**
0.95	**0.884**	**0.672**	**0.828**	**0.826**	**0.840**	**0.747**	**0.936**	**0.945**	**0.746**	**0.684**	**0.812**	**0.908**	**0.861**
1.00	**0.901**	**0.745**	**0.861**	**0.822**	**0.851**	**0.745**	**0.971**	**0.976**	**0.738**	**0.703**	**0.849**	**0.922**	**0.894**
1.05	**0.893**	**0.745**	**0.871**	**0.789**	**0.863**	**0.728**	**0.966**	**0.972**	**0.729**	0.758	**0.874**	0.913	0.902
1.10	**0.896**	**0.757**	**0.883**	**0.769**	**0.880**	**0.735**	**0.953**	0.960	0.758	0.788	0.892	0.931	–
1.15	**0.908**	**0.785**	**0.882**	**0.765**	**0.886**	**0.778**	0.945	0.943	0.789	0.821	–	–	–
1.20	0.920	**0.831**	**0.883**	**0.785**	**0.890**	0.808	0.947	0.935	0.809	0.843	–	–	–
1.25	0.944	**0.873**	**0.894**	**0.813**	0.898	0.843	0.918	0.898	0.823	–	–	–	–
1.30	0.945	**0.890**	**0.913**	0.859	0.921	0.872	0.920	–	0.830	–	–	–	–
1.35	0.960	**0.906**	0.927	0.885	0.933	0.901	0.906	–	–	–	–	–	–
1.40	–	0.924	0.941	0.912	0.946	–	–	–	–	–	–	–	–

This table reports the contribution of USV proxied by the first few principal components of ATM straddle hedging errors in explaining the hedging errors of caps and cap straddles across moneyness and maturity. It reports the R^2s of regressions of hedging errors of caps across moneyness and maturity on changes of the three yield factors and the first five principal components of straddle hedging errors. The bold entries represent moneyness/maturity groups that have less than 10 % of missing values, and the rest are the ones with 10–50 % of missing values by explicitly incorporating USV factors. It also suggests that it might be more convenient to consider derivative pricing in the forward rate models of HJM (1992) or the random field models of Goldstein (2000) and Santa-Clara and Sornette (2001) because it is generally very difficult to introduce USV in DTSMs. For example, Collin-Dufresne and Goldstein (2002) show that highly restrictive assumptions on model parameters need to be imposed to guarantee that some state variables that are important for derivative pricing do not affect bond prices. In contrast, they show that it is much easier to introduce USV in the HJM and random field class of models: any HJM or random field model in which the forward rate has a stochastic volatility exhibits USV. While it has always been argued that HJM and random field models are more appropriate for pricing derivatives than DTSMs, the reasoning given here is quite different. That is, in addition to the commonly known advantages of these models (such as they can perfectly fit the initial yield curve while DTSMs generally cannot), another advantage of HJM and random field models is that they can easily accommodate USV (see Collin-Dufresne and Goldstein (2002) for illustration).

The existence of USV suggests that these models may not be directly applicable to derivatives because they all rely on the fundamental assumption that bonds and derivatives are driven by the same set of risk factors. In this paper, we provide probably the first empirical analysis of DTSMs in hedging interest rate derivatives and hope to resolve the controversy on USV through this exercise.

86.3 LIBOR Market Models with Stochastic Volatility and Jumps: Theory and Estimation

In this section, we develop a multifactor HJM model with stochastic volatility and jumps in LIBOR forward rates and discuss model estimation and comparison using a wide cross section of difference caps. Instead of modeling the unobservable instantaneous spot rate or forward rate, we focus on the LIBOR forward rates which are observable and widely used in the market.

86.3.1 Specification of the LIBOR Market Models

Throughout our analysis, we restrict the cap maturity T to a finite set of dates $0 = T_0 < T_1 < \cdots < T_K < T_{K+1}$ and assume that the intervals $T_{k+1} - T_k$ are equally spaced by δ, a quarter of a year. Let $L_k(t) = L(t, T_k)$ be the LIBOR forward rate for the actual period $[T_k, T_{k+1}]$, and similarly let $D_k(t) = D(t, T_k)$ be the price of a zero-coupon bond maturing on T_k: thus, we have

$$L(t, T_k) = \frac{1}{\delta}\left(\frac{D(t, T_k)}{D(t, T_{k+1})} - 1\right), \quad \text{for } k = 1, 2, \ldots K. \tag{86.19}$$

For LIBOR-based instruments, such as caps, floors, and swaptions, it is convenient to consider pricing under the forward measure. Thus, we will focus on the dynamics of the LIBOR forward rates $L_k(t)$ under the forward measure $\mathbb{Q}^{k+1}$, which is essential for pricing caplets maturing at T_{k+1}. Under this measure, the discounted price of any security using $D_{k+1}(t)$ as the numeraire is a martingale. Therefore, the time-t price of a caplet maturing at T_{k+1} with a strike price of X is

$$Caplet(t, T_{k+1}, X) = \delta D_{k+1}(t) E_t^{\mathbb{Q}^{k+1}}\left[(L_k(T_k) - X)^+\right], \tag{86.20}$$

where $E_t^{\mathbb{Q}^{k+1}}$ is taken with respect to $\mathbb{Q}^{k+1}$ given the information set at t. The key to valuation is modeling the evolution of $L_k(t)$ under $\mathbb{Q}^{k+1}$ realistically and yet parsimoniously to yield closed-form pricing formula. To achieve this goal, we rely on the flexible AJDs of Duffie et al. (2000) to model the evolution of LIBOR rates.

We assume that under the physical measure $\mathbb{P}$, the dynamics of LIBOR rates are given by the following system of SDEs, for $t \in [0, T_k)$ and $k = 1, \ldots, K$:

$$\frac{dL_k(t)}{L_k(t)} = \alpha_k(t)dt + \sigma_k(t)dZ_k(t) + dJ_k(t), \tag{86.21}$$

where $\alpha_k(t)$ is an unspecified drift term, $Z_k(t)$ is the k-th element of a K-dimensional correlated Brownian motion with a covariance matrix $\Psi(t)$, and $J_k(t)$ is the k-th element of a K-dimensional independent pure jump process assumed independent of $Z_k(t)$ for all k: to introduce stochastic volatility and correlation, we could allow the volatility of each LIBOR rate $\sigma_k(t)$ and each individual element of $\Psi(t)$ to follow a stochastic process. But such a model is unnecessarily complicated and difficult to implement. Instead, we consider a low-dimensional model based on the first few principal components of historical LIBOR forward rates. We assume that the entire LIBOR forward curve is driven by a small number of factors $N < K (N \leq 3$ in our empirical analysis). By focusing on the first N principal components of historical LIBOR rates, we can reduce the dimension of the model from K to N.

Following LSS (2001) and Han (2007), we assume that the instantaneous covariance matrix of changes in LIBOR rates shares the same eigenvectors as the historical covariance matrix. Suppose that the historical covariance matrix can be approximated as $H = U\Lambda_0 U'$, where Λ_0 is a diagonal matrix whose diagonal elements are the first N largest eigenvalues in descending order, and the N columns of U are the corresponding eigenvectors.[7] Our assumption means that the instantaneous covariance matrix of changes in LIBOR rates with fixed time to maturity, Ω_t, shares the same eigenvectors as H. That is,

$$\Omega_t = U\Lambda_t U', \tag{86.22}$$

[7]We acknowledge that with jumps in LIBOR rates, both the historical and instantaneous covariance matrices of LIBOR rates contain a component that is due to jumps. Our approach implicitly assumes that the first three principal components from the historical covariance matrix capture the variations in LIBOR rates due to continuous shocks and that the impact of jumps is only contained in the residuals.

where Λ_t is a diagonal matrix whose i-th diagonal element, denoted by $V_i(t)$, can be interpreted as the instantaneous variance of the i-th common factor driving the yield curve evolution at t. We assume that $V(t)$ follows the square-root process that has been widely used in the literature for modeling stochastic volatility (see, e.g., Heston 1993):

$$dV_i(t) = \kappa_i(\overline{v}_i - V_i(t))dt + \xi_i\sqrt{V_i(t)}d\widetilde{W}_i(t) \tag{86.23}$$

where $\widetilde{W}_i(t)$ is the i-th element of an N-dimensional independent Brownian motion assumed independent of $Z_k(t)$ and $J_k(t)$ for all k.[8]

While Eqs. 86.4 and 86.5 specify the instantaneous covariance matrix of LIBOR rates with fixed time to maturity, in applications we need the instantaneous covariance matrix of LIBOR rates with fixed maturities $\sum_t$. At $t = 0$, Σ_t coincides with Ω_t; for $t > 0$, we obtain $\sum_t$ from Ω_t through interpolation. Specifically, we assume that $U_{S,j}$ is piecewise constant,[9] i.e., for time to maturity $s \in (T_k, T_{k+1})$:

$$U_s^2 = \frac{1}{2}\left(U_k^2 + U_k^2{+}1\right). \tag{86.24}$$

We further assume that $U_{s,j}$ is constant for all caplets belonging to the same difference cap. For the family of the LIBOR rates with maturities $T = T_1, T_2, \ldots, T_K$, we denote U_{T-t} the time-t matrix that consists of rows of U_{T_k-t}, and therefore we have the time-t covariance matrix of the LIBOR rates with fixed maturities:

$$\Sigma_t = U_{T-t}\Lambda_t U_{T-t}^{'} \tag{86.25}$$

To stay within the family of AJDs, we assume that the random jump times arrive with a constant intensity λ_J, and conditional on the arrival of a jump, the jump size follows a normal distribution $N(\mu_J, \sigma^2{}_J)$. Intuitively, the conditional probability at time t of another jump within the next small time interval Δt is $\lambda_J\Delta t$, and conditional on a jump event, the mean relative jump size is $\mu = \exp\left(\mu_J + \frac{1}{2}\sigma_J^2\right) - 1$.[10] We also assume that the shocks driving LIBOR rates, volatility, and jumps (both jump time and size) are mutually independent from each other.

[8]Many empirical studies on interest rate dynamics (see, e.g., Andersen and Lund 1997; Ball and Torous 1999; Chen and Scott 2001) have shown that correlation between stochastic volatility and interest rates is close to zero. That is, there is not a strong "leverage" effect for interest rates as for stock prices. The independence assumption between stochastic volatility and LIBOR rates in our model captures this stylized fact.

[9]Our interpolation scheme is slightly different from that of Han (2002) for the convenience of deriving closed-form solution for cap prices.

[10]For simplicity, we assume that different forward rates follow the same jump process with constant jump intensity. It is not difficult to allow different jump processes for individual LIBOR rates and the jump intensity to depend on the state of the economy within the AJD framework.

Given the above assumptions, we have the following dynamics of LIBOR rates under the physical measure $\mathbb{P}$:

$$\frac{dL_k(t)}{L_k(t)} = \alpha_k(t)dt + \sum_{j=1}^{N} U_{T_k-t,j}\sqrt{V_j(t)}dW_j(t) + dJ_k(t), \quad k = 1,2,\ldots,K. \tag{86.26}$$

To price caps, we need the dynamics of LIBOR rates under the appropriate forward measure. The existence of stochastic volatility and jumps results in an incomplete market and hence the nonuniqueness of forward martingale measures. Our approach for eliminating this nonuniqueness is to specify the market prices of both the volatility and jump risks to change from the physical measure $\mathbb{P}$ to the forward measure $\mathbb{Q}^{k+1}$.[11] Following the existing literature, we model the volatility risk premium as $\eta_j^{k+1}\sqrt{V_j(t)}$, for $j = 1, \ldots, N$. For the jump risk premium, we assume that under the forward measure $\mathbb{Q}^{k+1}$, the jump process has the same distribution as that under P, except that the jump size follows a normal distribution with mean μ_J^{k+1} and variance σ_J^2. Thus, the mean relative jump size under $\mathbb{Q}^{k+1}$ is $\mu^{k+1} = \exp(\mu_J^{k+1} + \frac{1}{2}\sigma_J^2) - 1$. Our specification of the market prices of jump risks allows the mean relative jump size under $\mathbb{Q}^{k+1}$ to be different from that under $\mathbb{P}$, accommodating a premium for jump size uncertainty. This approach, which is also adopted by Pan (2002), artificially absorbs the risk premium associated with the timing of the jump by the jump size risk premium. In our empirical analysis, we make the simplifying assumption that the volatility and jump risk premiums are linear functions of time to maturity, i.e., $\eta_j^{k+1} = c_{jv}(T_k - 1)$ and $\mu_J^{k+1} = \mu_J + c_J(T_k - 1)$.[12] Due to the no arbitrage restriction, the risk premiums of shocks to LIBOR rates for different forward measures are intimately related to each other. If shocks to volatility and jumps are also correlated with shocks to LIBOR rates, then both volatility and jump risk premiums for different forward measures should also be closely related to each other. However, in our model shocks to LIBOR rates are independent of that to volatility and jumps, and as a result, the change of measure of LIBOR shocks does not affect that of volatility and jump shocks. Due to stochastic volatility and jumps, the underlying LIBOR market is no longer complete, and there is no unique forward measure. This gives us the freedom to choose the functional forms of η_j^{k+1} and μ_J^{k+1}. See Andersen and Brotherton-Ratcliffe (2001) for similar discussions.

[11]The market prices of interest rate risks are defined in such a way that the LIBOR rate is a martingale under the forward measure.

[12]In order to estimate the volatility and jump risk premiums, we need a joint analysis of the dynamics of LIBOR rates under both the physical and forward measures, as in Chernov and Ghysels (2000), Pan (2002), and Eraker (2004). In our empirical analysis, we only focus on the dynamics under the forward measure. Therefore, we can only identify the differences in the risk premiums between forward measures with different maturities. Our specifications of both risk premiums implicitly use the 1-year LIBOR rate as a reference point.

Given the above market prices of risks, we can write down the dynamics of log $(L_k(t))$ under forward measure $\mathbb{Q}^{k+1}$:

$$d\log(L_k(t)) = -\left(\lambda_{J\mu}{}^{k+1} + \frac{1}{2}\sum_{j=1}^{N} U^2_{T_k-t,j} V_j(t)\right)dt + \sum_{j=1}^{N} U_{T_k-t,j}\sqrt{V_j(t)}dW_j^{\mathbb{Q}^{k+1}}(t) + dJ_k^{\mathbb{Q}^{k+1}}(t). \tag{86.27}$$

For pricing purpose, the above process can be further simplified to the following one which has the same distribution:

$$d\log(L_k(t)) = -\left(\lambda_{J\mu}{}^{k+1} + \frac{1}{2}\sum_{j=1}^{N} U^2_{T_k-t,j} V_j(t)\right)dt + \sqrt{\sum_{j=1}^{N} U^2_{T_{k-t,j}} V_j(t)dZ_k^{\mathbb{Q}k+1}(t) + dJ_k^{\mathbb{Q}k+1}(t)}, \tag{86.28}$$

where $Z_k^{\mathbb{Q}^{k+1}}(t)$ is a standard Brownian motion under $\mathbb{Q}^{k+1}$. Now the dynamics of $V_i(t)$ under $\mathbb{Q}^{k+1}$ becomes

$$dV_i(t) = \kappa_i^{k+1}\left(v_i^{-k+1} - V_i(t)\right)dt + \xi_i\sqrt{V_i(t)}d\widetilde{W}_i^{\mathbb{Q}^{k+1}}(t) \tag{86.29}$$

where $\widetilde{W}^{\mathbb{Q}^{Qk+1}}$ is independent of $Z^{\mathbb{Q}^{k+1}}$, $\kappa_j^{k+1} = \kappa_j - \xi_j\eta_j^{k+1}$, and $v_j^{-k+1} = \frac{\kappa_j\overline{v}_j}{\kappa_j - \xi_j\eta_j^{k+1}}$, $j = 1, \ldots N$. The dynamics of $L_k(t)$ under the forward measure $\mathbb{Q}^{k+1}$ are completely captured by Eqs. 86.28 and 86.29.

Given that LIBOR rates follow AJDs under both the physical and forward measures, we can directly apply the transform analysis of Duffie et al. (2000) to derive closed-form formula for cap prices. Denote the state variables at t as $Y_t = (\log(L_k(t)), V_t)'$ and the time-t expectation of $e^{u\ YT}{}_k$ under the forward measure $\mathbb{Q}^{k+1}$ as $\psi(u, Y_t, t, T_k) \triangleq E_t^{\mathbb{Q}^{k+1}}[e^{u \cdot YT}\mathrm{k}]$. Let $u = (u_0, 0_{1xN})'$, and then the time-t expectation of LIBOR rate at T_k equals

$$\begin{aligned} E_t^{\mathbb{Q}^{k+1}}\{\exp[u_0\log(L_k(T_k))]\} &= \psi\left(u_0, Y_t, t, T_k\right) \\ &= \exp\left[a(s) + u_0\log(L_k(t)) + B(s)'V_t\right], \end{aligned} \tag{86.30}$$

where $s = T_k - t$ and closed-form solutions of $a(s)$ and $B(s)$ (an N-by-1 vector) are obtained by solving a system of Riccati equations in the appendix.

Following Duffie et al. (2000), we define

$$G_{a,b}\left(y; Y_t, T_k, \mathbb{Q}^{k+1}\right) = E_t^{\mathbb{Q}^{k+1}}\left[e^{a\cdot\log(L_k(T_k))}1_{\{b\cdot\log(L_k(T_k))\le y\}}\right], \tag{86.31}$$

and its Fourier transform

$$\begin{aligned}\mathcal{G}_{a,b}\left(v;Y_t,T_k,\mathbb{Q}^{k+1}\right) &= \int_R^{e^{ivy}dGa,b(y)} \\ &= E_t^{\mathbb{Q}k+1}\left[e^{(a+ivb)\cdot\log(L_k(T_k))}\right] \\ &= \psi(a+ivb,Y_t,t,T_k).\end{aligned} \tag{86.32}$$

Levy's inversion formula gives

$$G_{a,b}\left(y;Y_t,T_k,\mathbb{Q}^{k+1}\right) = \frac{\psi(a+ivb,Y_t,t,T_k)}{2} - \frac{1}{\pi}\int_0^\infty \frac{\mathrm{Im}[\psi(a+ivb,Y_t,t,T_k)e^{-ivy}]}{v}dv. \tag{86.33}$$

The time-0 price of a caplet that matures at T_{k+1} with a strike price of X equals

$$Caplet(0,T_{k+1},X) = \delta D_{k+1}(0)E_0^{\mathbb{Q}^{k+1}}\left[(L_k(T_k)-X)^+\right], \tag{86.34}$$

where the expectation is given by the inversion formula

$$\begin{aligned}E_0^{\mathbb{Q}^{k+1}}[L_k(T_k)-X]^+ &= G_{1,-1}\left(-\ln X;Y_0,T_k,\mathbb{Q}^{k+1}\right) \\ &\quad - XG_{0,-1}\left(-\ln X;Y_0,T_k,\mathbb{Q}^{k+1}\right).\end{aligned} \tag{86.35}$$

The new models developed in this section nest some of the most important models in the literature, such as LSS (2001) (with constant volatility and no jumps) and Han (2007) (with stochastic volatility and no jumps). The closed-form formula for cap prices makes an empirical implementation of our model very convenient and provides some advantages over existing methods. For example, Han (2007) develops approximations of ATM cap and swaption prices using the techniques of Hull and White (1987). However, such an approach might not work well for away-from-the-money options. In contrast, our method would work well for all options, which is important for explaining the volatility smile.

In addition to introducing stochastic volatility and jumps, our multifactor HJM models also have advantages over the standard LIBOR market models of Brace et al. (1997) and Miltersen et al. (1997) and their extensions often applied to caps in practice.[13] While our models provide a unified multifactor framework to characterize the evolution of the whole yield curve, the LIBOR market models typically make separate specifications of the dynamics of LIBOR rates with different maturities.

[13] Andersen and Brotherton-Ratcliffe (2001) and Glasserman and Kou (2003) develop LIBOR models with stochastic volatility and jumps, respectively.

As suggested by LSS (2001), the standard LIBOR models are "more appropriately viewed as a collection of different univariate models, where the relationship between the underlying factors is left unspecified." In contrast, the dynamics of LIBOR rates with different maturities under their related forward measures are internally consistent with each other given their dynamics under the physical measure and the market prices of risks. Once our models are estimated using one set of prices, they can be used to price and hedge other fixed-income securities.

86.3.2 Estimation Method and Results

We estimate our new market model using prices from a wide cross section of difference caps with different strikes and maturities. Every week we observe prices of difference caps with 10 moneyness and 13 maturities. However, due to changing interest rates, we do not have enough observations in all moneyness/maturity categories throughout the sample. Thus, we focus on the 53 moneyness/maturity categories that have less than 10 % of missing values over the sample estimation period. The moneyness and maturity of all difference caps belong to the following sets {0.7, 0.8, 0.9, 1.0, 1.1} and {1.5, 2.0, 2.5, 3.0, 3.5, 4.0, 4.5, 5.0, 6.0, 7.0, 8.0, 9.0, 10.0} (unit in years), respectively. The difference caps with time to maturity less than or equal to 5 years represent portfolios of two caplets, while those with time to maturity longer than 5 years represent portfolios of four caplets.

We estimate the model parameters by minimizing the sum of squared percentage pricing errors (SSEs) of all relevant difference caps.[14] Consider the time-series observations $t = 1, \ldots, \mathcal{T}$, of the prices of 53 difference caps with moneyness m_i and time to maturities τ_i, $i = 1, \ldots, M = 53$. Let θ represent the model parameters which remain constant over the sample period. Let $C(t, m_i, \tau_i)$ be the observed price of a difference cap with moneyness m_i and time to maturity τ_i and let $\hat{C}(t, \tau_i, m_i, V_t(\theta), \theta)$ denote the corresponding theoretical price under a given model, where $V_t(\theta)$ is the model implied instantaneous volatility at t given model parameters θ. For each i and t, denote the percentage pricing error as

$$u_{i,t}(\theta) = \frac{C(t, m_i, \tau_i) - \hat{C}(t, m_i, \tau_i, V_t(\theta), \theta)}{C(t, m_i, \tau i)}, \tag{86.36}$$

where $V_t(\theta)$ is defined as

$$V_t(\theta) = \arg\min_{\{V_t\}} \sum_{i=1}^{M} \left[\frac{C(t, m_i, \tau_i) - \hat{C}(t, m_i, \tau_i, V_t, \theta)}{C(t, m_i, \tau_i)} \right]^2. \tag{86.37}$$

[14] Due to the wide range of moneyness and maturities of the difference caps involved, there could be significant differences in the prices of difference caps. Using percentage pricing errors helps to mitigate this problem.

We provide empirical evidence on the performance of six different models in capturing the cap volatility smile. The first three models, denoted as SV1, SV2, and SV3, allow one, two, and three principal components to drive the forward rate curve, respectively, each with its own stochastic volatility. The next three models, denoted as SVJ1, SVJ2, and SVJ3, introduce jumps in LIBOR rates in each of the previous SV models. SVJ3 is the most comprehensive model and nests all the others as special cases. We first examine the separate performance of each of the SV and SVJ models, and then we compare performance across the two classes of models. The estimation of all models is based on the principal components extracted from historical LIBOR forward rates between June 1997 and July 2000.[15]

The SV models contribute to cap pricing in four important ways. First, the three principal components capture variations in the levels of LIBOR rates caused by innovations in the "level," "slope," and "curvature" factors. Second, the stochastic volatility factors capture the fluctuations in the volatilities of LIBOR rates reflected in the Black implied volatilities of ATM caps.[16] Third, the stochastic volatility factors also introduce fatter tails in LIBOR rate distributions than implied by the lognormal model, which helps capture the volatility smile. Finally, given our model structure, innovations of stochastic volatility factors also affect the covariances between LIBOR rates with different maturities. The first three factors, however, are more important for our applications, because difference caps are much less sensitive to time-varying correlations than swaptions.[17] Our discussion of the performance of the SV models focuses on the estimates of the model parameters and the latent volatility variables and the time-series and cross-sectional pricing errors of difference caps.

A comparison of the parameter estimates of the three SV models in Table 86.7 shows that the "level" factor has the most volatile stochastic volatility, followed, in decreasing order, by the "curvature" and "slope" factor. The long-run mean $(\overline{v}_1)$ and volatility of volatility (ξ_1) of the first volatility factor are much bigger than that of the other two factors. This suggests that the fluctuations in the volatilities of LIBOR rates are mainly due to the time-varying volatility of the "level" factor. The estimates of the volatility risk premium of the three models are significantly negative, suggesting that the stochastic volatility factors of longer maturity LIBOR rates under the forward measure are less volatile with lower long-run mean and faster speed of mean reversion. This is consistent with the fact that the Black implied volatilities of longer maturity *difference caps* are less volatile than that of short-term *difference caps*.

[15]The LIBOR forward curve is constructed from weekly LIBOR and swap rates from Datastream following the bootstrapping procedure of LSS (2001).

[16]Throughout our discussion, volatilities of LIBOR rates refer to market implied volatilities from cap prices and are different from volatilities estimated from historical data.

[17]See Han (2002) for more detailed discussions on the impact of time-varying correlations for pricing swaptions.

Table 86.7 Parameter estimates of stochastic volatility models

	SV1		SV2		SV3	
Parameter	Estimate	Std. err	Estimate	Std. err	Estimate	Std. err
κ_1	0.0179	0.0144	0.0091	0.0111	0.0067	0.0148
κ_2			0.1387	0.0050	0.0052	0.0022
κ_3					0.0072	0.0104
$\bar{v}_1$	1.3727	1.1077	1.7100	2.0704	2.1448	4.7567
$\bar{v}_2$			0.0097	0.0006	0.0344	0.0142
$\bar{v}_3$					0.1305	0.1895
ζ_1	1.0803	0.0105	0.8992	0.0068	0.8489	0.0098
ζ_2			0.0285	0.0050	0.0117	0.0065
ζ_3					0.1365	0.0059
c_{1v}	−0.0022	0.0000	−0.0031	0.0000	−0.0015	0.0000
c_{2v}			−0.0057	0.0010	−0.0007	0.0001
c_{3v}					−0.0095	0.0003
Objective function	0.0834		0.0758		0.0692	

This table reports parameter estimates and standard errors of the one-, two-, and three-factor stochastic volatility models. The estimates are obtained by minimizing the sum of squared percentage pricing errors (SSEs) of difference caps in 53 moneyness and maturity categories observed on a weekly frequency from August 1, 2000, to September 23, 2003. The objective functions reported in the table are rescaled SSEs over the entire sample at the estimated model parameters and are equal to RMSE of difference caps. The volatility risk premium of the ith stochastic volatility factor for forward measure Q^{k+1} is defined as $\eta_i^{k+1} = c_{iv}(T_k - 1)$

Our parameter estimates are consistent with the volatility variables inferred from the prices of *difference caps*. The volatility of the "level" factor is the highest among the three (although at lower absolute levels in the more sophisticated models). It starts at a low level and steadily increases and stabilizes at a high level in the later part of the sample period. The volatility of the "slope" factor is much lower and relatively stable during the whole sample period. The volatility of the "curvature" factor is generally between that of the first and second factors. The steady increase of the volatility of the "level" factor is consistent with the increase of Black implied volatilities of ATM difference caps throughout our sample period. In fact, the correlation between the Black implied volatilities of most difference caps and the implied volatility of the "level" factor is higher than 0.8. The correlation between Black implied volatilities and the other two volatility factors is much weaker. The importance of stochastic volatility is obvious: the fluctuations in Black implied volatilities show that a model with constant volatility simply would not be able to capture even the general level of cap prices.

The other aspects of model performance are the time-series and cross-sectional pricing errors of *difference caps*. The likelihood ratio tests in Panel A of Table 86.8 overwhelmingly reject SV1 and SV2 in favor of SV2 and SV3, respectively. The Diebold-Mariano statistics in Panel A of Table 86.8 also show that SV2 and SV3 have significantly smaller SSEs than SV1 and SV2, respectively, suggesting that the

Table 86.8 Comparison of the performance of stochastic volatility models

Models	D-M stats	Likelihood ratio stats $x^2(168)$
SV2-SV1	−1.931	1,624
SV3-SV2	−6.351	1,557

Moneyness	1.5 year	2 year	2.5 year	3 year	3.5 year	4 year	4.5 year	5 year	6 year	7 year	8 year	9 year	10 year
Panel A. Likelihood ratio and Diebold-Mariano statistics for overall model performance based on SSEs													
0.7	–	–	–	–	**2.433**	**2.895**	**2.385**	**5.414**	**4.107**	**5.701**	**2.665**	−1.159	−1.299
0.8	–	–	−0.061	0.928	1.838	1.840	**2.169**	**6.676**	**3.036**	**2.274**	−0.135	−1.796	−1.590
0.9	–	−1.553	**−1.988**	**−2.218**	−1.064	−1.222	**−3.410**	−1.497	0.354	−0.555	−1.320	−1.439	−1.581
1.0	−0.295	**−5.068**	**−2.693**	−1.427	−1.350	−1.676	**−3.498**	**−3.479**	**−2.120**	−1.734	−1.523	−0.133	**−2.016**
1.1	−1.260	**−4.347**	−1.522	0.086	−1.492	**−3.134**	**−3.439**	**−3.966**	–	–	–	–	–
Panel B. Diebold-Mariano statistics between SV3 and SV2 for individual difference caps based on squared percentage pricing errors													
0.7	–	–	–	–	1.493	1.379	0.229	−0.840	**−3.284**	**−5.867**	**−4.280**	−0.057	**−2.236**
0.8	–	–	**−3.135**	−1.212	1.599	1.682	−0.052	0.592	**−3.204**	**−6.948**	**−4.703**	1.437	−1.079
0.9	–	**−2.897**	**−3.771**	**−3.211**	1.417	1.196	**−2.237**	−1.570	−1.932	**−6.920**	−1.230	**−2.036**	−1.020
1.0	−0.849	**−3.020**	**−3.115**	−0.122	0.328	**−3.288**	**−3.342**	**−3.103**	1.351	1.338	0.139	**−4.170**	−0.193
1.1	0.847	**−2.861**	0.675	0.315	**−3.650**	**−3.523**	**−2.923**	**−2.853**	–	–	–	–	–

This table reports model comparison based on likelihood ratio and Diebold-Mariano statistics. The total number of observations (both cross sectional and time series), which equals 8,545 over the entire sample, times the difference between the logarithms of the SSEs between two models follows a χ^2 distribution asymptotically. We treat implied volatility variables as parameters. Thus, the degree of freedom of the χ^2 distribution is 168 for the pairs of SV2/SV1 and SV3/SV2, because SV2 and SV3 have four more parameters and 164 additional implied volatility variables than SV1 and SV2, respectively. The 1 % critical value of $\chi^2(168)$ is 214. The Diebold-Mariano statistics are calculated according to Eq. 86.14 with a lag order q of 40 and follow an asymptotic standard normal distribution under the null hypothesis of equal pricing errors. A negative statistic means that the more sophisticated model has smaller pricing errors. Bold entries mean that the statistics are significant at the 5 % level more sophisticated SV models improve the pricing of all caps. The time series of RMSEs of the three SV models over our sample period[18] suggest that except for two special periods where all models have extremely large pricing errors, the RMSEs of all models are rather uniform with the best model (SV3) having RMSEs slightly above 5 %. The two special periods with high pricing errors cover the period between the second half of December 2000 and the first half of January 2001, and the first half of October 2001, and coincide with high prepayments in mortgage-backed securities (MBS). Indeed, the MBAA refinancing index and prepayment speed (see Fig. 3 of Duarte 2004) show that after a long period of low prepayments between the middle of 1999 and late 2000, prepayments dramatically increased at the end of 2000 and the beginning of 2001. There is also a dramatic increase of prepayments at the beginning of October 2001. As widely recognized in the fixed-income market,[19] excessive hedging demands for prepayment risk using interest rate derivatives may push derivative prices away from their equilibrium values, which could explain the failure of our models during these two special periods.[20]

In addition to overall model performance as measured by SSEs, we also examine the cross-sectional pricing errors of difference caps with different moneyness and maturities. We first look at the squared percentage pricing errors, which measure both the bias and variability of the pricing errors. Then we look at the average percentage pricing errors (the difference between market and model prices divided by the market price) to see whether SV models can on average capture the volatility smile in the cap market.

The Diebold-Mariano statistics of squared percentage pricing errors of individual *difference caps* between SV2 and SV1 in Panel B of Table 86.8 show that SV2 reduces the pricing errors of SV1 for some but not all *difference caps*. SV2 has the most significant reductions in pricing errors of SV1 for mid- and short-term around-the-money difference caps. On the other hand, SV2 has larger pricing errors for deep ITM *difference caps*. The Diebold-Mariano statistics between SV3 and SV2 in Panel C of Table 86.8 show that SV3 significantly reduces the pricing errors of many short- (2–3 years) and midterm around-the-money and long-term (6–10 years) ITM *difference caps*.

Table 86.9 reports the average percentage pricing errors of all *difference caps* under the three SV models. Panel A of Table 86.9 shows that, on average, SV1 underprices short-term and overprices mid- and long-term ATM *difference caps* and underprices ITM and overprices OTM difference caps. This suggests that SV1 cannot generate enough skewness in the implied volatilities to be consistent with

---

[18]RMSE of a model at t is calculated as $\sqrt{u_t'\left(\hat{\theta}\right)u_t\left(\hat{\theta}\right)/M}$. We plot RMSEs instead of SSEs because the former provides a more direct measure of average percentage pricing errors of difference caps.

[19]We would like to thank Pierre Grellet Aumont from Deutsche Bank for his helpful discussions on the influence of MBS markets on OTC interest rate derivatives.

[20]While the prepayment rates were also high in later part of 2002 and for most of 2003, they might not have come as surprises to participants in the MBS markets given the two previous special periods.

Table 86.9 Average percentage pricing errors of stochastic volatility models

Moneyness	1.5 year	2 year	2.5 year	3 year	3.5 year	4 year	4.5 year	5 year	6 year	7 year	8 year	9 year	10 year
Panel A. Average percentage pricing errors of SV1													
0.7	–	–	–	–	0.034	0.0258	0.0122	0.0339	0.0361	0.0503	0.0344	0.0297	0.0402
0.8	–	–	0.0434	0.0412	0.0323	0.018	0.0106	0.0332	0.0322	0.0468	0.0299	0.0244	0.0325
0.9	–	0.1092	0.0534	0.0433	0.0315	0.01	0.0003	0.0208	0.0186	0.0348	0.0101	0.0062	0.0158
1.0	0.0293	0.1217	0.0575	0.0378	0.0227	−0.0081	−0.0259	−0.0073	−0.0079	0.0088	−0.0114	−0.0192	−0.0062
1.1	−0.1187	0.0604	−0.0029	−0.0229	−0.034	−0.0712	−0.0815	−0.0562	–	–	–	–	–
Panel B. Average percentage pricing errors of SV2													
0.7	–	–	–	–	0.0482	0.0425	0.0304	0.0524	0.0544	0.0663	0.0456	0.0304	0.0378
0.8	–	–	0.0509	0.051	0.0443	0.032	0.0258	0.0486	0.0472	0.0586	0.0344	0.0138	0.0202
0.9	–	0.1059	0.0498	0.0421	0.0333	0.0145	0.0069	0.0284	0.0265	0.0392	0.0054	−0.0184	−0.008
1.0	−0.0002	0.0985	0.0369	0.0231	0.0134	−0.0123	−0.0261	−0.005	−0.0042	0.008	−0.024	−0.0572	−0.0403
1.1	−0.1056	0.0584	−0.0085	−0.026	−0.0326	−0.0653	−0.0721	−0.0454	–	–	–	–	–
Panel C. Average percentage pricing errors of SV3													
0.7	–	–	–	–	0.0489	0.0437	0.0308	0.0494	0.0431	0.0466	0.031	0.03	0.028
0.8	–	–	0.044	0.0476	0.0462	0.0378	0.0322	0.0506	0.0367	0.0365	0.0226	0.0249	0.0139
0.9	–	0.0917	0.0367	0.0379	0.0398	0.0288	0.0226	0.0377	0.0178	0.0145	−0.0026	0.0068	−0.0109
1.0	−0.0126	0.0782	0.0198	0.0194	0.0252	0.0105	−0.0012	0.011	−0.0129	−0.0221	−0.0299	−0.0192	−0.0432
1.1	−0.1184	0.0314	−0.0323	−0.0336	−0.0212	−0.0397	−0.0438	−0.0292	–	–	–	–	–

This table reports average percentage pricing errors of difference caps with different moneyness and maturities of the three stochastic volatility models. Average percentage pricing errors are defined as the difference between market price and model price divided by the market price the data. Panel B shows that SV2 has some improvements over SV1, mainly for some short-term (less than 3.5 years) ATM and midterm (3.5–5 years) slightly OTM *difference caps*. But SV2 has worse performance for most deep ITM ($m = 0.7$ and 0.8) difference caps: it actually worsens the underpricing of ITM caps. Panel C of Table 86.9 shows that relative to SV1 and SV2, SV3 has smaller average percentage pricing errors for most long-term (7–10 years) ITM, midterm (3.5–5 years) OTM, and short-term (2–2.5 years) ATM *difference caps* and bigger average percentage pricing errors for midterm (3.5–6 years) ITM difference caps. There is still significant underpricing of ITM and overpricing of OTM difference caps under SV3.

Overall, the results show that stochastic volatility factors are essential for capturing the time-varying volatilities of LIBOR rates. The Diebold-Mariano statistics in Table 86.8 shows that in general more sophisticated SV models have smaller pricing errors than simpler models, although the improvements are more important for close-to-the-money difference caps. The average percentage pricing errors in Table 86.9 show that, however, even the most sophisticated SV model cannot generate enough volatility skew to be consistent with the data. While previous studies, such as Han (2007), have shown that a three-factor stochastic volatility model similar to ours performs well in pricing ATM caps and swaptions, our analysis shows that the model fails to completely capture the volatility smile in the cap markets. Our findings highlight the importance of studying the relative pricing of caps with different moneyness to reveal the inadequacies of existing term structure models; the same inadequacies cannot be obtained from studying only ATM options.

One important reason for the failure of SV models is that the stochastic volatility factors are independent of LIBOR rates. As a result, the SV models can only generate a symmetric volatility smile, but not the asymmetric smile or skew observed in the data. The pattern of the smile in the cap market is rather similar to that of index options: ITM calls (and OTM puts) are overpriced, and OTM calls (and ITM puts) are underpriced relative to the Black model. Similarly, the smile in the cap market could be due to a market expectation of dramatically declining LIBOR rates. In this section, we examine the contribution of jumps in LIBOR rates in capturing the volatility smile. Our discussion of the performance of the SVJ models parallels that of the SV models.

Parameter estimates in Table 86.10 show that the three stochastic volatility factors of the SVJ models resemble that of the SV models closely. The “level” factor still has the most volatile stochastic volatility, followed by the “curvature” and the “slope” factor. With the inclusion of jumps, the stochastic volatility factors in the SVJ models, especially that of the “level” factor, tend to be less volatile than that of the SV models (lower long-run mean and volatility of volatility). Negative estimates of the volatility risk premium show that the volatility of the longer maturity LIBOR rates under the forward measure has lower long-run mean and faster speed of mean reversion.

Most importantly, we find overwhelming evidence of strong negative jumps in LIBOR rates under the forward measure. To the extent that cap prices reflect market expectations of future evolutions of LIBOR rates, the evidence suggests that the market expects a dramatic decline in LIBOR rates over our sample

Table 86.10 Parameter estimates of stochastic volatility and jump models

	SVJ1		SVJ2		SVJ3	
Parameter	Estimate	Std. err	Estimate	Std. err	Estimate	Std. err
κ_1	0.1377	0.0085	0.0062	0.0057	0.0069	0.0079
κ_2			0.0050	0.0001	0.0032	0.0000
κ_3					0.0049	0.0073
$\bar{v}_1$	0.1312	0.0084	0.7929	0.7369	0.9626	1.1126
$\bar{v}_2$			0.3410	0.0030	0.2051	0.0021
$\bar{v}_3$					0.2628	0.3973
ζ_1	0.8233	0.0057	0.7772	0.0036	0.6967	0.0049
ζ_2			0.0061	0.0104	0.0091	0.0042
ζ_3					0.1517	0.0035
c_{1v}	−0.0041	0.0000	−0.0049	0.0000	−0.0024	0.0000
c_{2v}			−0.0270	0.0464	−0.0007	0.0006
c_{3v}					−0.0103	0.0002
λ	0.0134	0.0001	0.0159	0.0001	0.0132	0.0001
μ_J	−3.8736	0.0038	−3.8517	0.0036	−3.8433	0.0063
c_J	0.2632	0.0012	0.3253	0.0010	0.2473	0.0017
σ_J	0.0001	3.2862	0.0003	0.8723	0.0032	0.1621
Objective function	0.0748		0.0670		0.0622	

This table reports parameter estimates and standard errors of the one-, two-, and three-factor stochastic volatility and jump models. The estimates are obtained by minimizing the sum of squared percentage pricing errors (SSEs) of difference caps in 53 moneyness and maturity categories observed on a weekly frequency from August 1, 2000, to September 23, 2003. The objective functions reported in the table are rescaled SSEs over the entire sample at the estimated model parameters and are equal to RMSE of difference caps. The volatility risk premium of the ith stochastic volatility factor and the jump risk premium for forward measure Q^{k+1} are defined as $\eta_i^{k+1} = c_{iv}(T_k - 1)$ and $\mu_J{}^{k+1} = \mu_J + c_J(T_k - 1)$, respectively

period. Such an expectation might be justifiable given that the economy has been in recession during a major part of our sample period. This is similar to the volatility skew in the index equity option market, which reflects investors' fear of the stock market crash such as that of 1987. Compared to the estimates from index options (see, e.g., Pan 2002), we see lower estimates of jump intensity (about 1.5 % per annual), but much higher estimates of jump size. The positive estimates of a jump risk premium suggest that the jump magnitude of longer maturity forward rates tends to be smaller. Under SVJ3, the mean relative jump sizes, exp $(\mu_J + c_J(T_k - 1) + \sigma_J^2/2) - 1$, for 1-, 5-, and 10-year LIBOR rates are −97 %, −94 %, and −80 %, respectively. However, we do not find any incidents of negative moves in LIBOR rates under the physical measure with a size close to that under the forward measure. This big discrepancy between jump sizes under the physical and forward measures resembles that between the physical and risk-neutral measures for index options (see, e.g., Pan 2002). This could be a result of a huge jump risk premium.

The likelihood ratio tests in Panel A of Table 86.11 again overwhelmingly reject SVJ1 and SVJ2 in favor of SVJ2 and SVJ3, respectively. The Diebold-Mariano

Table 86.11 Comparison of the performance of stochastic volatility and jump models

Models	D-M stats	Likelihood ratio stats $\chi^2(168)$
Panel A. Likelihood ratio and Diebold-Mariano statistics for overall model performance based on SSEs		
SVJ2-SVJ1	**−2.240**	**1,886**
SVJ3-SVJ2	**−7.149**	**1,256**

Moneyness	1.5 year	2 year	2.5 year	3 year	3.5 year	4 year	4.5 year	5 year	6 year	7 year	8 year	9 year	10 year
Panel B. Diebold-Mariano statistics between SVJ2 and SVJ1 for individual difference caps based on squared percentage pricing errors													
0.7	–	–	–	–	−0.234	−0.308	−1.198	−0.467	−0.188	0.675	−0.240	−0.774	−0.180
0.8	–	–	−1.346	−0.537	−0.684	−1.031	−1.892	−1.372	−0.684	−0.365	−0.749	−1.837	−1.169
0.9	–	−1.530	−1.914	−0.934	−1.036	−1.463	**−3.882**	**−3.253**	−0.920	−1.588	**−2.395**	**−3.287**	−0.686
1.0	−0.094	**−3.300**	**−2.472**	−1.265	−1.358	−1.647	**−2.186**	−2.**020**	−0.573	−1.674	−1.396	**−2.540**	−0.799
1.1	−1.395	**−5.341**	−0.775	0.156	−0.931	**−3.141**	**−3.000**	**−2.107**	–	–	–	–	–
Panel C. Diebold-Mariano statistics between SVJ3 and SVJ2 for individual difference caps based on squared percentage pricing errors													
0.7	–	–	–	–	0.551	0.690	−1.023	−1.023	−1.133	**−2.550**	−1.469	−0.605	−1.920
0.8	–	–	−1.036	−0.159	1.650	1.609	−1.714	−1.898	−0.778	**−3.191**	**−3.992**	**−2.951**	**−3.778**
0.9	–	−1.235	**−2.057**	−0.328	2.108	1.183	**−2.011**	−1.361	−0.249	**−2.784**	−1.408	**−3.411**	**−2.994**
1.0	−1.594	−1.245	**−2.047**	−0.553	−0.289	−0.463	**−2.488**	−1.317	2.780	0.182	−0.551	−1.542	−1.207
1.1	−0.877	−1.583	−0.365	−0.334	−1.088	−2.040	**−3.302**	−1.259	–	–	–	–	–

This table reports model comparison based on likelihood ratio and Diebold-Mariano statistics. The total number of observations (both cross sectional and time series), which equals 8,545 over the entire sample, times the difference between the logarithms of the SSEs between two models follows a χ^2 distribution asymptotically. We treat implied volatility variables as parameters. Thus, the degree of freedom of the χ^2 distribution is 168 for the pairs of SVJ2/SVJ1 and SVJ3/SVJ2, because SVJ2 and SVJ3 have four more parameters and 164 additional implied volatility variables than SVJ1 and SVJ2, respectively. The 1 % critical value of χ^2 (168) is 214. The Diebold-Mariano statistics are calculated according to Eq. 86.14 with a lag order q of 40 and follow an asymptotic standard normal distribution under the null hypothesis of equal pricing errors. A negative statistic means that the more sophisticated model has smaller pricing errors. Bold entries mean that the statistics are significant at the 5 % level statistics in Panel A of Table 86.11 show that SVJ2 and SVJ3 have significantly smaller SSEs than SVJ1 and SVJ2, respectively, suggesting that the more sophisticated SVJ models significantly improve the pricing of all *difference caps*. The Diebold-Mariano statistics of squared percentage pricing errors of individual *difference caps* in Panel B of Table 86.11 show that SVJ2 significantly improves the performance of SVJ1 for long-, mid-, and short-term around-the-money *difference caps*. The Diebold-Mariano statistics in Panel C of Table 86.11 show that SVJ3 significantly reduces the pricing errors of SVJ2 for long-term ITM and some mid- and short-term around-the-money *difference caps*. Table 86.12 shows the average percentage pricing errors also improve over the SV models.

Table 86.13 compares the performance of the SVJ and SV models. During the first 20 weeks of our sample, the SVJ models have much higher RMSEs than the SV models. As a result, the likelihood ratio and Diebold-Mariano statistics between the three pairs of SVJ and SV models over the entire sample are somewhat smaller than that of the sample period without the first 20 weeks. Nonetheless, all the SV models are overwhelmingly rejected in favor of their corresponding SVJ models by both tests. The Diebold-Mariano statistics of individual *difference caps* in Panels B, C, and D show that the SVJ models significantly improve the performance of the SV models for most *difference caps* across moneyness and maturity. The most interesting results are in Panel D, which show that SVJ3 significantly reduces the pricing errors of most ITM *difference caps* of SV3, strongly suggesting that the negative jumps are essential for capturing the asymmetric smile in the cap market.

Our analysis shows that a low-dimensional model with three principal components driving the forward rate curve, stochastic volatility of each component, and strong negative jumps captures the volatility smile in the cap markets reasonably well. The three yield factors capture the variations of the levels of LIBOR rates, while the stochastic volatility factors are essential to capture the time-varying volatilities of LIBOR rates. Even though the SV models can price ATM caps reasonably well, they fail to capture the volatility smile in the cap market. Instead, significant negative jumps in LIBOR rates are needed to capture the smile. These results highlight the importance of studying the pricing of caps across moneyness: the importance of negative jumps is revealed only through the pricing of away-from-the-money caps. Excluding the first 20 weeks and the two special periods, SVJ3 has a reasonably good pricing performance with an average RMSE of 4.5 %. Given that the bid-ask spread is about 2–5 % in our sample for ATM caps, and because ITM and OTM caps tend to have even higher percentage spreads,[21] this can be interpreted as a good performance.

Despite its good performance, there are strong indications that SVJ3 is misspecified and the inadequacies of the model seem to be related to MBS markets. For example, while SVJ3 works reasonably well for most of the sample period, it has large pricing errors in several special periods coinciding with high prepayment activities in the MBS markets. Moreover, even though we assume that the

[21]See, for example, Deuskar et al. (2003).

Table 86.12 Average percentage pricing errors of stochastic volatility and jump models

Moneyness	1.5 year	2 year	2.5 year	3 year	3.5 year	4 year	4.5 year	5 year	6 year	7 year	8 year	9 year	10 year
Panel A. Average percentage pricing errors of SVJ1													
0.7	–	–	–	–	0.0164	0.0073	−0.0092	0.01	0.0102	0.0209	−0.0001	−0.0061	0.0077
0.8	–	–	0.014	0.0167	0.0116	−0.0014	−0.0091	0.0111	0.007	0.0207	−0.0009	−0.0076	0.0053
0.9	–	0.0682	0.0146	0.0132	0.0112	−0.0035	−0.0103	0.0104	0.0038	0.0204	−0.0062	−0.0114	0.0042
1.0	−0.009	0.0839	0.0233	0.016	0.0158	−0.0004	−0.0105	0.0105	0.0062	0.0194	0.0013	−0.0083	0.0094
1.1	−0.098	0.0625	−0.0038	−0.0144	−0.0086	−0.0255	−0.0199	0.0094	–	–	–	–	–
Panel B. Average percentage pricing errors of SVJ2													
0.7	–	–	–	–	0.0243	0.0148	−0.0008	0.0188	0.0175	0.0279	0.0116	0.0106	0.0256
0.8	–	–	0.0232	0.0271	0.0211	0.0062	−0.0035	0.0172	0.0137	0.0255	0.0081	0.0061	0.0139
0.9	–	0.0698	0.019	0.0205	0.0172	−0.0012	−0.0119	0.0068	0.0039	0.0198	−0.0041	−0.0047	−0.002
1.0	−0.0375	0.0668	0.013	0.0131	0.015	−0.0058	−0.0214	−0.0047	−0.0054	0.0127	−0.0058	−0.0112	−0.0128
1.1	−0.089	0.0612	−0.0048	−0.0094	0.0003	−0.0215	−0.0273	−0.0076	–	–	–	–	–
Panel C. Average percentage pricing errors of SVJ3													
0.7	–	–	–	–	0.0261	0.0176	0.0008	0.017	0.0085	0.0167	0.0008	−0.0049	−0.0021
0.8	–	–	0.0222	0.0249	0.0223	0.0115	0.0027	0.0185	0.0016	0.0131	0.004	−0.0008	−0.0063
0.9	–	0.0713	0.014	0.0155	0.0182	0.0073	−0.0002	0.0129	−0.0108	0.0072	0.0044	0.0048	−0.0092
1.0	−0.0204	0.0657	0.005	0.0054	0.0142	0.0033	−0.0068	0.0047	−0.0232	−0.001	0.019	0.0206	−0.0058
1.1	−0.0688	0.0528	−0.02	−0.0242	−0.0085	−0.0199	−0.0182	−0.0028	–	–	–	–	–

This table reports average percentage pricing errors of difference caps with different moneyness and maturities of the three stochastic volatility and jump models. Average percentage pricing errors are defined as the difference between market price and model price divided by market price

Table 86.13 Comparison of the performance of SV and SVJ models

Models	D-M stats (whole sample)	D-M stats (without first 20 weeks)	Likelihood ratio stats χ^2 (4) (whole sample)	Likelihood ratio stats χ^2 (4) (without first 20 weeks)
Panel A. Likelihood ratio and Diebold-Mariano statistics for overall model performance based on SSEs				
SVJ1-SV1	**−2.972**	**−3.006**	**1,854**	**2,437**
SVJ2-SV2	**−3.580**	**−4.017**	**2,115**	**2,688**
SVJ3-SV3	**−3.078**	**−3.165**	**1,814**	**2,497**

Moneyness	1.5 year	2 year	2.5 year	3 year	3.5 year	4 year	4.5 year	5 year	6 year	7 year	8 year	9 year	10 year
Panel B. Diebold-Mariano statistics between SVJ1 and SV1 for individual difference caps based on squared percentage pricing errors (without first 20 weeks)													
0.7	–	–	–	–	**−3.050**	**−3.504**	−0.504	−1.904	−4.**950**	**−3.506**	**−3.827**	**−2.068**	**−2.182**
0.8	–	–	**−3.243**	**−12.68**	**−9.171**	−1.520	−0.692	−1.195	−2.**245**	**−1.986**	−1.920	−1.353	−1.406
0.9	–	**−7.162**	**−9.488**	**−2.773**	−1.948	−1.030	−1.087	−0.923	−0.820	−0.934	−1.176	−1.109	−1.166
1.0	0.670	**−5.478**	**−2.844**	−1.294	−1.036	**−4.001**	**−3.204**	−1.812	**−2.331**	−1.099	−1.699	**−2.151**	**−2.237**
1.1	−0.927	−0.435	−0.111	−0.261	−1.870	**−2.350**	−1.710	−0.892	–	–	–	–	–
Panel C. Diebold-Mariano statistics between SVJ2 and SV2 for individual difference caps based on squared percentage pricing errors (without first 20 weeks)													
0.7	–	–	–	–	**−9.696**	−7.**714**	**−2.879**	**−3.914**	**−14.01**	−7.**387**	**−7.865**	**−3.248**	**−1.637**
0.8	–	–	**−7.353**	**−5.908**	**−5.612**	**−2.917**	−1.669	**−2.277**	**−7.591**	**−4.610**	**−4.182**	0.397	2.377
0.9	–	**−7.013**	**−6.271**	**−2.446**	**−2.145**	−1.246	−1.047	−1.309	**−2.856**	−1.867	−0.183	0.239	3.098
1.0	1.057	−6.**025**	**−2.736**	−1.159	−0.688	**−4.478**	**−4.410**	**−3.754**	−0.404	−0.416	−0.881	**−2.504**	0.023
1.1	−0.969	−0.308	0.441	−1.284	**−2.179**	**−3.148**	**−2.874**	**−2.267**	–	–	–	–	–

(*continued*)

Table 86.13 (continued)

Moneyness	1.5 year	2 year	2.5 year	3 year	3.5 year	4 year	4.5 year	5 year	6 year	7 year	8 year	9 year	10 year
Panel D. Diebold-Mariano statistics between SVJ3 and SV3 for individual difference caps based on squared percentage pricing errors (without first 20 weeks)													
0.7	–	–	–	–	−7.**040**	−**8.358**	−**7.687**	−**10.49**	−**7.750**	−5.**817**	−5.**140**	−**3.433**	−**3.073**
0.8	–	–	−**4.732**	−**7.373**	−**21.74**	−**7.655**	−**5.145**	−**4.774**	−**6.711**	−3.030	−**2.650**	−**2.614**	−1.239
0.9	–	−**1.980**	−**2.571**	−**2.501**	−**6.715**	−**3.622**	−**1.985**	−**2.384**	−1.938	−1.114	−0.768	−4.119	−1.305
1.0	−0.530	−1.124	−1.305	−1.353	−1.909	−0.880	1.023	0.052	0.543	−**2.110**	−0.359	−0.492	−**2.417**
1.1	−1.178	1.395	−1.424	**2.218**	−1.834	−**2.151**	−1.537	−0.337	–	–	–	–	–

This table reports model comparison based on likelihood ratio and Diebold-Mariano statistics. The total number of observations (both cross sectional and time series), which equals 8,545 over the entire sample and 7,485 without the first 20 weeks, times the difference between the logarithms of the SSEs between two models follows a χ^2 distribution asymptotically. We treat implied volatility variables as parameters. Thus, the degree of freedom of the χ^2 distribution is 4 for the pairs of SVJ/SV1, SVJ2/SV2, and SVJ3/SV3, because SVJ models have four more parameters and equal number of additional implied volatility variables as the corresponding SV models. The 1 % critical value of χ^2 (4) is 13. The Diebold-Mariano statistics are calculated according to Eq. 86.14 with a lag order q of 40 and follow an asymptotic standard normal distribution under the null hypothesis of equal pricing errors. A negative statistic means that the more sophisticated model has smaller pricing errors. Bold entries mean that the statistics are significant at the 5 % level

Table 86.14 Correlations between LIBOR rates and implied volatility variables

	L(t, 1)	L(t, 3)	L(t, 5)	L(t, 7)	L(t, 9)	V1(t)	V2(t)	V3(t)
V1(t)	−0.8883	−0.8772	−0.8361	−0.7964	−0.7470	1	−0.4163	0.3842
V2(t)	0.1759	0.235	0.2071	0.1545	0.08278	−0.4163	1	−0.0372
V3(t)	−0.5951	−0.485	−0.4139	−0.3541	−0.3262	0.3842	−0.0372	1

This table reports the correlations between LIBOR rates and implied volatility variables from SVJ3. Given the parameter estimates of SVJ3 in Table 86.4, the implied volatility variables are estimated at t by minimizing the SSEs of all difference caps at t

stochastic volatility factors are independent of LIBOR rates, Table 86.14 shows strong negative correlations between the implied volatility variables of the first factor and the LIBOR rates. This result suggests that when interest rate is low, cap prices become too high for the model to capture and the implied volatilities have to become abnormally high to fit the observed cap prices. One possible explanation of the "leverage" effect is that higher demands for caps to hedge prepayments from MBS markets in low interest rate environments could artificially push up cap prices and implied volatilities. Therefore, extending our models to incorporate factors from MBS markets seems to be a promising direction of future research.

86.4 Nonparametric Estimation of the Forward Density

For LIBOR-based instruments such as caps, floors, and swaptions, it is convenient to consider pricing using the forward measure approach. We will therefore focus on the dynamics of LIBOR forward rate $L_k(t)$ under the forward measure $\mathbb{Q}^{k+1}$, which is essential for pricing caplets maturing at T_{k+1}. Under this measure, the discounted price of any security using $D_{k+1}(t)$ as the numeraire is a martingale. Thus, the time-t price of a caplet maturing at T_{k+1} with a strike price of X is

$$C(L_k(t), X, t, T_k) = \delta D_{k+1}(t) \int_X^{\infty} (y - X) p^{\mathbb{Q}^{k+1}}(L_k(T_k) = y | L_k(t)) dy, \qquad (86.38)$$

where $p^{\mathbb{Q}^{k+1}}(L_k(T_k) = y | L_k(t))$ is the conditional density of $L_k(T_k)$ under forward measure $\mathbb{Q}^{k+1}$. Once we know the forward density, we can price any security whose payoff on T_{k+1} depends only on $L_k(t)$ by discounting its expected payoff under $\mathbb{Q}^{k+1}$ using $D_{k+1}(t)$.

Existing term structure models rely on parametric assumptions on the distribution of $L_k(t)$ to obtain closed-form pricing formulae for caplets. For example, the standard LIBOR market models of Brace et al. (1997) and Miltersen et al. (1997) assume that $L_k(t)$ follows a lognormal distribution and price caplet using the Black formula. The models of Jarrow et al. (2007) assume that $L_k(t)$ follows affine jump diffusions of Duffie et al. (2000).

86.4.1 Nonparametric Method

We estimate the distribution of $L_k(t)$ under $\mathbb{Q}^{k+1}$ using the prices of a cross section of caplets that mature at T_{k+1} and have different strike prices. Following Breeden and Litzenberger (1979), we know that the density of $L_k(t)$ under $\mathbb{Q}^{k+1}$ is proportional to the second derivative of $C(L_k(t), t, T_k, X)$ with respect to X:

$$p^{\mathbb{Q}^{k+1}}(L_k(T_k)|L_k(t)) = \frac{1}{\delta D_{k+1}(t)} \frac{\partial^2 C(L_k(t), t, T_k, X)}{\partial X^2} |X = L_k(T_k). \tag{86.39}$$

In standard LIBOR market models, it is assumed that the conditional density of $L_k(T_k)$ depends only on the current LIBOR rate. This assumption, however, can be overly restrictive given the multifactor nature of term structure dynamics. For example, while the level factor can explain a large fraction (between 80 % and 90 %) of the variations of LIBOR rates, the slope factor still has significant explanatory power of interest rate variations. Moreover, there is overwhelming evidence that interest rate volatility is stochastic,[22] and it has been suggested that interest rate volatility is unspanned in the sense that it cannot be fully explained by the yield curve factors such as the level and slope factors.

One important innovation of our study is that we allow the volatility of $L_k(t)$ to be stochastic and the conditional density of $L_k(T_k)$ to depend on not only the level but also the slope and volatility factors of LIBOR rates. Denote the conditioning variables as $Z(t) = \{s(t), v(t)\}$, where $s(t)$ (the slope factor) is the difference between the 10- and 2-year LIBOR forward rates and $v(t)$ (the volatility factor) is the first principal component of EGARCH-filtered spot volatilities of LIBOR rates across all maturities. Under this generalization, the conditional density of $L_k(T_k)$ under the forward measure $\mathbb{Q}^{k+1}$ is given by

$$p^{\mathbb{Q}^{k+1}}(L_k(T_k)|L_k(t), Z(t)) = \frac{1}{\delta D_{k+1}(t)} \frac{\partial^2 C(L_k(t), X, t, T_k, Z(t))}{\partial X^2} |X = L_k(T_k). \tag{86.40}$$

Next we discuss how to estimate the SPDs by combining the forward and physical densities of LIBOR rates. Denote an SPD function as π. In general, π depends on multiple economic factors, and it is impossible to estimate it using interest rate caps alone. Given the available data, all we can estimate is the projection of π onto the future spot rate $L_k(T_k)$:

$$\pi_k(L_k(T_k); L_k(t), Z(t)) = E_t^{\mathbb{P}}[\pi | L_k(T_k); L_k(t), Z(t)], \tag{86.41}$$

where the expectation is taken under the physical measure. Then the price of the caplet can be calculated as

[22]See Andersen and Lund (1997), Ball and Torous (1999), Brenner et al. (1996), Chen and Scott (2001), and many others.

$$C(L_k(t), X, t, T_k, Z(t)) = \delta E_t^{\mathbb{P}}\left[\pi \cdot (L_k(T_k) - X)^+\right]$$
$$= \delta \int_X^{\infty} \pi_k(y)(y - X)p^{\mathbb{P}}(L_k(T_k) = y|L_k(t), Z(t))dy, \tag{86.42}$$

where the second equality is due to iterated expectation and $p^{\mathbb{P}}(L_k(T_k) = y|L_k(t), Z(t))$ is the conditional density of $L_k(T_k)$ under the physical measure.

Comparing Eqs. 86.2 and 86.6, we have

$$\pi_k(L_k(T_k); L_k(t), Z(t)) = D_{k+1}(t)\frac{p^{\mathbb{Q}^{k+1}}(L_k(T_k)|L_k(t), Z(t))}{p^{\mathbb{P}}(L_k(T_k)|L_k(t), Z(t))}. \tag{86.43}$$

Therefore, by combining the densities of $L_k(T_k)$ under $\mathbb{Q}^{k+1}$ and $\mathbb{P}$, we can estimate the projection of π onto $L_k(T_k)$. The SPDs contain rich information on how risks are priced in financial markets. While Ait-Sahalia and Lo (1998, 2000); Jackwerth (2000), Rosenberg and Engle (2002), and others estimate the SPDs using index options (i.e., the projection of π onto index returns), our analysis based on interest rate caps documents the dependence of the SPDs on term structure factors.

Similar to many existing studies, to reduce the dimensionality of the problem, we further assume that the caplet price is homogeneous of degree 1 in the current LIBOR rate:

$$C(L_k(t), X, t, T_k, Z(t)) = \delta D_{k+1}(t)L_k(t)C_M(M_k(t), t, T_k, Z(t)), \tag{86.44}$$

where $M_k(t) = X/L_k(t)$ represents the moneyness of the caplet. Hence, for the rest of the paper, we estimate the forward density of $L_k(T_k)/L_k(t)$ as the second derivative of the price function C_M with respect to M:

$$p^{\mathbb{Q}^{k+1}}\left(\frac{L_k(T_k)}{L_k(t)}|Z(t)\right) = \frac{1}{\delta D_{k+1}(t)}\frac{\partial^2 C_M(M_k(t), t, T_k, Z(t))}{\partial M^2}|M = L_k(T_k)/L_k(t). \tag{86.45}$$

86.4.2 Empirical Results

In this section, we present nonparametric estimates of the probability densities of LIBOR rates under physical and forward martingale measures. In particular, we document the dependence of the forward densities on the slope and volatility factors of LIBOR rates.

Figure 86.2 presents nonparametric estimates of the forward densities at different levels of the slope and volatility factors at 2-, 3-, 4-, and 5-year maturities. The two levels of the slope factor correspond to a flat and a steep forward curve, while the two levels of the volatility factor represent low and high volatility of LIBOR rates. The 95 % confidence intervals are obtained through simulation. The forward densities should have a zero mean since LIBOR rates under appropriate forward measures are martingales. The expected log percentage changes of the LIBOR rates are slightly negative due to an adjustment from the Jensen's inequality. We normalize the forward densities so that they integrate to one. However, we do not have enough data at the right tail of the distribution at 4- and 5-year maturities. We do not extrapolate the data to avoid potential biases.

Figure 86.2 documents three important features of the nonparametric LIBOR forward densities. First, the lognormal assumption underlying the popular LIBOR market models is grossly violated in the data, and the forward densities across all maturities are significantly negatively skewed. Second, all the forward densities depend significantly on the slope of the term structure. For example, moving from a flat to a steep term structure, the forward densities across all maturities become much more dispersed and more negatively skewed. Third, the forward densities also depend on the volatility factor. Under both flat and steep term structures, the forward densities generally become more compact when the volatility factor increases. This is consistent with a mean-reverting volatility process: high volatility right now leads to low volatility in the future and more compact forward densities.

To better illustrate the dependence of the forward densities on the two conditioning variables, we also regress the quantiles of the forward densities on the two factors. We choose quantiles instead of moments of the forward densities in our regressions for two reasons. First, quantiles are much easier to estimate. While quantiles can be obtained from the CDF function, which is the first derivative of the price function, moments require integrations of the forward density, which is the second derivative of the price function. Second, a wide range of quantiles provide a better characterization of the forward densities than a few moments, especially for the tail behaviors of the densities.

Suppose we consider I and J levels of the transformed slope and volatility factors in our empirical analysis. For a given level of the two conditioning variables (s_i, v_j), we first obtain a nonparametric estimate of the forward density at a given maturity and its quantiles $Q_x(s_i, v_j)$, where x can range from 0 % to 100 %. Then we consider the following regression model:

$$Q_x(S_i, vj) = b_{0x} + b_{1x} \cdot s_i + b_{2x} \cdot v_j + b_{3x} \cdot s_i \cdot v_j + \varepsilon_x, \tag{86.46}$$

where $i = 1, 2, \ldots, I$, and $j = 1, 2, \ldots, J$. We include the interaction term to capture potential nonlinear dependence of the forward densities on the two conditioning variables.

Figure 86.3 reports regression coefficients of the slope and volatility factors for the most complete range of quantiles at each maturity, i.e., b_{1x} and b_{2x} as a function of x. While Fig. 86.3 includes only the slope and volatility factors as explanatory variables, Fig. 86.4 contains their interaction term as well. Though in results not reported here we also include lagged conditioning variables in our regressions, their coefficients are generally not statistically significant.

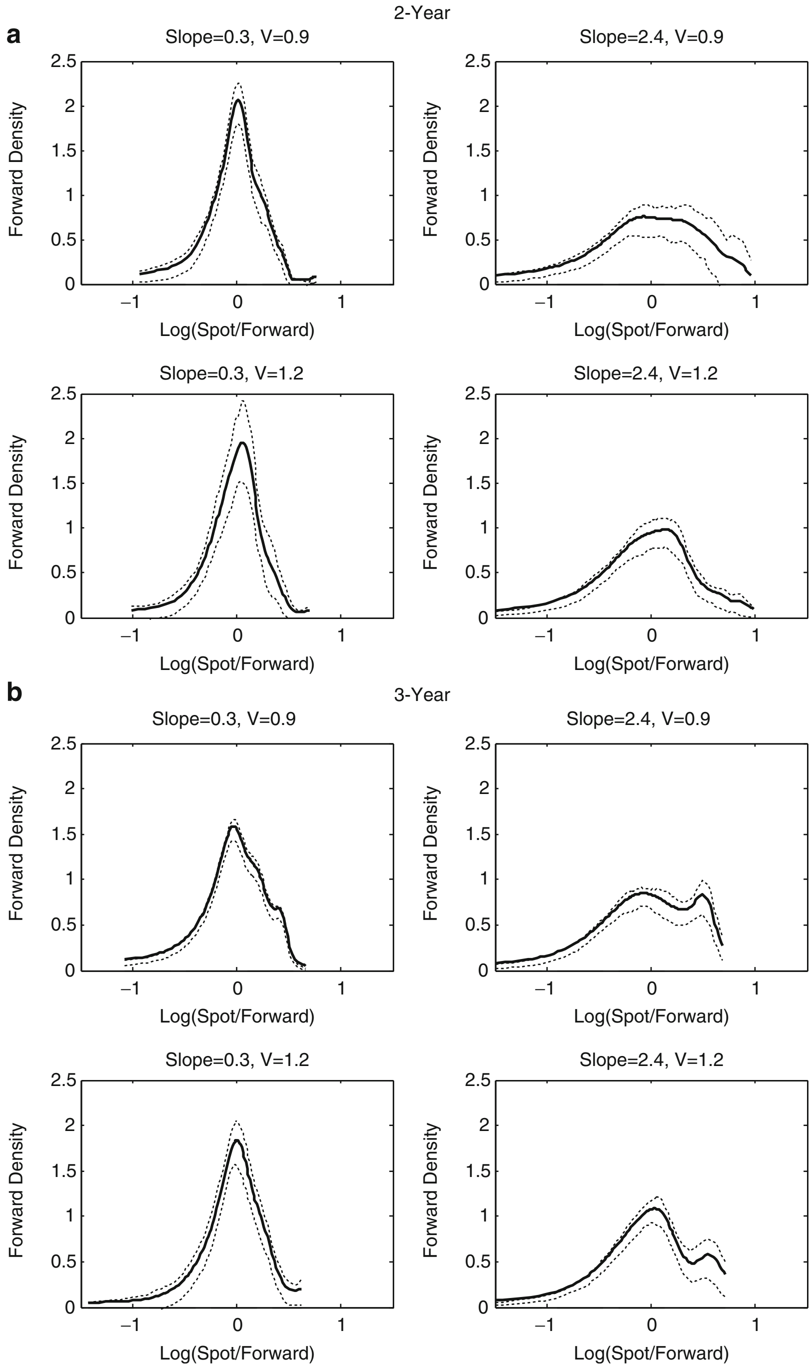


Fig. 86.2 (continued)

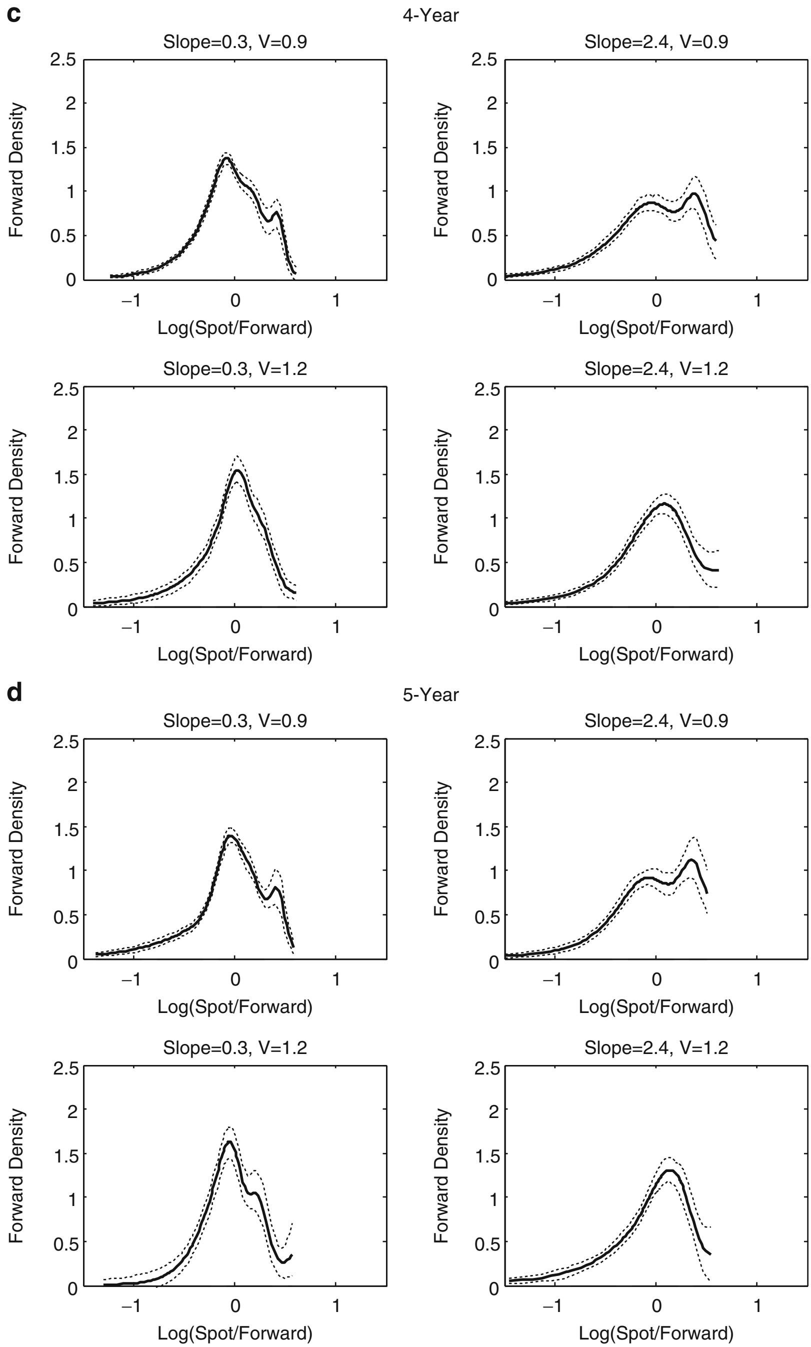


Fig. 86.2 Nonparametric estimates of the LIBOR forward densities at different levels of the slope and volatility factors. The slope factor is defined as the difference between the 10- and 2-year

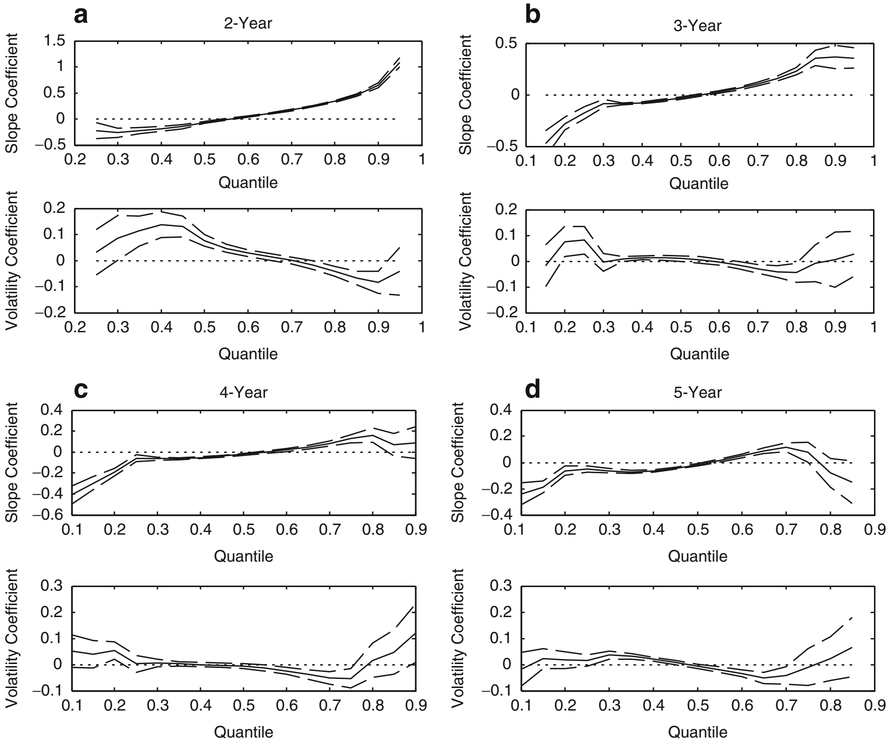


Fig. 86.3 Impacts of the slope and volatility factors on LIBOR forward densities. This figure reports regression coefficients of different quantiles of the forward densities at 2-, 3-, 4-, and 5-year maturities on the slope and volatility factors of LIBOR rates in Eq. 86.27 without the interaction term

The regression results in Fig. 86.3 are generally consistent with the main findings in Fig. 86.2. The slope coefficients are generally negative (positive) for the left (right) half of the distribution and become more negative or positive at both tails. Consistent with Fig. 86.2, this result suggests that when the term structure steepens, the forward densities become more dispersed and the effect is more pronounced at both tails. One exception to this result is that the slope coefficients become negative and statistically insignificant at the right tail at 5-year maturity. The coefficients of the volatility factor are generally positive (negative) for the left (right) half of the distribution. Although the volatility coefficients start to turn positive at the right tail of the distribution, they

Fig. 86.2 (continued) 3-month LIBOR forward rates. The volatility factor is defined as the first principal component of EGARCH-filtered spot volatilities and has been normalized to a mean that equals one. The two levels of the slope factor correspond to flat and steep term structures, while the two levels of the volatility correspond to low and high levels of volatility

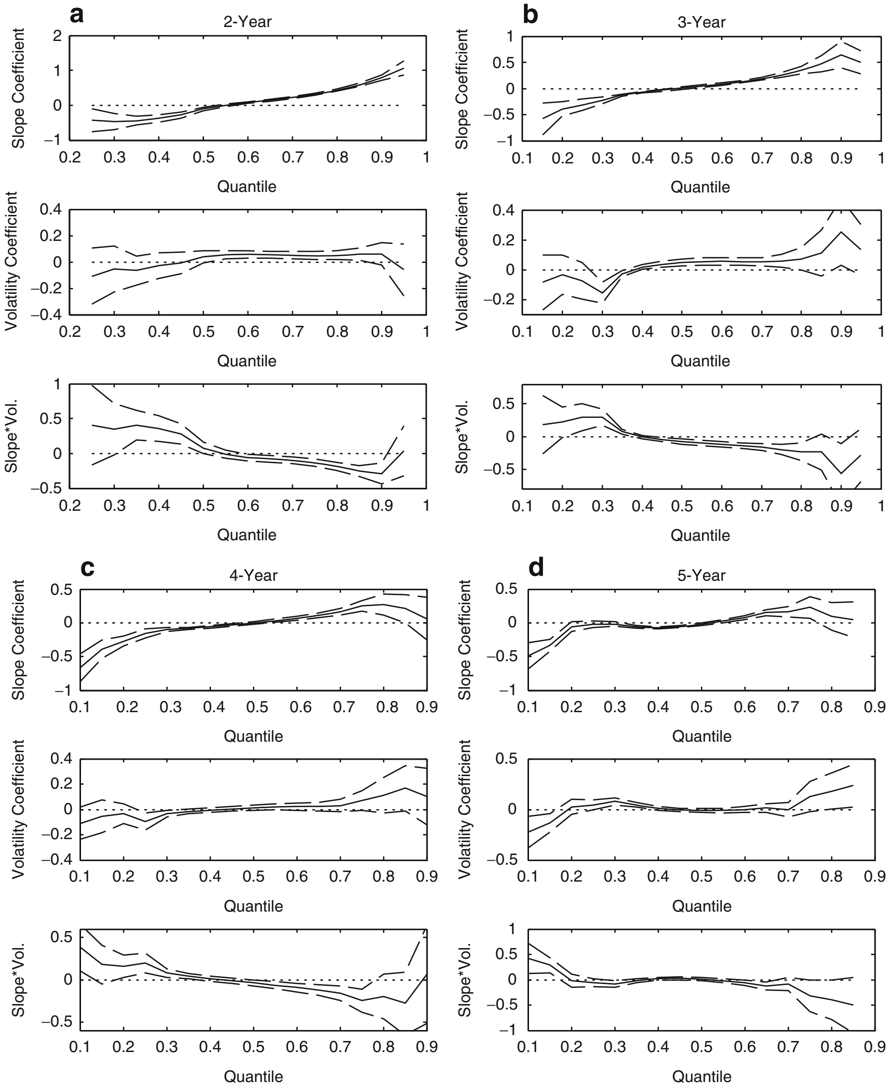


Fig. 86.4 Impacts of the slope and volatility factors (with their interaction term) on LIBOR forward densities. This figure reports regression coefficients of different quantiles of the forward densities at 2-, 3-, 4-, and 5-year maturities on the slope and volatility factors of LIBOR rates and their interaction term in Eq. 86.27

are not statistically significant. These results suggest that higher volatility leads to more compact forward densities, a result that is generally consistent with that in Fig. 86.2.

In Fig. 86.4, although the slope coefficients exhibit similar patterns as that in Fig. 86.3, the interaction term changes the volatility coefficients quite significantly.

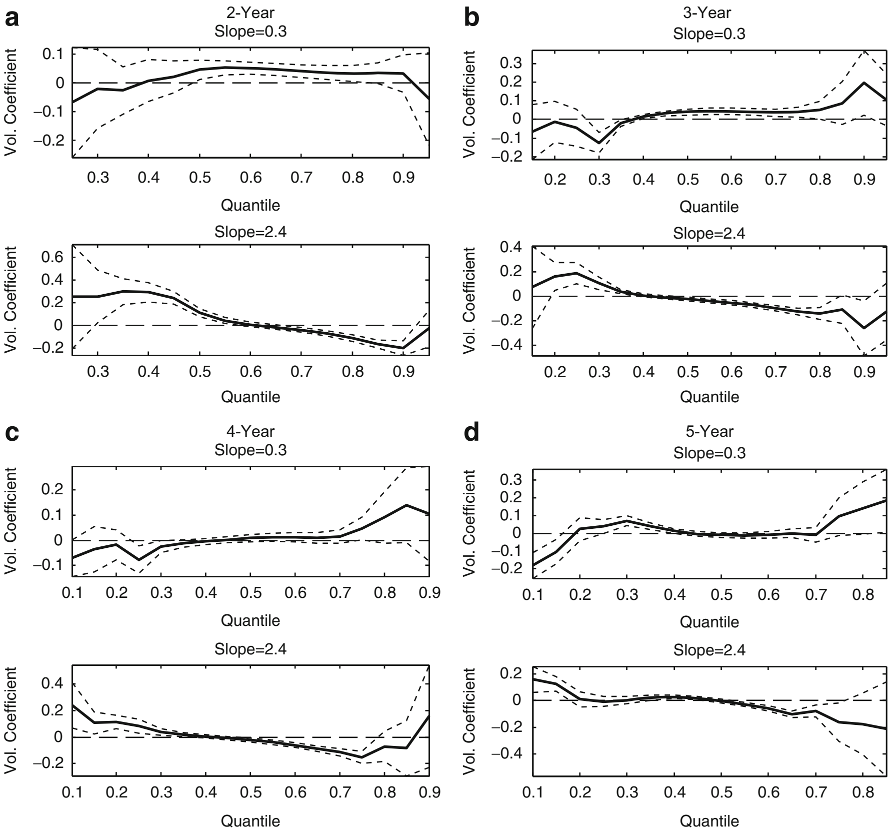


Fig. 86.5 Nonlinear dependence of LIBOR forward densities on the volatility factor of LIBOR rates. This figure presents regression coefficients of quantiles of LIBOR forward densities on the volatility factor at different levels of the slope factor. The two levels of the slope factor represent flat and steep term structures

The volatility coefficients become largely insignificant and exhibit quite different patterns than those in Fig. 86.3. For example, the volatility coefficients at 2- and 3-year maturities are largely constant across different quantiles. At 4- and 5-year maturities, they even become negative (positive) for the left (right) half of the distribution. On the other hand, the coefficients of the interaction term exhibit similar patterns as that of the volatility coefficients in Fig. 86.3. These results suggest that the impacts of volatility on the forward densities depend on the slope of the term structure.

Figure 86.5 presents the volatility coefficients at different levels of the slope factor (i.e., $\hat{b}_{2x} + \hat{b}_{3x} \cdot s_i$, where $s_i = 0.3$ or 2.4). We see clearly that the impact of volatility on the forward densities depends significantly on the slope factor. With a flat term structure, the volatility coefficients generally increase with the quantiles, especially at 3-, 4-, and 5-year maturities. The volatility coefficients are generally negative (positive) for the left (right) tail of the distribution, although not all of them are statistically significant. However, with a steep term structure, the volatility coefficients are generally positive (negative) for the left (right) half of the distribution for most maturities. Therefore, if the current volatility is high and the term structure is flat (steep), then volatility is likely to increase (decline) in the future. We observe flat term structure during early part of our sample when the Fed has raised interest rate to slow down the economy. It could be that the market was more uncertain about future state of the economy because it felt that recession was imminent. On the other hand, we observe steep term structure after the Internet bubble bursted and the Fed has aggressively cut interest rate. It could be that the market felt that the worst was over and thus was less uncertain about future state of the economy.

Our nonparametric analysis reveals important nonlinear dependence of the forward densities on both the slope and volatility factors of LIBOR rates. These results have important implications for one of the most important and controversial topics in the current term structure literature, namely, the USV puzzle. While existing studies on USV mainly rely on parametric methods, our results provide nonparametric evidence on the importance of USV: even after controlling for important bond market factors, such as level and slope, the volatility factor still significantly affects the forward densities of LIBOR rates. Even though many existing term structure models have modeled volatility as a mean-reverting process, our results show that the speed of mean reversion of volatility is nonlinear and depends on the slope of the term structure.

Some recent studies have documented interactions between activities in mortgage and interest rate derivatives markets. For example, in an interesting study, Duarte (2008) shows that ATM swaption implied volatilities are highly correlated with prepayment activities in the mortgage markets. Duarte (2008) extends the string model of Longstaff et al. (2001) by allowing the volatility of LIBOR rates to be a function of the prepayment speed in the mortgage markets. He shows that the new model has much smaller pricing errors for ATM swaptions than the original model with a constant volatility or a CEV model. Jarrow et al. (2007) also show that although their LIBOR model with stochastic volatility and jumps can price caps across moneyness reasonably well, the model pricing errors are unusually large during a few episodes with high prepayments in MBS. These findings suggest that if activities in the mortgage markets, notably the hedging activities of government-sponsored enterprises, such as Fannie Mae and Freddie Mac, affect the supply/demand of interest rate derivatives, then this source of risk may not be fully spanned by the factors driving the evolution of the term structure.[23]

In this section, we provide nonparametric evidence on the impact of mortgage activities on LIBOR forward densities. Our analysis extends Duarte (2008) in several important dimensions. First, by considering caps across moneyness, we examine the impacts of mortgage activities on the entire forward densities. Second, by explicitly allowing LIBOR forward densities to depend on the slope and volatility factors of

[23]See Jaffee (2003) and Duarte (2006) for excellent discussions on the use of interest rate derivatives by Fannie Mae and Freddie Mac in hedging interest rate risks.

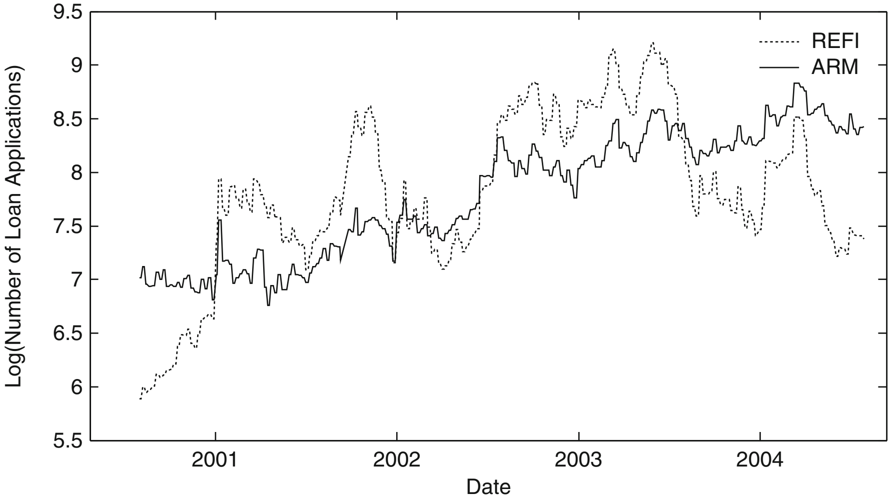


Fig. 86.6 Mortgage Bankers Association of America (MBAA) weekly refinancing and ARMs indexes. This figure reports the logs of the refinance and ARMs indexes obtained by weekly surveys at the Mortgage Bankers Association of America (MBAA)

LIBOR rates, we examine whether prepayment still has incremental contributions in explaining interest rate option prices in the presence of these two factors.[24] Finally, in addition to prepayment activities, we also examine the impacts of ARMs origination on the forward densities. Implicit in any ARM is an interest rate cap, which caps the mortgage rate at a certain level. Since lenders of ARMs implicitly sell a cap to the borrower, they might have incentives to hedge such exposures.[25]

Our measures of prepayment and ARMs activities are the weekly refinancing and ARMs indexes based on the weekly surveys conducted by MBAA, respectively. The two indexes, as plotted in Fig. 86.6, tend to be positively correlated with each other. There is an upward trend in ARMs activities during our sample period, which is consistent with what happened in the housing market in the past few years.

To examine the impacts of mortgage activities on LIBOR forward densities, we repeat the above regressions by including two additional explanatory variables that measure refinance and ARMs activities. Specifically, we refer to the top 20 % of the observations of the refinance (ARMs) index as the high prepayment (ARMs) group. After obtaining a nonparametric forward density at a particular level of the two conditioning variables, we define two new variables "Refi" and "ARMs," which

[24]While the slope factor can have nontrivial impact on prepayment behavior, the volatility factor is crucial for pricing interest rate options.

[25]We thank the referee for the suggestion of examining the effects of ARMs origination on the forward densities.

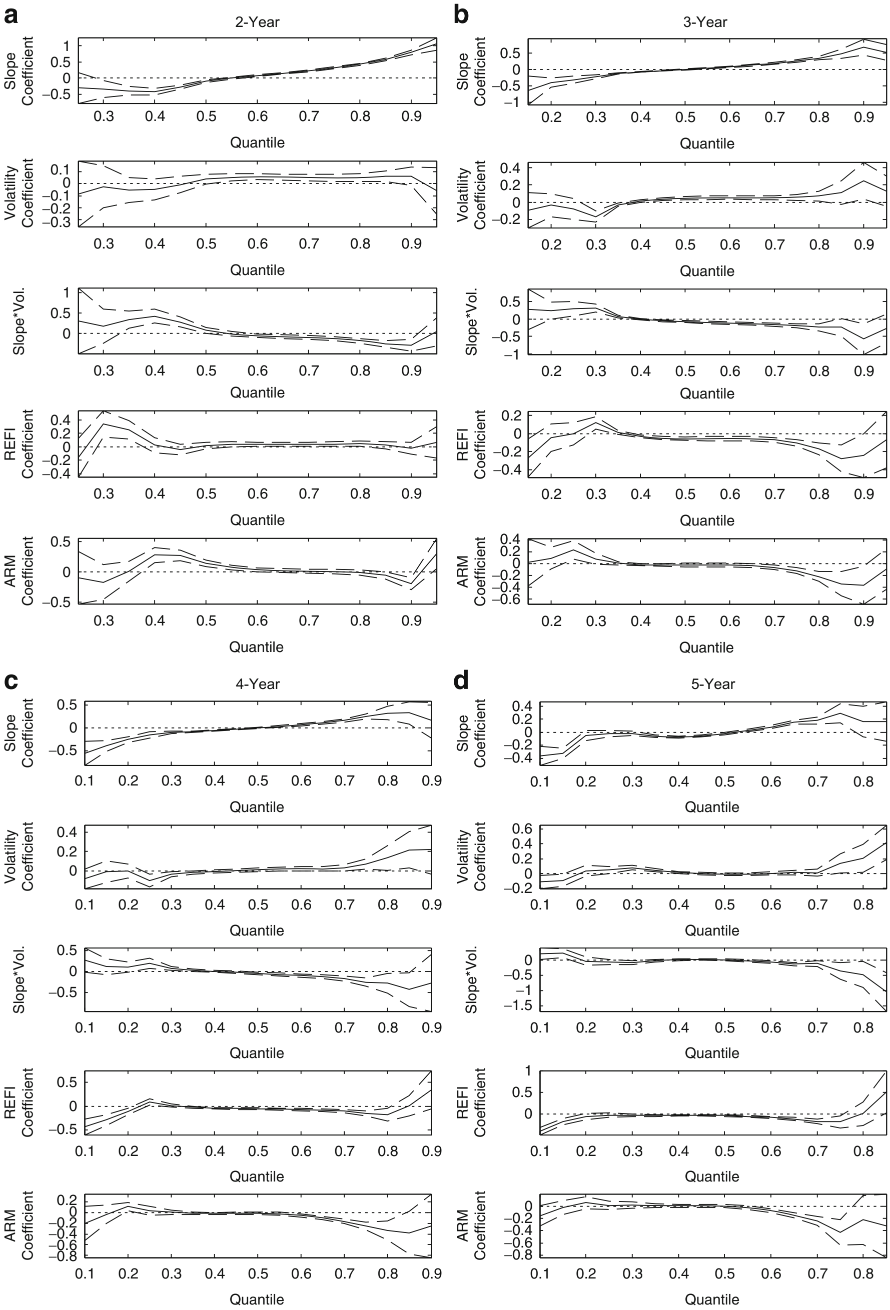


Fig. 86.7 Impacts of refinance and ARMs activities on LIBOR forward densities. In this figure, for each quantile of LIBOR forward densities at 2-, 3-, 4-, and 5-year maturities, we report regression coefficients of the quantile on (i) the slope and volatility factors and their interaction term as in Eq. 86.27 and (ii) refinance and ARMs activities

measure the percentages of observations used in estimating the forward density that belong to the high prepayment and ARMs groups, respectively. These two variables allow us to test whether the forward densities behave differently when prepayment/ARMs activities are high. To control for potential collinearity among the explanatory variables, we have orthogonalized any new explanatory variable with respect to existing ones.

Figure 86.7 contains the new regression results with "Refi" and "ARMs" for the four maturities. The coefficients of the slope, volatility, and the interaction term exhibit similar patterns as that in Fig. 86.4.[26]

The strongest impacts of ARMs on the forward densities occur at 2-year maturity, as shown in Panel A of Fig. 86.7. Therefore, high ARMs origination shifts the median and the right tail of the forward densities at 2-year maturity toward the right. This finding is consistent with the notion that hedging demands from ARMs lenders for the cap they have shorted might increase the price of OTM caps. One possible reason that the effects of ARMs are more pronounced at 2-year maturity than at 3-, 4-, and 5-year maturities is that most ARMs get reset within the first 2 years.

While high ARMs activities shift the forward density at 2-year maturity to the right, high refinance activities shift the forward densities at 3-, 4-, and 5-year maturities to the left. We see that the coefficients of Refi at the left tail are significantly negative. While the coefficients also are significantly negative for the middle of the distribution (40–70 % quantiles), the magnitude of the coefficients is much smaller. These can be seen in Panels B, C, and D of Fig. 86.7. Therefore, high prepayment activities lead to much more negatively skewed forward densities. This result is consistent with the notion that investors in MBS might demand OTM floors to hedge their potential losses from prepayments. The coefficients of Refi are more significant at 4- and 5-year maturities because the duration of most of MBS are close to 5 years.

Our results confirm and extend the findings of Duarte (2008) by showing that mortgage activities affect the entire forward density and consequently the pricing of interest rate options across moneyness. While prepayment activities affect the left tail of the forward densities at intermediate maturities, ARMs activities affect the right tail of the forward densities at short maturity. Our findings hold even after controlling for the slope and volatility factors and suggest that part of the USV factors could be driven by activities in the mortgage markets.

86.5 Conclusion

The unspanned stochastic volatility puzzle is one of the most important topics in the current term structure modeling. Similar to the stochastic volatility in the equity options literature, the existence of USV challenges the benchmark in the current

[26]In results not reported, we find that the nonlinear dependence of the forward densities on the volatility factor remains the same as well.

term structure literature, the dynamic term structure models. But it also in part explains why the practitioners generally apply the HJM type of models for interest rate derivatives, where sometimes the models are applied in an inconsistent manner across securities. Unlike the stock options models where the underlying stock price usually follows a univariate process, it is more challenging to argue that the stochastic volatilities of yields are not spanned by the existing yield curve factors. We in this paper review the current literature, which is mostly in support of the USV using either bonds data or both bonds and derivatives data. We present the results in Li and Zhao (2006) that the DTSMs have serious difficulty in hedging against the interest rate caps. We also present the results from Li and Zhao (2009) where they show nonparametrically both the actual volatility of interest rates and the liquidity component of the implied volatility affect the derivative prices after controlling for the yield curve factors. This paper also presents the model developed in Jarrow et al. (2007), which is quite rich parametrically to capture a spectrum of derivative prices. We can expect that the USV will have the similar effect on interest rate derivatives as the stochastic volatility on the equity options literature with many more issues to be addressed in the future.

Appendix 1: The Derivation for QTSMs

To guarantee the stationarity of the state variables, we assume that ξ permits the following eigenvalue decomposition:

$$\xi = U\Lambda U^{-1}$$

where Λ is the diagonal matrix of the eigenvalues that take negative values, $\Lambda \equiv \text{diag}[\lambda_i]_N$, and U is the matrix of the eigenvectors of ξ, $U \equiv [u_1 u_2 \ldots u_N]$. The conditional distribution of the state variables X_t is multivariate Gaussian with conditional mean

$$E[X_{t+\Delta t}|X_t] = U\Lambda^{-1}[\Phi - I_N]U^{-1}\mu + U\Lambda^{-1}[\Phi - I_N]U^{-1}X_t \tag{44.47}$$

and conditional variance

$$\text{var}[X_{t+\Delta t}|X_t] = U\Theta U' \tag{44.48}$$

where Φ is a diagonal matrix with elements $\exp(\lambda_i \Delta t)$ for $i = 1, \ldots, N$, Θ is an N-by-N matrix with elements

$$\left[\frac{v_{ij}}{\lambda_i + \lambda_j}\left(e^{\Delta t\left(\lambda_i + \lambda_j\right)} - 1\right)\right],$$

where $[v_{ij}]_{N \times N} = U^{-1} \sum \sum' U'^{\,-1}$.

With the specification of market price of risk, we can relate the risk-neutral measure Q to the physical one P as follows:

$$E\left[\frac{dQ}{dP}|\mathcal{F}_t\right] = \exp\left[-\int_0^t \zeta(X_s)'dW_s - \frac{1}{2}\int_0^t \zeta(X_s)'\zeta(X_s)ds\right], \quad \text{for } t \le T. \tag{44.49}$$

Applying Girsanov's theorem, we obtain the risk-neutral dynamics of the state variables

$$dX_t = [\delta + \gamma X_t]dt + \sum dW_t{}^Q$$

where $\delta = \mu - \sum \eta_0, \gamma = \xi - \sum \eta_1$, and $W_t{}^Q$ is an N-dimensional standard Brownian motion under measure Q.

Under the above assumptions, a large class of fixed-income securities can be priced in (essentially) closed form (see Leippold and Wu 2002). We discuss the pricing of zero-coupon bonds below and the pricing of caps. Let $V(t, \tau)$ be the time-t value of a zero-coupon bond that pays 1 dollar at time $T(\tau = T - t)$. In the absence of arbitrage, the discounted value process $\exp(-\int_0^t r(X_s)ds)V(t,\tau)$ is a Q-martingale. Thus, the value function must satisfy the fundamental PDE, which requires the bond's instantaneous return equals the risk-free rate:

$$\frac{1}{2}tr\left(\sum\sum{}'\frac{\partial^2 V(t,\tau)}{\partial X_t \partial X_t'}\right) + \frac{\partial V(t,\tau)}{\partial X_t'}(\delta + \gamma X_t) + \frac{\partial V(t,\tau)}{\partial t} = r_t V(t,\tau) \tag{44.50}$$

with the terminal condition $V(t, 0) = 1$. The solution takes the form

$$V(t,\tau) = \exp\left[-X_t'A(\tau)X_t - b(\tau)'X_t - c(\tau)\right],$$

where $A(\tau)$, $b(\tau)$, and $c(\tau)$ satisfy the following system of ordinary differential equations (ODEs):

$$\begin{aligned}
\frac{\partial A(\tau)}{\partial \tau} &= \psi + A(\tau)\gamma + \gamma'A(\tau) - 2A(\tau)\sum\sum{}'A(\tau);\\
\frac{\partial b(\tau)}{\partial \tau} &= \beta + 2A(\tau)\delta + \gamma'b(\tau) - 2A(\tau)\sum\sum{}'b(\tau);\\
\frac{\partial c(\tau)}{\partial \tau} &= \alpha + b(\tau)'\delta - \frac{1}{2}b(\tau)'\sum\sum{}'b(\tau) + tr\left[\sum\sum{}'A(\tau)\right];\\
&\text{with } A(0) = 0_{N\times N;b(0)} = 0_{N;c(0)} = 0.
\end{aligned}$$

Consequently, the yield to maturity, $y(t, \tau)$, is a quadratic function of the state variables:

$$y(t,\tau) = \frac{1}{\tau}\left[X_t'A(\tau)X_t + b(\tau)'X_t + c(\tau)\right].$$

In contrast, in the ATSMs the yields are linear in the state variables, and therefore the correlations among the yields are solely determined by the correlations of the state variables. Although the state variables in the QTSMs follow multivariate Gaussian process, the quadratic form of the yields helps to model the time-varying volatility and correlation of bond yields.

Leippold and Wu (2002) show that a large class of fixed-income securities can be priced in closed form in the QTSMs using the transform analysis of Duffie et al. (2000). They show that the time-t value of a contract that has an exponential quadratic payoff structure at terminal time T, i.e.,

$$\exp(-q(X_T)) = \exp\left(-X_T^{'}\overline{A}X_T - \overline{b}^{'}X_T - \overline{c}\right)$$

has the following form:

$$\begin{aligned}\Psi(q, X_t, t, T) &= E_Q\left(e^{-\int_t^T r(X_s)ds} e^{-q(X_T)}|\mathcal{F}_t\right). \\ &= \exp\left[-X_t A(T-t)X_t - b(T-t)'X_t - c(T-t)\right]. \end{aligned} \tag{44.51}$$

where $A(.)$, $b(.)$, and $c(.)$ satisfy the ODEs (86.4), (86.5), and (86.6) with the initial conditions $A(0) = \overline{A}, b(0) = \overline{b}$ and $c(0) = \overline{c}$. The time-t price a call option with payoff $\left(e^{-q(X_T)} - y\right)^+$ at $T = t + \tau$ equals

$$\begin{aligned}C(q, y, X_t, \tau) &= EQ\left(e^{-\int_t^T r(X_s)ds}\left(e^{-q(X_T)} - y\right)^+|\mathcal{F}t\right). \\ &= EQ\left(e^{-\int_t^T r(X_s)ds}\left(e^{-q(X_T)} - y\right)1\{-q(X_T) \geq \ln(y)\}|\mathcal{F}t\right). \\ &= G_{q,q}(-\ln(y), X_t, \tau) - yG_{0,q}(-\ln(y), X_t, \tau),\end{aligned}$$

where $G_{q1,q2}(y, X_t, \tau) = E_Q\left[e^{-\int_t^T r(X_s)ds} e^{-q1(X_T)}1\{q_2(X_T) \leq y\}|\mathcal{F}t\right]$ and can be computed by the inversion formula

$$\begin{aligned}G_{q1,q2}(y, X_t, \tau) = &\frac{\psi(q_1, X_t, t, T)}{2} \\ &- \frac{1}{\pi}\int_0^\infty \frac{e^{ivy}\psi(q_1 + ivq_2) - e^{-ivy}\psi(q_1 - ivq_2)}{iv}dv.\end{aligned} \tag{44.52}$$

Similarly, the price of a put option is

$$P(q, y, \tau, X_t) = yG_{0,-q}(\ln(y), X_t, \tau) - G_{q,-q}(\ln(y), X_t, \tau).$$

We are interested in pricing a cap which is portfolio of European call options on future interest rates with a fixed strike price. For simplicity, we assume the face value is 1 and the strike price is $\overline{r}$. At time 0, let τ, 2τ, . . ., $n\tau$ be the fixed dates for future interest payments. At each fixed date $k\tau$, the $\overline{r}$-capped interest payment is given by $\tau(R((k-1)\tau, k\tau) - \overline{r})^+$, where $R((k-1)\tau, k\tau)$ is the τ-year floating interest rate at time $(k-1)\tau$, defined by

$$\begin{aligned}\frac{1}{1+\tau R((k-1)\tau, k\tau)} &= q((k-1)\tau, k\tau)\\ &= E^Q\left(\exp\left(-\int_{(k-1)\tau}^{k\tau} r(X_s)ds\right)|\mathcal{F}(k-1)\tau\right).\end{aligned}$$

The market value at time 0 of the caplet paying at date $k\tau$ can be expressed as

$$\begin{aligned}Caplet(k) &= E^Q\left[\exp\left(-\int_0^{k\tau} r(X_s)ds\right)\tau(R((k-1)\tau, k\tau) - \overline{r})^+\right]\\ &= (1+\tau\overline{r})E^Q\left[\exp\left(-\int_0^{(k-1)\tau} r(X_s)ds\right)\left(\frac{1}{(1+\tau\overline{r})} - q((k-1)\tau, k\tau)\right)^+\right].\end{aligned}$$

Hence, the pricing of the k-th caplet is equivalent to the pricing of an $(k-1)\tau$ − for − τ put struck at $K = \frac{1}{(1+\tau\overline{r})}$. Therefore,

$$\begin{aligned}Caplet(k) = G_{0,-q\tau}&\left(\ln K, X_{(k-1)\tau}, (k-1)\tau\right)\\ &-\frac{1}{K}G_{q\tau,-q\tau}\left(\ln K, X_{(k-1)\tau}, (k-1)\tau\right).\end{aligned} \tag{44.53}$$

Similarly for the *k-th* floorlet,

$$\begin{aligned}Floorlet(k) = -G_{0,q\tau}&\left(-\ln K, X_{(k-1)\tau}, (k-1)\tau\right)\\ &+\frac{1}{K}G_{q\tau,q\tau}\left(-\ln K, X_{(k-1)\tau}, (k-1)\tau\right).\end{aligned} \tag{44.54}$$

Appendix 2: The Implementation of the Kalman Filter

To implement the extended Kalman filter, we first recast the QTSMs into a state-space representation. Suppose we have a time series of observations of yields of

L zero-coupon bonds with maturities $\Gamma = (\tau_1, \tau_2, \ldots, \tau_L)$. Let Ξ be the set of parameters for QTSMs and $Y_k = f(X_k, \Gamma; \Xi)$ be the vector of the L observed yields at the discrete time points $k\Delta t$, for $k = 1, 2, \ldots, K$, where Δt is the sample interval (1 day in our case). After the following change of variable,

$$Z_k = U^{-1}\left(\xi^{-1}\mu + X_k\right),$$

we have the state equation

$$Z_k = \Phi Z_{k-1} + w_k, \qquad w_k \sim N(0, \Theta)$$

where Φ and Θ are first introduced in Eqs. 86.4 and 86.5, and measurement equation

$$Y_k = d_k + H_k Z_k + v_k, \quad v_k \sim N(0, R^v)$$

where the innovations in the state and measurement equations w_k and v_k follow serially independent Gaussian processes and are independent from each other. The time-varying coefficients of the measurement equation d_k and H_k are determined at the ex ante forecast of the state variables:

$$\begin{aligned} H_k &= \frac{\partial f\left(Uz - \xi^{-1}\mu, \Gamma\right)}{\partial z} |z = Z_k|_{k-1} \\ d_k &= f\left(UZ_{k|k-1} - \xi^{-1}\mu, \Gamma\right) - H_k Z_{k|k-1} + B_k, \end{aligned}$$

where $Z_{k|k-1} = \Phi Z_{k-1}$.

In the QTSMs, the transition density of the state variables is multivariate Gaussian under the physical measure. Thus, the transition equation in the Kalman filter is exact. The only source of approximation error is due to the linearization of the quadratic measurement equation. As our estimation uses daily data, the approximation error, which is proportional to 1-day-ahead forecast error, is likely to be minor. Furthermore, we can minimize the approximation error by introducing the correction term B_k.[27] The Kalman filter starts with the initial state variable $Z_0 = E(Z_0)$ and the variance-covariance matrix P_0^Z:

$$P_0^Z = E\left[(Z_0 - E(Z_0))(Z_0 - E(Z_0))'\right].$$

These unconditional mean and variance have closed-form expressions that can be derived using Eqs. 86.4 and 86.5 by letting Δt goes to infinity. Given the set of

[27]The differences between parameter estimates with and without the correction term are very small, and we report those estimates with the correction term B_k.

filtering parameters, $\{\Xi, R^v\}$, we can write down the log-likelihood of observations based on the Kalman filter K:

$$
\begin{aligned}
\log \mathscr{L}(Y;\Xi) &= \sum_{k=1}^{K} \log f(Y_k; Y_{k-1}, \{\Xi, R^v\}) \\
&= -\frac{LK}{2}\log(2\pi) - \frac{1}{2}\sum_{k=1}^{K}\log\left|P^Y_{k|k-1}\right| \\
&\quad -\frac{1}{2}\sum_{k=1}^{K}\left[\left(Y_k - Y_{k|k-1}\right)'\left(P^Y_{k|k-1}\right)^{-1}\left(Y_k - Y_{k|k-1}\right)\right]
\end{aligned}
$$

with $\mathcal{Y}_{k-1}$ is the information set at time (k–1) Δt, and is the time $P^Y_{k|k-1}$ (k–1) Δt, conditional variance of Y_k

$$
\begin{aligned}
P^Y_{k|k-1} &= H_k P^Z_{k|k-1} H_k' + R^v; \\
P^Z_{k|k-1} &= \Phi P^Z_{k-1} \Phi' + \Theta.
\end{aligned}
$$

Appendix 3: Derivation of the Characteristic Function

The solution to the characteristic function of log ($L_k(T_k)$),

$$
\psi(u_0, Y_t, t, T_k) = \exp\left[a(s) + u_0 \log(L_k(t)) + B(s)'V_t\right],
$$

$\alpha(s)$ and $B(s)$, $0 \le s \le T_k$ satisfy the following system of Riccati equations:

$$
\begin{aligned}
\frac{dB_j(s)}{ds} &= -\kappa_j^{k+1} B_j(s) + \frac{1}{2}B_j^2(s)\xi_j^2 + \frac{1}{2}\left[u_0^2 - u_0\right]U^2_{s,j}, \qquad 1 \le j \le N, \\
\frac{da(s)}{ds} &= \sum_{j=1}^{N}\kappa_j^{k+1}\theta_j^{k+1}B_j(s) + \lambda_J[\Gamma(u_0) - 1 - u_0(\Gamma(1) - 1)],
\end{aligned}
$$

where the function Γ is

$$
\Gamma(c) = \exp\left(\mu_J^{k+1}c + \frac{1}{2}\sigma_J^2 c^2\right).
$$

The initial conditions are $B(0) = 0_{N\times 1}$, $\alpha(0) = 0$, and κ_j^{k+1} and θ_j^{k+1} are the parameters of $V_j(t)$ process under $\mathbb{Q}^{k+1}$.

For any $l < k$, given that $B(T_l) = B_0$ and $a(T_l) = a_o$, we have the closed-form solutions for $B(T_{l+1})$ and $a(T_{l+1})$. Define constants $p = \left[u_0^2 - u_0\right] U_{s,j}^2, q = \sqrt{\left(\kappa_j^{k+1}\right)^2 + p\xi_j^2}, c = \frac{p}{q-\kappa_j^{k+1}}$ and $d = \frac{p}{q+\kappa_j^{k+1}}$. Then we have

$$B_j(T_{l+1}) = c - \frac{(c+d)\left(c - B_{j0}\right)}{\left(d + B_{j0}\right)\exp(-q\delta) + \left(c - B_{j0}\right)}, \qquad 1 \leq j \leq N,$$

$$a(T_{l+1}) = \alpha_0 - \sum_{j=1}^{N} \left[\kappa_j^{k+1}\theta_j^{k+1}\left(d\delta + \frac{2}{\xi_j^2}\ln\left(\frac{\left(d + B_{j0}\right)\exp(-q\delta) + \left(c - B_{j0}\right)}{c+d}\right)\right)\right] + \lambda_J\delta[\Gamma(u_0) - 1 - u_0(\Gamma(1) - 1)],$$

if $p \neq 0$ and $B_j(T_{l+1}) = B_{j0}$, $a(T_{l+1}) = a_0$ otherwise. $B(T_k)$ and $a(T_k)$ can be computed via iteration.

References

Ahn, D. H., & Gao, B. (1999). A parametric nonlinear model of term structure dynamics. *The Review of Financial Studies, 12*(4), 721–762.

Ahn, D.-H., Dittmar, R. F., Ronald Gallant, A., & Gao, B. (2003). Purebred or hybrid?: Reproducing the volatility in term structure dynamics. *Journal of Econometrics, 116*, 147–180.

Ahn, D.-H., Dittmar, R. F., & Ronald Gallant, A. (2002). Quadratic term structure models: Theory and evidence. *Review of Financial Studies, 15*, 243–288.

Ai't-Sahalia, Y., & Duarte, J. (2003). Nonparametric option pricing under shape restrictions. *Journal of Econometrics, 116*, 9–47.

Ai't-Sahalia, Y., & Lo, A. W. (1998). Nonparametric estimation of state-price densities implicit in financial asset prices. *Journal of Finance, American Finance Association, 53*(2), 499–547.

Ait-Sahalia, Y., & Lo, A. W. (2000). Nonparametric risk management and implied risk aversion. *Journal of Econometrics, 94*, 9–51.

Andersen, T. G., & Benzoni, L. (2006). *Do bonds span volatility risk in the U.S. Treasury market? A specification test of affine term structure models* (Working Paper). Northwestern University.

Andersen, L., & Brotherton-Ratcliffe, R. (2001). *Extended LIBOR market models with stochastic volatility* (Working Paper). Gen Re Securities.

Andersen, T. G., & Lund, J. (1997). Estimating continuous time stochastic volatility models of the short term interest rate. *Journal of Econometrics, 77*, 343–378.

Bakshi, G., Cao, C., & Chen, Z. (1997). Empirical performance of alternative option pricing models. *Journal of Finance, 52*, 2003–2049.

Ball, C., & Torous, W. (1999). The stochastic volatility of short-term interest rates: Some international evidence. *Journal of Finance, 54*, 2339–2359.

Bansal, R., & Zhou, H. (2002). Term structure of interest rates with regime shifts. *Journal of Finance, 57*, 1997–2043.

Bikbov, R., & Chernov, M. (2004). *Term structure and volatility: Lessons from the Eurodollar Markets* (Working Paper). Columbia: GSB.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy, 81*, 637–654.

Black, F. (1976). The pricing of commodity contracts. *Journal of Financial Economics, 3*, 167–179.

Brace, A., Gatarek, D., & Musiela, M. (1997). The market model of interest rate dynamics. *Mathematical Finance, 7*, 127–155.

Brandt, M. W., & David, C. (1978). Prices of state-contingent claims implicit in option prices. *Journal of Business, 51*, 621–651. "Comparing Multifactor Models of the Term Structure," Working paper. Breeden, Douglas T., Robert H. Litzenberger.

Brenner, R., Harjes, R., & Kroner, K. (1996). Another look at alternative models of short-term interest rate. *Journal of Financial and Quantitative Analysis, 31*, 85–107.

Campbell, J., Lo, A., & MacKinlay, C. (1997). *The econometrics of financial markets*. New Jersey: Princeton University Press.

Casassus, J., Collin-Dufresne, P., & Goldstein, R. S. (2005). Unspanned stochastic volatility and fixed income derivative pricing. *Journal of Banking and Finance, 29*, 2723–2749.

Chacko, G., & Das, S. (2002). Pricing interest rate derivatives: A general approach. *Review of Financial Studies, 15*, 195–241.

Chen, R., & Scott, L. (2001). *Stochastic volatility and jumps in interest rates: An empirical analysis* (Working Paper). Rutgers University.

Chernov, M., & Ghysels, E. (2000). A study towards a unified approach to the joint estimation of objective and risk neutral measures for the purpose of options valuation. *Journal of Financial Economics, 56*, 407–458.

Collin-Dufresne, P., & Goldstein, R. S. (2002). Do bonds span the fixed income markets? Theory and evidence for unspanned stochastic volatility. *Journal of Finance, 57*, 1685–1729.

Collin-Dufresne, P., & Goldstein, R. S. (2003). *Stochastic correlation and the relative pricing of caps and swaptions in a generalized affine framework* (Working paper). Carnegie Mellon University.

Collin-Dufresne, P., Goldstein, R., & Hugonnier, J. (2004). A general formula for valuing defaultable securities. *Econometrica, 72*(5), 1377–1407.

Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985). A theory of the term structure of interest rates. *Econometrica, 53*, 385–407.

Dai, Q., & Singleton, K. (2003). Term structure dynamics in theory and reality. *Review of Financial Studies, 16*, 631–678.

Dai, Q., & Singleton, K. (2002). Expectation puzzles, time-varying risk premia, and affine models of the term structure. *Journal of Financial Economics, 63*(3), 415–441.

Dai, Q., & Singleton, K. (2000). Specification analysis of affine term structure models. *Journal of Finance, 50*(5), 1943–1978.

Das, S. (2002). The surprise element: Jumps in interest rates. *Journal of Econometrics, 106*, 27–65.

Deuskar, P., Gupta, A., & Subrahmanyam, M. (2003). *Liquidity effects and volatility smiles in interest rate option markets* (Working Paper). New York University.

Diebold, F. X., & Mariano, R. S. (1995). Comparing predictive accuracy. *Journal of Business and Economic Statistics, 13*, 253–265.

Duarte, J. (2004). Evaluating an alternative risk preference in affine term structure models. *Review of Financial Studies, 17*, 379–404.

Duarte, J. (2008). Mortgage-backed securities refinancing and the arbitrage in the swaption market. *Review of Financial Studies, 21*, 1689–1731.

Duffie, D. (2002). *Dynamic asset pricing theory*. New Jersey: Princeton University Press.

Duffie, D., & Lando, D. (2001). Term structures of credit spreads with incomplete accounting information. *Econometrica, 69*(3), 633–664.

Duffie, D., & Kan, R. (1996). A yield-factor model of interest rates. *Mathematical Finance, 6*, 379–406.

Duffie, D., Pan, J., & Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica, 68*, 1343–1376.

Duffee, G. R. (2002). Term premia and interest rate forecasts in affine models. *Journal of Finance, 57*, 405–443.

Eraker, B. (2003). Do stock prices and volatility jump? Reconciling evidence from spot and option prices. *Journal of Finance, 59*, 1367–1404.

Fan, R., Gupta, A., & Ritchken, P. (2003). Hedging in the possible presence of unspanned stochastic volatility: Evidence from swaption markets. *Journal of Finance, 58*, 2219–2248.

Gallant, A. R. (1987). *Nonlinear statistical models*. New York: Wiley.

Gallant, A. R., & Tauchen, G. (1998). Reprojecting partially observed systems with applications to interest rate diffusions. *Journal of American Statistical Association, 93*, 10–24.

Glasserman, P., & Kou, S. (2003). The term structure of simple forward rates with jump risk. *Mathematical Finance, 13*(3), 383–410.

Goldstein, R. S. (2000). The term structure of interest rates as a random field. *Review of Financial Studies, 13*, 365–384.

Gupta, A., & Subrahmanyam, M. (2001). *An examination of the static and dynamic performance of interest rate option pricing models in the dollar cap-floor markets* (Working Paper). Case Western Reserve University.

Han, B. (2007). Stochastic volatilities and correlations of bond yields. *Journal of Finance, 62*, 1491–1524.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica, 50*, 1029–1054.

Harris, L. (2003). *Trading and exchanges: Market microstructure for practitioners*. Oxford: Oxford University Press.

Heath, D., Jarrow, R., & Morton, A. (1992). Bond pricing and the term structure of interest rates: A new methodology. *Econometrica, 60*, 77–105.

Heidari, M., & Wu, L. (2003). Are interest rate derivatives spanned by the term structure of interest rates? *Journal of Fixed Income, 13*, 75–86.

Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies, 6*, 327–343.

Hull, J., & White, A. (1987). The pricing of options on assets with stochastic volatilities. *Journal of Finance, 42*, 281–300.

Hull, J., & White, A. (2000). Forward rate volatilities, swap rate volatilities, and the implementation of the LIBOR market models. *Journal of Fixed Income, 10*, 46–62.

Jackwerth, J. C. (2000). Recovering risk aversion from options prices and realized returns. *Review of Financial Studies, 13*, 433–451.

Jaffee, D. (2003). The interest rate risk of Fannie Mae and Freddie Mac. *Journal of Financial Services Research, 24*(1), 5–29.

Jagannathan, R., Kaplin, A., & Sun, S. (2003). An evaluation of multi-factor CIR models using LIBOR, swap rates, and cap and swaption prices. *Journal of Econometrics, 116*, 113–146.

Jarrow, R., Li, H., & Zhao, F. (2007). Interest rate caps "Smile" too! But can the LIBOR market models capture it? *Journal of Finance, 62*, 345–382.

Johannes, M. (2004). The statistical and economic role of jumps in interest rates. *Journal of Finance, 59*, 227–260.

Leippold, M., & Wu, L. (2002). Asset pricing under the quadratic class. *Journal of Financial and Quantitative Analysis, 37*, 271–295.

Li, H., & Zhao, F. (2006). Unspanned stochastic volatility: Evidence from hedging interest rate derivatives. *Journal of Finance, 61*, 341–378.

Li, H., & Zhao, F. (2009). Nonparametric estimation of state-price densities implicit in interest rate cap prices. *Review of Financial Studies, 22*(11), 4335–4376.

Litterman, R., & Scheinkman, J. (1991). Common factors affecting bond returns. *Journal of Fixed Income, 1*, 62–74.

Longstaff, F., & Schwartz, E. (1992). Interest rate volatility and the term structure: A two-factor general equilibrium model. *Journal of Finance, 47*, 1259–1282.

Longstaff, F., Santa-Clara, P., & Schwartz, E. (2001). The relative valuation of caps and swaptions: Theory and evidence. *Journal of Finance, 56*, 2067–2109.

Merton, R. (1973). The theory of rational option pricing. *Bell Journal of Economics and Management Science, 4*, 141–183.

Miltersen, M., Sandmann, K., & Sondermann, D. (1997). Closed-form solutions for term structure derivatives with lognormal interest rates. *Journal of Finance, 52*, 409–430.

Newey, W., & West, K. (1987). A simple, positive semi-definite, heteroskedasticity and autocorrelation consistent covariance matrix. *Econometrica, 55*, 703–708.

Pan, J. (2002). The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of Financial Economics, 63*, 3–50.

Piazzesi, M. (2003). Affine term structure models. In (2010) *Handbook of financial econometrics*. North Holland.

Piazzesi, M. (2005). Bond yields and the Federal Reserve. *Journal of Political Economy, 113*(2), 311–386.

Protter, P. (1990). Stochastic integration and differential equations: A new approach. *Journal of Financial Economics, 64*(3), 341–372. Springer-Verlag. Rosenberg, J, Robert Engle. 2002. "Empirical pricing kernels".

Rosenberg, J. V., & Engle, R. F. (2002). Empirical pricing kernels. *Journal of Financial Economics, 64*(3), 341–372.

Santa-Clara, P., & Sornette, D. (2001). The dynamics of the forward interest rate curve with stochastic string shocks. *Review of Financial Studies, 14*, 2001.

Thompson, S. (2008). Identifying term structure volatility from the LIBOR-swap curve. *Review of Financial Studies, 21*, 819–854.

Trolle, A. B., & Schwartz, E. S. (2009). A general stochastic volatility model for the pricing of interest rate derivatives. *Review of Financial Studies, 22*(5), 2007–2057.

Alternative Equity Valuation Models

87

Hong-Yi Chen, Cheng-Few Lee, and Wei K. Shih

Contents

H.-Y. Chen (✉)
Department of Finance, National Central University, Taoyuan, Taiwan
e-mail: fnhchen@ncu.edu.tw

C.-F. Lee
Department of Finance and Economics, Rutgers Business School, Rutgers, The State University of New Jersey, Piscataway, NJ, USA

Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
e-mail: lee@business.rutgers.edu; cflee@business.rutgers.edu

W.K. Shih
Bates White Economic Consulting, Washington, DC, USA
e-mail: wkshih@gmail.com

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_87,

Abstract

This chapter examines alternative equity valuation models and their ability to forecast future stock prices. Equity valuation models included Ohlson's (1995) Model, Feltham and Ohlson's (1995) Model, and Warren and Shelton's (1971) Model. Five research hypotheses are developed to examine whether different estimation techniques, earnings measures, and combined forecasting methods can improve the ability to predict future stock prices. We find that the simultaneous equation estimation procedure can produce more accurate future stock price forecasts than the traditional single equation estimation method in terms of smaller prediction errors. In addition, the combined forecast method can further reduce the prediction errors by using combination of individual forecasts. Empirical evidence also shows that investors can use comprehensive earnings to more accurately forecast future stock prices in these valuation models.

We use simultaneous equation estimation technique to investigate the stock price forecast ability of Ohlson's Model, Feltham and Ohlson's Model, and Warren and Shelton's (1971) Model. Moreover, we use the combined forecasting methods proposed by Granger and Newbold (1973) and Granger and Ramanathan (1984) to form combined stock price forecasts from individual models. Finally, we examine whether comprehensive earnings can provide incremental price-relevant information beyond net income.

Keywords

Ohlson Model • Feltham and Ohlson Model • Warren and Shelton Model • Equity valuation models • Simultaneous equation estimation • Fundamental analysis • Financial statement analysis • Financial planning and forecasting • Combined forecasting • Comprehensive earnings • Abnormal earnings • Operating earnings • Accounting earnings

87.1 Introduction

In this chapter, we investigate new estimation procedures and combination of forecast methods of alternative equity valuation models, namely, the Ohlson (1995) Model, the Feltham and Ohlson (1995) Model (FO Model henceforth), and the Warren and Shelton (1971) Model (WS Model henceforth). The Ohlson Model and the FO Model are valuation models based on information obtained from income statement and balance sheet, i.e., earnings and book value per share, while the WS Model accounts for the overall operating and financial environments of the firm. This chapter first uses simultaneous equation estimation procedures to estimate the information dynamics in these models and compare their forecast ability of future stock prices. Moreover, we examine the stock price forecasts ability of the Ohlson and FO Model by using other forms of earnings. Given the Ohlson Model and FO Model are derived based on the clean surplus relation (CSR), the *earnings* or *income* in CSR should include all changes in equity during a period except those resulting from investments by owners and distributions to owners. The *earnings* under this concept are closer to *comprehensive income* rather than the bottom-line net income that had been frequently used in previous empirical studies. Given its consistency with the accounting-based valuation theory, we further investigate the stock price forecast ability of the Ohlson and FO Model by using the comprehensive income in the empirical model specification. Finally, we employ forecast combination methods to integrate the individual stock price forecasts from these models and explore possible improvement in terms of producing smaller prediction errors.

Our empirical results suggest that the simultaneous equation estimation of the information dynamics improves the explanatory power of the models. Prior literature in testing the accounting-based valuation model did not consider the feedback effect between the variables of interest. In other words, the traditional linear information dynamics conjectured that the earnings dynamics is an AR(1) process which is determined by past earnings and other value-relevant information such as the analyst earnings forecasts. However, following Tsay et al. (2008), we incorporate the feedback effect between the earnings and value-relevant information variables, such as analyst earnings forecasts, and improve the predictability of future stock prices in terms of better forecast accuracy. We find that simultaneous equation estimation procedure produces smaller mean forecast errors than the single equation estimation procedure by 5.14 % (3.14 %) on average in our sample period for the Ohlson (FO) Model. Moreover, with the addition of other accounting information variables in the WS Model, we find further improvement in forecasting future stock prices. The future stock price forecasts from the WS Model are smaller than those predicted by the Ohlson (FO) Model by 5.26 % (2.23 %) annually in our sample period. By considering the firm's overall operating and financial environment, the model produces smaller prediction errors to the previous two models considering only earnings and book value per share.

Given the future stock price forecasts from these models, we employ forecast combination methods to integrate these individual forecasts in order to generate more accurate future stock price forecasts. We find that the combined future stock price forecasts based on weighted least square regression methods with the geometric weighting scheme produce smaller prediction errors than individual forecasts from the Ohlson, FO, and WS Model. The combined forecasts using the Ohlson (FO) Model with the WS Model generate smaller prediction errors than the individual forecasts by 7.16 % (4.07 %) annually. These results suggest superior accuracy of the combined forecast methods compared to the individual forecasts in terms of future stock price prediction errors. If comprehensive (operating) income is used in the Ohlson (FO) Model, we find that the prediction errors of future stock price forecast can be reduced up to 3.30 % (2.44 %) compared to using net income. Moreover, when comprehensive (operating) income-based Ohlson (FO) Model forecasts are combined with WS Model forecasts, the reduction in pricing errors can reach up to 17.85 % (15.96 %). We find that comprehensive income indeed has incremental price-relevant information beyond bottom-line net income in terms of generating more accurate stock price forecasts in these valuation models. Our results provide new evidence in the value relevance of comprehensive income and shed light on how the issuance of Statement of Financial Accounting Standard No. 130, *Reporting Comprehensive Income*, can help investors better assess the overall performance of the corporation.

The remainder of this chapter is organized as follows. Section 87.2 provides literature review in accounting-based valuation models and the financial planning and forecast models. The theoretical development and empirical implementation of these models are discussed. Section 87.3 presents the sample selection criteria, model specification of the linear information dynamics, and the research hypotheses for the empirical tests. Section 87.4 discusses the empirical results of individual forecasts and combined forecasts for future stock prices from these models. Section 87.5 provides the summary of this chapter.

87.2 Literature Review

In this section, we first review the theoretical development and the empirical assessment of the Ohlson Model and the FO Model. We then review the financial planning and forecasting model developed by Warren and Shelton (1971). We also provide the background and prior academic research on the comprehensive earnings reporting issues. Finally, we will review the combined forecasting methods proposed by Granger and Newbold (1973), Granger and Ramanathan (1984), and Diebold and Pauly (1987).

87.2.1 Ohlson Model (1995) and Feltham-Ohlson Model (1995)

The Ohlson Model provides a theoretical framework linking the valuation to the reported financial statement variables. The traditional dividend discount model states the following relations:

$$P_t = \sum_{\tau=1}^{\infty} R_f^{-\tau}\left(\widetilde{d}_{t+\tau}\right) \tag{87.1}$$

where P_t is the price of the firm's equity at time t, $\widetilde{d}_{t+\tau}$ is the dividends paid at time t, and R_f is the risk-free rate plus one. The restrictive nature of this relation is that Eq. 87.1 does not relate the reported financial statement numbers to firm value. In Eq. 87.1, the value depends on the accounting data that influences the present value of expected future dividends. Since this distribution of wealth eventually converge with the creation of wealth, the Ohlson Model considers how the current value depends on accounting measures of wealth creation process. The Ohlson Model introduced the clean surplus relations (CSR) assumption requiring that income over a period equals net dividends and the change in book value of equity. CSR ensures that all changes in shareholder equity that do not result from transactions with shareholders (such as dividends, share repurchases, or share offerings) are reflected in the income statement. In other words, CSR is an accounting system recognizing that the periodical value created is distinguished from the value distributed.

Let x_t denote the earnings for period $(t-1, t)$, y_t denote the book value at time t, and $x_t^a = x_t - (R_f - 1)y_{t-1}$ denote the abnormal earnings at time t. The clean surplus relations $y_t = y_{t-1} + x_t - d_t$ imply that

$$P_t = y_t + \sum_{\tau=1}^{\infty} R_f^{-\tau} E_t\left[\widetilde{x}_{t+\tau}^a\right] \tag{87.2}$$

the firm's value is equal to its book value adjusted for the present value of expected future abnormal earnings. The variables on the right-hand side of (87.2) are still forecasts, not past realizations. To deal with this problem, the Ohlson Model introduced the information dynamics to link the value to the contemporaneous accounting data. Assume $\left\{\widetilde{x}_t^a\right\}_{\tau \geq 1}$ follows the stochastic process

$$\begin{aligned} \widetilde{x}_{t+1}^a &= \omega x_t^a + v_t + \widetilde{\varepsilon}_{1,t+1} \\ \widetilde{v}_{t+1} &= \gamma v_t + \widetilde{\varepsilon}_{2,t+1} \end{aligned} \tag{87.3}$$

where v_t is value-relevant information other than abnormal earnings and $0 \leq \omega$, $\gamma \leq 1$. Based on Eqs. 87.2 and 87.3, the Ohlson Model demonstrated that the value of the equity is a function of contemporaneous accounting variables as follows:

$$P_t = y_t + \hat{\alpha}_1 x_t^a + \hat{\alpha}_2 v_t \tag{87.4}$$

where $\hat{\alpha}_1 = \hat{\omega}/\left(R_f - \hat{\omega}\right)$ and $\hat{\alpha}_2 = R_f/\left(R_f - \hat{\omega}\right)\left(R_f - \hat{\gamma}\right)$. Or equivalently,

$$P_t = \kappa(\varphi x_t - d_t) + (1-\kappa)y_t + \alpha_2 v_t \tag{87.5}$$

where $\kappa = \left(R_f - 1\right)\hat{\omega}/\left(R_f - \hat{\omega}\right)$ and $\varphi = R_f/(R_f - 1)$. Equations 87.4 and 87.5 imply that the market value of the equity is equal to the book value adjusted for (i) the current profitability as measured by abnormal earnings and (ii) other information that modifies the prediction of future profitability. One major limitation of the Ohlson Model is that it assumed unbiased accounting. In Eq. 87.3, since both abnormal earnings and other information follow an AR(1) process, over time their averages are zero and thus the average abnormal earnings are zero as well. If given biased (conservative) accounting, the average abnormal earnings will be nonzero. Consequently, the future growth in book value will become an important factor. This motivated Feltham and Ohlson (1995) to introduce additional dynamics to deal with this issue.

The FO Model analyzes how firm value relates to the accounting information that discloses the results from both operating and financial activities. For the financial activities, there are relatively perfect markets and the accounting measures for book value and market value of these assets are reasonably close. However for the operating assets, accrual accounting usually results in difference between the book value and the market value of these assets since they are not traded in the market. Accrual accounting for the operating assets consequently results in discrepancy between their book value and market value and thus influences the goodwill of the firm. Similar to the Ohlson Model, the information dynamics in the FO Model is

$$\begin{aligned}\widetilde{ox}^a_{t+1} &= \omega_{10} + \omega_{11} ox^a_t + \omega_{12} oa_t + \omega_{13} v_{1t} + \widetilde{\varepsilon}_{1t+1}\\ \widetilde{oa}_{t+1} &= \omega_{20} + \omega_{22} ox^a_t + \omega_{24} v_{2t} + \widetilde{\varepsilon}_{2t+1}\\ \widetilde{v}_{1t+1} &= \omega_{30} + \omega_{33} v_{1t} + \widetilde{\varepsilon}_{3t+1}\\ \widetilde{v}_{2t+1} &= \omega_{40} + \omega_{44} v_{2t} + \widetilde{\varepsilon}_{4t+1}\end{aligned} \tag{87.6}$$

where ox^a_t is the abnormal operating earnings, oa_t is the operating assets, v_{1t} and v_{2t} are the other value-relevant information variables for firm at time t, respectively. The derived implied pricing function is

$$P_t = y_t + \hat{\lambda}_0 + \hat{\lambda}_1 ox^a_t + \hat{\lambda}_2 oa_t + \hat{\lambda}_3 v_{1t} + \hat{\lambda}_4 v_{2t} \tag{87.7}$$

where

$$\begin{aligned}\hat{\lambda}_0 &= \frac{(1+r)\begin{bmatrix}\hat{\omega}_{10}(1+r-\hat{\omega}_{22})(1+r-\hat{\omega}_{33})(1+r-\hat{\omega}_{44})\\ +\hat{\omega}_{12}\hat{\omega}_{20}(1+r-\hat{\omega}_{33}) + \hat{\omega}_{13}\hat{\omega}_{30}(1+r-\hat{\omega}_{22})\\ +\hat{\omega}_{14}\hat{\omega}_{40}(1+r-\hat{\omega}_{44})\end{bmatrix}}{r(1+r-\hat{\omega}_{11})(1+r-\hat{\omega}_{22})(1+r-\hat{\omega}_{33})(1+r-\hat{\omega}_{44})}\\ \hat{\lambda}_1 &= \frac{\hat{\omega}_{11}}{r(1+r-\hat{\omega}_{11})}\\ \hat{\lambda}_2 &= \frac{(1+r)\hat{\omega}_{12}}{(1+r-\hat{\omega}_{11})(1+r-\hat{\omega}_{22})}\\ \hat{\lambda}_3 &= \frac{(1+r)\hat{\omega}_{13}}{(1+r-\hat{\omega}_{11})(1+r-\hat{\omega}_{33})}\\ \hat{\lambda}_4 &= \frac{(1+r)\hat{\omega}_{14}}{(1+r-\hat{\omega}_{11})(1+r-\hat{\omega}_{44})}\end{aligned} \tag{87.8}$$

Or equivalently,

$$P_t = k(\phi x_t - d_t) + (1-\kappa) y_t + \hat{\alpha}_2 oa_t + \hat{\lambda}_3 v_{1t} + \hat{\lambda}_4 v_{2t} \tag{87.9}$$

where $\kappa = \left(R_f - 1\right)\hat{\omega}_{11}/\left(R_f - \hat{\omega}_{11}\right)$ and $\phi = R_f/(R_f - 1)$. The implied valuation function in Eqs. 87.7 and 87.9 is a weighted average of the firm's operating earnings, the firm's book value, and the other value-relevant information with an adjustment for the understatement of the operating assets resulting from accrual accounting. The major contribution of the FO Model is that it considered the accounting conservatism in the equity valuation.

A complementary study by Lehman (1993) also examined the similar problem in the computation share value. Essentially this chapter aimed at identifying cash flows more fundamental than dividends in the computation of the present values. Specifically, the cash flows that were the focus of this chapter have the following properties: (i) Similar to dividends, current prices of the shares are the risk-adjusted present value of the cash flows. (ii) Unlike the dividends, these cash flows have to be invariant to the changes in dividend policy. Lehman (1993) provided three building blocks to identify cash flows that are more fundamental than dividends and with the property that they are invariant to the changes in the dividend policy. The first building block is to transform the present value computation using the future dividends into a relation using future values of arbitrary cash flows. The second building block is the restriction of the changes in the dividend policy to be zero net present value alteration of future dividend stream that leave the risk premium of the firm unaffected. In other words, this chapter made the assumption that only the changes in investment and financial policies would alter the riskiness of the firm but not the changes in the dividend policy. The final building block is the translation of economic earnings into accounting earnings that can be obtained from the financial statement. Essentially this assumption states that the accounting earnings used in computing the present value of the shares have to include all relevant capital gains and losses from investment associated with dividend policy.

Lehman (1993) is different from the Ohlson Model and FO Model in determining the user cost of capital. Lehman (1993) generalizes the assumption that expected earnings are linearly related to lagged earnings, dividends, and book value. Although it also imposes CSR, the earnings are not confined to accounting earnings. Moreover, Lehman (1993) also differs from the FO Model in the specification of the book value of equity. It uses an aggregate book value approach which does not separate the effects of net financial assets from the net operating assets. In the absence of conservative accounting, there will be no differences between two valuation methods given that book value of equity is equal to the market value of equity. However, given that the current GAAP is biased towards conservative accounting, the valuation implications on equity shares are different between the two models. Finally, Lehman (1993) relaxes Ohlson/FO assumption that changes in dividend payments change next period's earnings by one plus the risk-free discount rate by using a general risk-adjusted rate to discount dividends.

The line of research in empirically testing the Ohlson Model and FO Model has been growing large since its introduction. Previous empirical literature focused on either the value-relevant information variables (Abarbanell and Bushee 1997; Myers 1999; Dechow et al. 1999; Liu et al. 2002; Begley and Feltham 2002) or the dynamics of the earnings process (Morel 2001; Callen and Morel 2001). However, none of them documented empirical validity of the Ohlson Model. Callen and Segal (2005) showed that the nested Ohlson Model is rejected in favor of the FO Model but it did not improve the predictability power of the future stock prices. Based on these previous studies, we will examine whether the simultaneous equation approach in estimating the information dynamics can improve the predictability power of the Ohlson and FO Model.

The other potential cause of the lack of empirical validity of the residual income valuation models is the use of net income as the earnings measure in the linear information dynamics. Given these models are based on the clean surplus relation, the earnings measure should include all changes in equity except those resulting from investments by owners and distributions to owners. Consequently, the determination of stock price in the valuation functions of Eqs. 87.4 and 87.7 cannot be complete unless the earnings include all the value-added activities in the firm (Linsmeier et al. 1997a). Comprehensive income is defined as the change in equity (net assets) of a business enterprise during a period from transactions and other events and circumstances from nonowner sources (Statements of Financial Accounting Concepts No. 6, *Elements of Financial Statements*, 1985). Accounting standards in the USA sometimes allow nonowner changes in asset and liabilities, such as foreign currency translation adjustments, available-for-sale marketable securities adjustments, and minimum required pension liability adjustments, to bypass the income statement. The exclusion of these value-relevant items in financial reporting might mislead the users of financial statement information in assessing the value of the firm.

The debate of whether the firms should report comprehensive income or more streamlined bottom-line income can be traced back to the 1930s (Brief and Peasnell 1996). The supporters of Reporting Comprehensive Income argue that it captures all sources of value creation within a firm. Comprehensive income allows the users of the financial statement information to consider all relevant factors for earnings forecasting and firm value assessment. It also grants less leeway for the managers to engage in earnings management which could potentially distort the actual performance of the firms.[1] On the other hand, the opponents to the comprehensive income reporting point out that it includes many items that are transitory in nature which are not representative of the core operation of the firms. The inclusion of these nonrecurring and extraordinary items hinders income measure to reflect the firm's long-term cash flow prospects (Dhaliwal et al. 1999). It is also argued that these items add noises to reported earnings and make it difficult for forecasts. The users of the financial statements should be able to focus on a single measurement that summarizes all the value-relevant information without much manager's discretion in reporting this figure.[2]

The debate of whether firms in the USA should employ comprehensive income reporting led to the issuance of Statement of Financial Accounting Standard No. 130 (SFAS 130), *Reporting Comprehensive Income*, by Financial Accounting Standards Board. SFAS 130 requires firms to report comprehensive income in their primary financial statements. In the pre- and post-SFAS 130 era, there were many studies examining the value relevance of this requirement to report comprehensive income. Cheng et al. (1996) evaluate the usefulness of different earnings definition and find that the conventional income measures such as operating income and net income provide better explanatory power for residual security returns than comprehensive income. Dhaliwal et al. (1999) examine the value relevance of the major three components in comprehensive income required by SFAS 130. Their results

[1]Previous literatures in advocating comprehensive income reporting include Robinson (1991), Johnson et al. (1995), Beresford et al. (1996), and Smith and Reither (1996).

[2]Previous literatures in advocating current operating performance concepts of reporting income include Kiger and Williams (1977), Black (1993), Brief and Peasnell (1996), and Holthausen and Watts (2001).

suggest that only the marketable securities adjustment item improves the income and returns association. More importantly, they fail to find support to show comprehensive income is a better measurement for firm performance than net income and raise questions about the reporting requirement in SFAS 130. Biddle and Choi (2006) on the other hand find that comprehensive income outperforms other income measure in explaining equity returns and predicting future income and operating cash flows. Chambers et al. (2007) argue that prior studies in the pre-SFAS 130 era suffer from the measurement error problem in calculating the comprehensive earning. These authors show that the aforementioned three major components in the comprehensive income are indeed priced by the market in the post-SFAS 130 era when these items are specifically reported by requirement. In addition to the US findings, there are also many other international studies regarding the value relevance of comprehensive income.[3] The results, however, are mixed because of the different local accounting standards and time period within which the requirements are implemented.

87.2.2 Warren and Shelton Model (1971)

In addition to the accounting-based valuation model discussed above, operational financial planning models can also be used to forecast stock prices. One of such mathematical models is the Warren and Shelton (1971) Model. The WS Model uses a simultaneous equation approach to analyze important operating and financial variables. The WS Model considers the overall operational and financial environment of the firm. It is flexible so that it can be adapted and extended to meet various circumstances. The model accounts for the interrelations between investment activity, financing activity, dividend policy, and the production decision of the firm and their influences on the market value of the firm.

The critical inputs in the WS Model are the sales growth rate forecast and several operating ratios. The WS Model has four segments including 20 equations simultaneously determining 20 unknowns. The four segments are corresponding to the firm's sales, investments, financing, and return to investment concepts in the financial theory. The model first generates the sales and earnings forecasts given the historical data. Further, the model calculates the total assets required to support these sales and earnings forecasts and the venue through which these assets are to be financed. Finally, given these operation and financing decisions, the models determine the stock per share data.

Essentially the WS Model generates the price per share estimate from the assumptions about the firm's future growth in revenue, its financial and operating policies, and the overall economic environment. The model provides a valuation

[3] International evidence of the debate of comprehensive income reporting in countries such as UK and New Zealand can be found in O'Hanlon and Pope (1999), Cahan et al. (2000), Brimble and Hodgson (2007), and Lin (2006).

framework that includes many different parameter inputs which generate a wide range of estimates that can be presented statistically by a distribution with a mean and variance. The user of the model can easily examine the implication of alternative underlying firm-level and economy-wide environment changes on the equity share prices.

87.2.3 Forecast Combination

Forecast evaluation is of interest in many areas of empirical finance research, such as market efficiency (Fama 1970, 1991), volatility of the observed asset returns (Shiller 1979; LeRoy and Porter 1981; Fama and French 1988), and forward exchange rates (Hansen and Hodrick 1980). Borrowing from Diebold and Lopez (1996) and Yee (2008), we consider two groups of combining forecast methods, i.e., the variance-covariance method and the regression method. Bates and Granger (1969) first proposed the variance-covariance method for forecast combination. Denote the one-period-ahead stock price forecast at time t, $\hat{y}^i_{t,t+1}$, from model $i \in \{1,2\}$; the combined forecast can be formed as the weighted average between the two forecasts:

$$\hat{y}^c_{t,t+1} = \omega \hat{y}^1_{t,t+1} + (1-\omega)\hat{y}^2_{t,t+1}$$

which is an unbiased forecast if the weights sum up to unity. Moreover, the composite forecast error has the same relation as the combined forecast

$$\varepsilon^c_{t,t+1} = \omega \varepsilon^1_{t,t+1} + (1-\omega)\varepsilon^2_{t,t+1}$$

and the variance of the combined forecast

$$\sigma^2_c = \omega^2 + \sigma^2_{11} + (1-\omega)^2\sigma^2_{22} + 2\omega(1-\omega)\sigma_{12}$$

where σ^2_{11}, σ^2_{22}, and σ^2_{12} are the variance of the forecast from model one, model two, and their covariance. The optimal weight to minimize the forecast error ω can be derived as

$$\omega = \frac{\sigma^2_{22} - \sigma_{12}}{\sigma^2_{11} + \sigma^2_{22} - 2\sigma_{12}}$$

which is determined by the variances of each individual forecast and the covariance between them. The asymptotic properties of the optimal weights are

$$\lim_{\sigma^2_{11}\to\infty} \omega = 0 \quad \lim_{\sigma^2_{22}\to\infty} \omega = 1 \quad \lim_{\sigma^2_{11}\to 0} \omega = 1 \quad \lim_{\sigma^2_{22}\to 0} \omega = 0$$

Therefore, the variance-covariance method places larger weight on the more reliable forecast in forecast combination.

The regression method in forecast combination suggests a regression model in which the realization of y_{t+1} is regressed on the past forecasts of y_{t+1} to determine the optimal weights (Chong and Hendry 1986; Fair and Shiller 1989, 1990):

$$y_{t+1} = \alpha_0 + \alpha_1 \hat{y}^1_{t,t+1} + \alpha_2 \hat{y}^2_{t,t+1} + \varepsilon_{t,t+1} \tag{87.10}$$

Granger and Ramanathan (1984) showed that the optimal weight determined in the variance-covariance method has a regression interpretation as the coefficient vector in regression model (87.10) of a linear projection of y_{t+1} onto the forecasts y^i_{t+1} subject to the constraints that the weights α_i sum to unity and the exclusion of the intercept term. However, a number of researchers have recognized that the true but unknown variance-covariance matrix in determining the optimal weight ω is not fixed over time. Therefore, the ensuing research in this literature focused on the time-varying combining weights (Granger and Newbold 1973; Diebold and Pauly 1987) which can be achieved by using the technique of weighted least square (WLS). Diebold and Pauly (1987) proposed a WLS estimator:

$$\hat{\alpha}_{WLS} = (X'W'X)^{-1} X'WY$$

where the weighting matrix can be considered for the following schemes:

1. Equal weight (standard regression-based combining): $w_{tt} = 1$ for all t. (W1)
2. Linear weighting: $w_{tt} = t$ for all t. (W2)
3. Geometric weighting: $w_{tt} = \lambda^{T-t}$, $0 < \lambda \leq 1$, or $w_{tt} = \lambda^t$, $\lambda > 1$. (W3)
4. t^λ (t-lambda): $w_{tt} = t^\lambda$, $\lambda \geq 0$. (W4)

The geometric weighting has the appealing property that the weights increase at an increasing rate as we get closer to the present time. This yields heavy weighting on the more recent observation which might provide better accuracy for forecast values. Moreover, the geometric weighting can provide a weighting scheme that dies out fairly quickly which might be useful in modeling forecasts under an unstable environment. Similarly, the t-lambda (t^λ) specification can also produce weights that die out quickly, but it has an even more appealing fundamental characteristic that its weighting scheme can increase at an either increasing or decreasing rate as we get closer to the present time. When $\lambda = 0$, one obtains the constant weighting scheme in case (1), while when $\lambda = 1$, the linear weighting scheme in case (2) emerges.

The forecast combination equation with two primary individual forecasts can be written as

$$\underset{(T\times 1)}{\mathrm{Y}} = \underset{(T\times 3)}{\mathrm{f}} \underset{(3\times 1)}{\alpha} + \underset{(T\times 1)}{\varepsilon}$$

where

$$\alpha^i = P^i(t) = p^i_0 + p^i_1 t + \cdots + p^i_\gamma, \quad i = 0, 1, 2$$

$f_t = (1, f_t^1, f_t^2)$, $\alpha = (\alpha^0, \alpha^1, \alpha^2)$ ', and **f** is the matrix with tth row f_t. The time-varying combining weights are deterministic nonlinear polynomial functions of time. The advantage of this regression-based deterministically time-varying parameters model over the weighted least square approach is that this method can explicitly model any parameter evolution in the forecast combination equation. This approach can also project the evolution in when the forecasts are combined. The general polynomial and unrestricted regression-based combination is the following:

$$\begin{aligned} y_t &= \left(p_0^0 + p_1^0 t\right) + \left(p_0^1 + p_1^1 t\right)_{t-1} f_t^1 + \left(p_0^2 + p_1^2 t\right)_{t-1} f_t^2 \\ &= p_0^0 + p_1^0 t + p_{0\,t-1}^1 f_t^1 + p_1^1\left({}_{t-1}f_t^1\right) + p_{0\,t-1}^2 f_t^2 + p_1^2\left({}_{t-1}f_t^2\right) \end{aligned} \tag{87.11}$$

Similarly, the forecast can be obtained after estimating the parameters $\hat{p}_0^i$ and $\hat{p}_1^i$:

$$\begin{aligned} {}_t\hat{y}_{t+1} = &\left(\hat{p}_0^0 + \hat{p}_0^1(t+1)\right) + \left(\hat{p}_0^1 + \hat{p}_1^1(t+1)\right)_t f_{t+1}^1 \\ &+ \left(\hat{p}_0^2 + \hat{p}_1^2(t+1)\right)_t f_{t+1}^2 \end{aligned} \tag{87.12}$$

The weighted least square approach can be further combined with the time-varying parameters to determine the optimal weight in the forecast combination equation. For example, one can use geometric weighting scheme λ^{T-t} to construct the weighting matrix W and then estimate the parameters in Eq. 87.11. The estimated parameters $\hat{p}_0^i$ and $\hat{p}_1^i$ are then used in Eq. 87.12 to compute the forecast values. We will explore more possible combination of different weighting schemes and the time-varying parameters model to examine the combination of primary individual forecasts.

87.3 Data and Methodology

In this section, we first introduce the sample selection criteria. Then we present the research hypotheses and model specifications for the empirical tests discussed in the next section.

87.3.1 Data

The data used in this chapter is obtained from the intersection of the following three data sets between 1980 and 2007: annual Compustat for historical accounting data, monthly Center for Research in Security Prices (CRSP) for stock returns, and analyst forecast file from I/B/E/S. The following Compustat data items are used to construct the variables required in the empirical analysis in this chapter: cash and cash equivalent (# 1), total assets (# 6), long-term debt (# 9), interest expense (# 15), investments and advancements (# 32), debt in current liabilities (# 34), interest income (# 62), preferred shares (# 130), short-term investments (# 193), total liabilities (# 181), and notes payable (# 206). Moreover, the book value per share and price data are obtained from Compustat as well. For the Ohlson Model, the data required are already available from the data obtained from the aforementioned sources. However, for empirically testing the FO Model, further distinction between the net operating assets and the net financial assets, and between the operating earnings and the financial earnings, has to be conducted. As discussed previously, the FO Model assumes that the conservative accounting only applies to the operating assets while financial activities are all zero net present value investments. Consequently, only operating assets generate the differences between their book value and the market value which is the goodwill. However, neither theory nor the empirical rules demonstrate how to distinguish between the financial and operating assets. We follow the procedure outlined in Penman (2000) and Callen and Segal (2005) to calculate the operating assets and the financial assets:

$$\begin{aligned}
\text{Financial Assets} &= \text{Cash and Cash Equivalent } (\#\,1) + \text{Investments and Advancements } (\#\,32)\\
&\quad + \text{Short} - \text{Term Investments } (\#\,193)\\
\text{Financial Liabilities} &= \text{Long} - \text{Term debt } (\#\,9) + \text{Debt in Current Liabilities } (\#\,34)\\
&\quad + \text{Notes Payable } (\#\,206)\\
\text{Operating Assets} &= \text{Total Assets } (\#\,6) - \text{Financial Assets}\\
\text{Operating Liabilities} &= \text{Preferred Shares } (\#\,130) + \text{Total Liabilities } (\#\,181)\\
&\quad - \text{Financial Liabilities}\\
\text{Net Operating Assets} &= \text{Operating Assets} - \text{Operating Liabilities}\\
\text{Net Financial Assets} &= \text{Financial Assets} - \text{Financial Liabilities}
\end{aligned}$$

We also use comprehensive (operating) earnings in the linear information dynamics and examine how it affects the accuracy in forecasting future stock prices. SFAS 130 is effectively adopted in 1998 before which firms were not required to report comprehensive income. We follow Cheng et al. (1993), Dhaliwal et al. (1999), and Biddle and Choi (2006) to measure the comprehensive income. We did not use the actual reported comprehensive income given the lack of consistency in reporting of firms in our sample period.[4] The definition of comprehensive income by SFAS 130 is the net income adjusted for "other comprehensive income" items. These items include (1) the change in the balance of unrealized and losses on available-for-sale marketable securities (MSA), (2) the change in cumulative foreign currency translation adjustments (RECTA), and (3) the change in additional minimum pension liability in excess of unrecognized prior service costs (PENADJ). All these variables are scaled by the beginning-of-period market value of equity and they are calculated by using the Compustat data. *MSA* and *RECTA*

[4]In the post-SFAS 130 periods, Compustat has not yet completely disclosed all components in comprehensive income. Currently, Compustat only reports some of the items related to comprehensive income and these data are only complete after year 2001. Given that our empirical tests require sufficient time series to conduct forecasting, we employ the measurement methodology in Cheng et al. (1993) and Dhaliwal et al. (1999) to estimate comprehensive income.

are items *marketable securities adjustment* and *Retained Earnings − Cumulative Translation Adjustment* obtained directly from Compustat. *PENADJ* is calculated as *Pension-Additional Minimum Liability (PADDML) − Pension-Unrecognized Prior Service Cost (PCUPSO)*. The comprehensive income defined in SFAS 130, NI_{130}, is equal to *NI+MSA+RECTA+PENADJ*:

$$\begin{aligned}\text{Comprehensive Income } (x_{130}) = &\text{ Net Income } (\#\,172) + MSA(\#\,238)\\ &+ RECTA(\#\ \ \#\,230) + PENADJ(\#\,297 - \#\,298)\end{aligned}$$

Given the adjustments in calculating the comprehensive income, we further define the comprehensive operating income (ox_{130}) which is used in the information dynamics of the FO Model. Following Nissim and Penman (2001), we define the comprehensive operating income as follows:

$$\begin{aligned}&\text{Comprehensive Operating Income } \ (ox_{130}) = \text{Comprehensive Income } \ (x_{130})\\ &+ \text{Comprehensive Net Financial Expenses } \ (NFE) - \text{Minority Interest in Income } (\#\,49)\end{aligned}$$

where

Comprehensive Net Financial Expenses (*NFE*)
= Core Net Financial Expenses (*Core NFE*) + Unusual Financial Expenses (*UFE*)
= After-Tax Interest Expense (# 15×(1-marginal tax rate[5])) + Preferred Dividends (# 19) – After-Tax Interest Income (# 62×(1-marginal tax rate)) + Change in MSA (Lag # 238 – # 238)

Several other variables used in this chapter are discussed below. The earning used in our empirical analysis is the earnings from the continued operations obtained from I/B/E/S. We follow Callen and Segal (2005) to use the earnings reported in the I/B/E/S because of their comparability with the analyst earnings forecasts. Moreover, the interest rate on debt is computed as the interest expense (DATA 15) divided by the average financial liabilities. The cost of equity capital is calculated by the Fama-French three-factor model and the annualized 3-month treasury bill rate. Finally, we exclude the financial institution (SIC 6000) from the sample because of their minimal level of operating assets and the additional regulatory requirements. Observations with market value of equity less than $10 million, with negative operating and financial assets (liabilities), and with negative net operating and financial earnings are excluded as well. Finally, firms whose empirical variables are less than two firm years are deleted.

[5]Borrowing from Nissim and Penman (2001), the marginal tax rate is the top statutory federal tax rate plus 2 % average state tax rate. For our sample periods, the top statutory federal tax rate was 46 % in 1979–1986, 40 % in 1987, 34 % in 1988–1992, 35 % in 1993–1999, 40 % in 2000–2002, and 35 % in 2003–2008.

87.3.2 Research Hypotheses

Following Dechow et al. (1999) and Callen and Segal (2005), the empirical works in testing the Ohlson Model and FO Model employ the analyst earnings forecast to be the proxy for the other value-relevant information variable. The FO Model further supplements the Ohlson Model with the adjustment for conservative accounting towards which US GAPP is biased. By using the data in the US market, the FO Model is expected to produce better stock price forecast accuracy in terms of smaller prediction errors. As a result, we formulate our first testable hypothesis as follows:

H_{10} *The Feltham and Ohlson* (1995) *Model provides more accurate future stock price forecasts, in terms of smaller mean forecast errors, than the Ohlson* (1995) *Model.*

The use of analyst earnings forecasts is commonly used by the practitioner as well since they capture the forward-looking estimation of the performance of the firm. Should one use the analyst earnings forecasts to predict the future earnings, it is possible that there exists a relation between the two variables. Moreover, for the FO Model, it is also possible that there exist feedback relations of the operating earnings and operating assets with the short-term and long-term analyst earnings forecasts. As a result, we follow Tsay et al. (2008) to use a simultaneous equation approach to estimate the linear information dynamics in both the Ohlson Model and FO Model. By employing the simultaneous equation estimation in the linear information dynamics, we expect to capture the interaction of the future-period earnings with the current-period analyst earnings forecasts. We conjecture that the earnings forecasts influence the future-period earnings and thus the valuation of the equity. At the same time, the current-period earnings also affect the earnings forecasts produced by the analysts. It is the interrelationships between these variables that determine the fundamental value of the equity shares. Thus, we develop our second testable hypothesis as follows:

H_{20} *Simultaneous equation estimation of the linear information dynamics generates more accurate future stock price forecasts, in terms of smaller mean forecast errors, than single equation estimation in both the Ohlson* (1995) *Model and Feltham and Ohlson* (1995) *Model.*

In addition to the Ohlson Model and FO Model discussed above, we next focus on the stock price forecasts ability of the WS Model. The WS Model uses a simultaneous equation approach to forecast future stock price by considering both operating and financing decision of the firms. The WS Model is more comprehensive than the residual income valuation models because it accounts for the interrelations between investment activity, financing activity, dividend policy, and the production decision of the firm. Given its flexibility, the WS Model is expected to better predict future stock prices than the Ohlson/FO Model discussed previously. Thus, our third testable hypothesis is the following:

H_{30} *The Warren and Shelton* (1971) *Model can generate more accurate future stock price forecasts, in terms of smaller mean forecast errors, than both the Ohlson* (1995) *Model and Feltham and Ohlson* (1995) *Model.*

After considering the future stock price forecasts from these valuation models, we further investigate whether they can be combined to form more accurate forecasts. We thus employ the combined forecast methods (Granger and Newbold 1973; Granger and Ramanathan 1984; Diebold and Pauly 1987) to examine whether forecast combination is more accurate than individual forecasts in terms of mean forecast errors. Therefore, the fourth testable hypothesis in this chapter is the following:

H_{40} *The combination of individual forecasts from the Ohlson Model/Feltham-Ohlson Model and the WS Model can generate more accurate future stock price forecasts, in terms of smaller mean forecast errors, than each individual forecasts.*

Finally, we investigate whether comprehensive (operating) earnings can provide incremental price-relevant information beyond bottom-line earnings. We employ the comprehensive (operating) earnings in the linear information dynamics of the Ohlson Model/Feltham-Ohlson Model and examine its effects on the future stock price forecasts. These forecasts are further combined with the WS Model forecasts and their forecast accuracy is examined. Thus, the fifth testable hypothesis can be stated as follows:

H_{50} *Using comprehensive earnings as the earnings measure in the linear information dynamics of the Ohlson Model and Feltham-Ohlson Model can generate more accurate future stock price forecasts, in terms of mean forecast error, than bottom-line earnings as the earnings measure.*

87.3.3 The Model Specifications

In this section, we propose two sets of the linear information dynamics in the Ohlson Model and FO Model. The estimated coefficients from these information dynamics are further used in the valuation function to forecast future stock prices. The first set of specifications includes the single equation and the simultaneous equation estimation with the analyst forecast of earnings in the linear information dynamics of the Ohlson Model. The single equation approach specified in this set is essentially the model tested in Dechow et al. (1999). Moreover, as Tsay et al. (2008) stated that there are feedback relations between the other value-relevant information and the earnings, we further employ the simultaneous equation linear information dynamics. More specifically, we examine whether the feedback effect between the current earnings, analyst forecasts, and the book value improves the predictability of the model.

1. Model Set I: In the first model specification, we test the Ohlson Model with the other value-relevant information variable. This variable essentially summarizes information that is captured in a firm's stock because it can predict future abnormal earnings but is not yet reflected in the financial statements. Here we test a modified version of Dechow et al. (1999) model in which the other value-relevant information variable is the analysts' earnings forecasts. The linear information dynamic is

$$\begin{aligned} \widetilde{x}^a_{i,t+1} &= \omega_{10} + \omega_{11} x^a_{i,t} + \omega_{12} v_{i,t} + \widetilde{\varepsilon}_{i1,t+1} \\ \widetilde{v}_{i,t+1} &= \omega_{20} + \omega_{22} v_{i,t} + \widetilde{\varepsilon}_{i2,t+1} \end{aligned} \tag{87.13}$$

where $\widetilde{x}^a_{i,t+1}$ is the abnormal earnings of firm i at time t, and $v_{i,t}$ is the difference between the conditional expectation of abnormal earnings for firm i at time period t+1 based on all available information and the expectation of abnormal earnings, i.e., $v_{i,t} = E_t\left[\widetilde{x}^a_{i,t+1}\right] - \omega_{11}x^a_{i,t}$. Following Dechow et al. (1999), the period t conditional expectation of period t+1 earnings is the median consensus analyst forecast of period t+1 earnings denoted by f_t, i.e., $E_t\left[\widetilde{x}^a_{i,t+1}\right] = f^a_{i,t} = f_{i,t} - ry_{i,t}$, where $f_{i,t}$ is the median consensus analyst earnings forecasts of next year's earnings measured at the first month after the publication of the annual financial report. Consequently, the other value-relevant information can thus be written as $v_{i,t} = f^a_{i,t} - \omega_{11}x^a_{i,t}$.

The simultaneous equation specification of the linear information dynamic on the other hand is

$$\begin{aligned}\widetilde{x}^a_{i,t+1} &= \omega_{10} + \omega_{11}x^a_{i,t} + \omega_{12}v_{i,t} + \widetilde{\varepsilon}_{i1,t+1}\\ \widetilde{v}_{i,t+1} &= \omega_{20} + \omega_{21}x^a_{i,t} + \omega_{22}v_{i,t} + \widetilde{\varepsilon}_{i2,t+1}\end{aligned} \tag{87.13}$$

where the coefficient ω_{21} represents the feedback effect from current-period abnormal earnings to next-period analyst earnings forecasts. Given the specification of the information dynamics in Eqs. 87.12 and 87.13, the implied valuation function can be written as

$$P_{i,t} = y_{i,t} + \hat{\beta}_0 + \hat{\beta}_1 x^a_{i,t} + \hat{\beta}_2 v_{i,t} \tag{87.14}$$

where the estimated coefficients are

$$\begin{aligned}\hat{\beta}_0 &= \frac{\left(1+\hat{r}_{i,t}\right)\left[\hat{\omega}_{10}\left(1+\hat{r}_{i,t}-\hat{\omega}_{22}\right)+\hat{\omega}_{12}\hat{\omega}_{20}\right]}{\hat{r}_{i,t}\left(1+\hat{r}_{i,t}-\hat{\omega}_{11}\right)\left(1+\hat{r}_{i,t}-\hat{\omega}_{22}\right)}\\ \hat{\beta}_1 &= \frac{\hat{\omega}_{11}}{\left(1+\hat{r}_{i,t}-\hat{\omega}_{11}\right)}\\ \hat{\beta}_2 &= \frac{\left(1+\hat{r}_{i,t}\right)\hat{\omega}_{12}}{\left(1+\hat{r}_{i,t}-\hat{\omega}_{11}\right)\left(1+\hat{r}_{i,t}-\hat{\omega}_{22}\right)}\end{aligned}$$

and $\hat{r}_{i,t}$ is the cost of equity capital for firm i at time t. Or equivalently,

$$P_{i,t} = \kappa\left(\varphi x_{i,t} - d_{i,t}\right) + (1-\kappa)y_{i,t} + \hat{\beta}_2 v_{i,t} \tag{87.15}$$

where $\kappa = \hat{r}_{i,t}\hat{\omega}_{11}/\left(1+\hat{r}_{i,t}-\hat{\omega}_{11}\right)$ and $\varphi = \left(1+\hat{r}_{i,t}\right)/\hat{r}_{i,t}$.

In the second set of the model specification, we test the FO Model with single equation and simultaneous equation linear information dynamics. The FO Model argues that it is important to separate the financial assets and the operating assets in the valuation function since only operating assets generate goodwill.

The FO Model considers the practice of accrual accounting and how it influences the equity valuation.

2. Model Set II: The single equation linear information dynamics in the FO Model is the following:

$$\begin{aligned}
\widetilde{ox}^a_{i,t+1} &= \omega_{10} + \omega_{11} ox^a_{i,t} + \omega_{12} oa_{i,t} + \omega_{13} v_{1i,t} + \widetilde{\varepsilon}_{1i,t+1} \\
\widetilde{oa}^a_{i,t+1} &= \omega_{20} + \omega_{22} oa_{i,t} + \omega_{24} v_{2i,t} + \widetilde{\varepsilon}_{2i,t+1} \\
\widetilde{v}_{1i,t+1} &= \omega_{30} + \omega_{33} v_{1i,t} + \widetilde{\varepsilon}_{3i,t+1} \\
\widetilde{v}_{2i,t+1} &= \omega_{40} + \omega_{44} v_{2i,t} + \widetilde{\varepsilon}_{4i,t+1}
\end{aligned} \tag{87.16}$$

and the simultaneous linear information dynamics is

$$\begin{aligned}
\widetilde{ox}^a_{i,t+1} &= \omega_{10} + \omega_{11} ox^a_{i,t} + \omega_{12} oa_{i,t} + \omega_{13} v_{1i,t} + \omega_{14} v_{2i,t} + \widetilde{\varepsilon}_{1i,t+1} \\
\widetilde{oa}^a_{i,t+1} &= \omega_{20} + \omega_{21} ox^a_{i,t} + \omega_{22} oa_{i,t} + \omega_{23} v_{1i,t} + \omega_{24} v_{2i,t} + \widetilde{\varepsilon}_{2i,t+1} \\
\widetilde{v}_{1i,t+1} &= \omega_{30} + \omega_{31} ox^a_{i,t} + \omega_{32} oa_{i,t} + \omega_{33} v_{1i,t} + \omega_{34} v_{2i,t} + \widetilde{\varepsilon}_{3i,t+1} \\
\widetilde{v}_{2i,t+1} &= \omega_{40} + \omega_{41} ox^a_{i,t} + \omega_{42} oa_{i,t} + \omega_{43} v_{1i,t} + \omega_{44} v_{2i,t} + \widetilde{\varepsilon}_{4i,t+1}
\end{aligned} \tag{87.17}$$

where $ox^a_{i,t}$ is the abnormal operating earnings for firm i at time t, and $oa_{i,t}$ is the operating assets for firm i at time t. Moreover, the value-relevant information variables $v_{1i,t}$ and $v_{2i,t}$ are the growth in expected operating earnings and the expected growth in operating assets, respectively. The expected operating earnings are measured as the difference between the median consensus analyst earnings forecast for next year and the expected net interest revenue (product of end-of-current-year financial liabilities and the interest on debt). The growth in expected operating earnings is defined as the expected change in operating earnings divided by operating assets, i.e., $v_{1i,t} = E_t[\Delta\widetilde{ox}_{t+1}]/oa_t = E_t[\widetilde{ox}_{t+1} - ox_t]/oa_t$, where the current-period operating earnings ox_t are calculated as actual earnings reported by the I/B/E/S minus the interest revenue (product of beginning-of-current-year financial liabilities and the interest on debt). The expected growth in net operating asset is defined as the change in expected operating asset divided by the operating asset, i.e., $v_{2i,t} = E_t[\Delta\widetilde{oa}_{t+1}]/oa_t$. Following Liu and Ohlson (2000), we use the analyst earnings forecasts of long-term earnings growth rates as a proxy for the expected growth in net operating assets.[6]

Given the specification of the information dynamics in Eqs. 87.16 and 87.17, the implied valuation function is

$$P_{i,t} = y_{i,t} + \hat{\lambda}_0 + \hat{\lambda}_1 ox^a_{i,t} + \hat{\lambda}_2 oa_{i,t} + \hat{\lambda}_3 v_{1i,t} + \hat{\lambda}_4 v_{2i,t} \tag{87.18}$$

[6]The long-term growth forecast generally represents an expected increase in operating earnings over the company's next full business cycle. Usually, these forecasts refer to a period of between 3 and 5 years. Thomson Financial recommends the median value for long-term growth forecast rather than the mean. The median value is less affected by outlier forecasts.

where

$$\hat{\lambda}_0 = \frac{\left(1+\hat{r}_{i,t}\right)\begin{bmatrix}\hat{\omega}_{10}\left(1+\hat{r}_{i,t}-\hat{\omega}_{22}\right)\left(1+\hat{r}_{i,t}-\hat{\omega}_{33}\right)\left(1+\hat{r}_{i,t}-\hat{\omega}_{44}\right)\\ +\hat{\omega}_{12}\hat{\omega}_{20}\left(1+\hat{r}_{i,t}-\hat{\omega}_{33}\right)+\hat{\omega}_{13}\hat{\omega}_{30}\left(1+\hat{r}_{i,t}-\hat{\omega}_{22}\right)\\ +\hat{\omega}_{14}\hat{\omega}_{40}\left(1+\hat{r}_{i,t}-\hat{\omega}_{44}\right)\end{bmatrix}}{\hat{r}_{i,t}\left(1+\hat{r}_{i,t}-\hat{\omega}_{11}\right)\left(1+\hat{r}_{i,t}-\hat{\omega}_{22}\right)\left(1+\hat{r}_{i,t}-\hat{\omega}_{33}\right)\left(1+\hat{r}_{i,t}-\hat{\omega}_{44}\right)}$$

$$\hat{\lambda}_1 = \frac{\hat{\omega}_{11}}{\left(1+\hat{r}_{i,t}-\hat{\omega}_{11}\right)}$$

$$\hat{\lambda}_2 = \frac{\left(1+\hat{r}_{i,t}\right)\hat{\omega}_{12}}{\left(1+\hat{r}_{i,t}-\hat{\omega}_{11}\right)\left(1+\hat{r}_{i,t}-\hat{\omega}_{22}\right)}$$

$$\hat{\lambda}_3 = \frac{\left(1+\hat{r}_{i,t}\right)\hat{\omega}_{13}}{\left(1+\hat{r}_{i,t}-\hat{\omega}_{11}\right)\left(1+\hat{r}_{i,t}-\hat{\omega}_{33}\right)}$$

$$\hat{\lambda}_4 = \frac{\left(1+\hat{r}_{i,t}\right)\hat{\omega}_{14}}{\left(1+\hat{r}_{i,t}-\hat{\omega}_{11}\right)\left(1+\hat{r}_{i,t}-\hat{\omega}_{44}\right)}$$

Or equivalently,

$$P_{i,t} = k\left(\phi x_{i,t} - d_{i,t}\right) + (1-\kappa)y_{i,t} + \hat{\lambda}_2 oa_{i,t} + \hat{\lambda}_3 v_{1i,t} + \hat{\lambda}_4 v_{2i,t} \qquad (87.19)$$

where $\kappa = \hat{r}_{i,t}\hat{\omega}_{11}/\left(1+\hat{r}_{i,t}-\hat{\omega}_{11}\right)$ and $\phi = \left(1+\hat{r}_{i,t}\right)/\hat{r}_{i,t}$.

On the basis of these model specifications, we empirically test the research hypotheses constructed in the previous section. For the model's ability to explain the cross section of the stock prices, we first estimate the parameters ω_{ii} in the linear information dynamics under both single equation and simultaneous equation model. Given these estimated coefficients ω_{ii}, we compute the theoretical stock prices implied by the pricing equations and compare the results to the observed prices. Pricing errors of the implied valuation function will be calculated and we examine the model's ability to explain the cross section of stock prices under different model specifications. For the predictability of future stock prices on the other hand, we run the regression of each valuation function to obtain the estimated coefficients. These estimated coefficients are then used with the observation in the future periods to compute the theoretical value of the equity in the future periods. Similarly, the prediction errors of these prices calculated from the implied valuation function are calculated for the comparison between various model specifications.

87.4 Empirical Results

Based on the research hypotheses constructed in the previous section, this chapter empirically tests the Ohlson Model and the FO Model under different linear information dynamics. We conjecture that the linear information dynamic including the value-relevant information variables such as analyst earnings forecasts and the insider transaction activity improves the power of the model to explain the cross section of stock prices and to predict future price movement. Furthermore, we also expect the simultaneous linear information dynamic to exhibit superior ability in pricing the equity share than its single equation counterpart.

In the following empirical analysis, we first use single equation approach and the simultaneous equation approach to estimate the coefficient in each linear information dynamic specification. The estimated coefficients are then used to compute the theoretical price of the equity implied by the valuation function. These implied values are then compared to the prices actually being observed to examine whether the model explains the cross section of the stock prices. If the Ohlson Model and the FO Model have empirical content, then the pricing errors produced by these specifications will not be statistically significant. Furthermore, we test the model's ability to forecast the future-period stock prices. We use the panel data random effect model in each implied valuation functions to estimate the coefficients. These estimated coefficients are then used with the future-period firm-level data to compute the forecasts of the stock prices. Similarly if the models have empirical content, they should produce minimal pricing errors compared to the model tested in the previous literature.

87.4.1 Summary Statistics

Table 87.1 provides the summary statistics of the variables used in our empirical analysis. We note that the mean earnings per share in our sample period is 0.415 while the abnormal earnings per share is −0.294. This indicates that the firms on average earned less than the required cost of equity capital of 0.146 in our sample period. Moreover, the analysts are on average optimistic about future earnings performance of the firms given that the mean analyst earnings per share forecast is a positive 0.172. The two value-relevant information variables in the FO Model, expected growth in operating earnings and in operating assets, have mean values of 0.031 and 0.018, respectively, in our sample period. Finally, the mean of the book value per share and stock price per share in our sample period is 9.015 and 21.361, respectively.

87.4.2 Time Series Behavior of Linear Information Dynamics

We next examined the time series behavior of the linear information dynamics in the Ohlson Model and FO Model in Tables 87.2 and 87.3, respectively.[7] Panel A1 in Table 87.2 shows that the abnormal earnings follow a stationary process since the coefficients of the lagged variables sum up less than one. Moreover, the first lag of

[7]Since the linear information dynamics contains lagged dependent variables, the OLS estimation is inconsistent. We proceed our estimation by the IV estimation and panel GMM proposed by Anderson and Hsiao (1981) and Arellano and Bond (1991), respectively. The panel GMM is more efficient than the IV estimator because of additional lags of dependent variable as instruments. The results from the two estimation methods are similar and we reported the results from panel GMM estimator.

Table 87.1 Descriptive statistics for variables used in sample (1980–2008)

Variables	Mean	Std. dev.	Q1	Median	Q3
Beta (β)	0.9926	2.4995	0.30	0.85	1.56
Cost of equity capital (r)	0.0342	0.0436	0.01	0.02	0.04
Earnings per share (x)	0.5545	1.0330	−0.02	0.60	1.24
Abnormal earnings per share (x^a)	0.3198	1.0269	−0.24	0.37	1.00
Growth in expected operating earnings (v_1)	0.1721	0.0357	−0.09	−0.04	0.05
Expected growth of operating assets (v_2)	0.2253	0.0362	−0.09	−0.04	0.05
Book value per share (b)	8.0327	5.4195	3.74	7.13	11.41
Analyst earning forecasts of abnormal earnings per share (v)	1.3031	0.0895	0.25	0.66	1.20
Stock price per share (P)	15.5368	11.0900	6.75	13.75	23.13
Comprehensive earnings per share (x_{130})	15.6134	13.2147	5.47	14.03	24.33
Marketable securities adj. (*MSA*)	0.0743	0.2161	0.02	0.09	0.16
Foreign currency translation adj. (*RECTA*)	0.0000	0.0087	−0.01	0.01	0.04
Pension requirement adj. (*PENADJ*)	0.0765	0.1357	0.02	0.08	0.05
Comprehensive operating earnings per share (ox_{130})	16.4327	12.3154	8.11	14.54	25.67

Our sample contains all available data in Compustat, CRSP, and I/B/E/S from 1980 to 2008. The firm's beta β is calculated based on the market model using the monthly rate of returns in the past 60 months. The cost of equity r is calculated based on the CAPM. Abnormal earnings per share (x^a) are the current-period earnings per share less the previous earnings per share growing at the cost of equity capital r. v is the analyst earnings forecasts of abnormal earnings per share. v_1 is the growth in expected operating earnings per share, measured as the difference between the median consensus analyst earnings per share forecast for next year and the expected net interest revenue (product of end-of-current-year financial liabilities and the interest on debt), divided by operating asset oa. v_2 is the expected growth in operating assets which is proxied by the analyst earnings per share forecasts of long-term earnings growth rate, divided by operating asset oa. x_{130} is the comprehensive earnings per share which is calculated as earnings per share adjusted for marketable securities adjustments (*MSA*) in Compustat, cumulative foreign currency translation adjustments (*RECTA*) in Compustat, and pension requirement adjustments (*PENADJ*). *PENADJ* is calculated as Pension-Additional Minimum Liability (*PADDML*) − Pension-Unrecognized Prior Service Cost (*PCUPSO*) in Compustat. *MSA*, *RECTA*, and *PENADJ* are all scaled by the beginning-of-period total shares outstanding. Comprehensive operating income (ox_{130}) is equal to comprehensive income (x_{130}) adjusted for comprehensive net financial expenses (*NFE*) and minority interest in income

the abnormal earnings accounts for the most serial correlation in the abnormal earning process. Even though the estimated coefficients for the other lagged variables are statistically significant at 5 %, the adjusted R^2 is approximately the same after considering these lags. Therefore, we found that the AR(1) process for abnormal earnings is sufficient for the linear information dynamics in the Ohlson Model and the FO Model as documented by previous literature. Furthermore, Panel B1 in Table 87.2 showed that after accounting for the analyst forecasts in the linear information dynamics for the abnormal earnings, the adjusted R^2 increases to 69 % from 40 %. This indicates that the analyst earnings forecast serves to be an appropriate value-relevant variable that adds explanatory power to the linear information dynamics.

Table 87.2 Time series behavior of Ohlson Model linear information dynamics

Estimated regression coefficients				
ω_{10}	ω_{11}	ω_{12}	ω_{13}	Adj. R^2
Panel A1: Autoregressive property of abnormal earnings with different lags				
Model: $\tilde{x}^a_{i,t+1} = \omega_{10} + \omega_{11}x^a_{i,t} + \tilde{\varepsilon}_{i1,t+1}$				
0.0338	0.7788			0.3719
(0.0425)	(<0.0001)			
Model: $\tilde{x}^a_{i,t+1} = \omega_{10} + \omega_{11}x^a_{i,t} + \omega_{12}x^a_{i,t-1} + \tilde{\varepsilon}_{i1,t+1}$				
0.0057	0.6111	0.2864		0.3811
(0.0412)	(<0.0001)	(<0.0001)		
Model: $\tilde{x}^a_{i,t+1} = \omega_{10} + \omega_{11}x^a_{i,t} + \omega_{12}x^a_{i,t-1} + \omega_{13}x^a_{i,t-2} + \tilde{\varepsilon}_{i1,t+1}$				
−0.1312	0.4889	0.2798	0.1622	0.3566
(0.0312)	(<0.0001)	(<0.0001)	(0.0655)	
Panel A2: Autoregressive property of abnormal comprehensive earnings with different lags				
Model: $\tilde{x}^a_{130i,t+1} = \omega_{10} + \omega_{11}x^a_{130i,t} + \tilde{\varepsilon}_{i1,t+1}$				
0.0155	0.8012			0.4114
(0.0231)	(0.0001)			
Model: $\tilde{x}^a_{130i,t+1} = \omega_{10} + \omega_{11}x^a_{130i,t} + \omega_{12}x^a_{130i,t-1} + \tilde{\varepsilon}_{i1,t+1}$				
0.0121	0.7411	0.1642		0.3978
(0.0013)	(<0.0001)	(0.0654)		
Model: $\tilde{x}^a_{103i,t+1} = \omega_{10} + \omega_{11}x^a_{103i,t} + \omega_{12}x^a_{103i,t-1} + \omega_{13}x^a_{130i,t-2} + \tilde{\varepsilon}_{i1,t+1}$				
−0.0004	0.7841	0.1681	0.0914	0.4001
(0.1124)	(0.0003)	(0.0241)	(0.1511)	
Panel B1: Autoregressive property of abnormal earnings with other information variables				
Model: $\tilde{x}^a_{i,t+1} = \omega_{10} + \omega_{11}x^a_{i,t} + \omega_{12}v_{i,t} + \tilde{\varepsilon}_{i1,t+1}$				
0.0524	0.6054	0.3052		0.4121
(<0.0001)	(0.0105)	(0.0211)		
Panel B2: Autoregressive property of abnormal comprehensive earnings with other information variables				
Model: $\tilde{x}^a_{130i,t+1} = \omega_{10} + \omega_{11}x^a_{130i,t} + \omega_{12}v_{i,t} + \tilde{\varepsilon}_{i1,t+1}$				
0.0214	0.6327	0.1864		0.4364
(0.0015)	(0.0064)	(0.0514)		
Panel C: Autoregressive property of analyst earnings per share forecast with different lags				
Model: $\tilde{v}_{i,t+1} = \omega_{10} + \omega_{11}v_{i,t} + \tilde{\varepsilon}_{i2,t+1}$				
0.1754	0.4185			0.6884
(<0.0001)	(<0.0001)			
Model: $\tilde{v}_{i,t+1} = \omega_{10} + \omega_{11}v_{i,t} + \omega_{12}v_{i,t-1} + \tilde{\varepsilon}_{i2,t+1}$				
0.2374	0.4339	0.2634		0.6791
(0.0050)	(<0.0001)	(<0.0001)		
Model: $\tilde{v}_{i,t+1} = \omega_{10} + \omega_{11}v_{i,t} + \omega_{12}v_{i,t-1} + \omega_{13}v_{i,t-2} + \tilde{\varepsilon}_{i2,t+1}$				
0.1499	0.3866	0.2561	0.1735	0.7015
(0.0015)	(<0.0001)	(<0.0001)	(0.0745)	

Panel GMM methodology proposed by Arellano and Bond (1991) is used to examine the autoregressive property of earnings and other value-relevant information variable in the Ohlson Model. The estimated coefficients ω_{ij} and the adjusted R^2 from the regression are provided in each panel. $x^a_{i,t}$ is the abnormal earnings per share for firm i at year t. $x^a_{130i,t}$ is the abnormal comprehensive earnings per share for firm i at time t. $v_{1i,t}$ is the consensus analyst earnings per share forecasts for year t+1 at year t. The p-values associated with each estimated coefficients are reported in the parenthesis below. Our sample contains all available data in Compustat, CRSP, and I/B/E/S from 1980 to 2008

Table 87.3 Time series behavior of Feltham-Ohlson Model linear information dynamics

Estimated regression coefficients				
ω_{10}	ω_{11}	ω_{12}	ω_{13}	Adj. R^2
Panel A1: Autoregressive property of abnormal operating earnings with different lags				
Model: $\widetilde{ox}^a_{i,t+1} = \omega_{10} + \omega_{11} ox^a_{i,t} + \widetilde{\varepsilon}_{1i,t+1}$				
−0.0099	0.2897			0.6434
(0.1241)	(<0.0001)			
Model: $\widetilde{ox}^a_{i,t+1} = \omega_{10} + \omega_{11} ox^a_{i,t} + \omega_{12} ox^a_{i,t-1} + \widetilde{\varepsilon}_{1i,t+1}$				
−0.0227	0.2567	0.0935		0.6413
−0.1762	(<0.0001)	(<0.0001)		
Model: $\widetilde{ox}^a_{i,t+1} = \omega_{10} + \omega_{11} ox^a_{i,t} + \omega_{12} ox^a_{i,t-1} + \omega_{13} ox^a_{i,t-2} + \widetilde{\varepsilon}_{1i,t+1}$				
−0.0339	0.1886	0.1048	0.0992	0.6505
(0.0513)	(<0.0001)	(<0.0001)	(<0.0001)	
Panel A2: Autoregressive property of abnormal comprehensive operating earnings with different lags				
Model: $\widetilde{ox}^a_{103i,t+1} = \omega_{10} + \omega_{11} ox^a_{130i,t} + \widetilde{\varepsilon}_{1i,t+1}$				
0.0132	0.4231			0.7211
(0.0845)	(<0.0001)			
Model: $\widetilde{ox}^a_{130i,t+1} = \omega_{10} + \omega_{11} ox^a_{130i,t} + \omega_{12} ox^a_{130i,t-1} + \widetilde{\varepsilon}_{1i,t+1}$				
−0.0017	0.3978	0.1412		0.6877
(0.0512)	(<0.0001)	(0.0014)		
Model: $\widetilde{ox}^a_{130i,t+1} = \omega_{10} + \omega_{11} ox^a_{130i,t} + \omega_{12} ox^a_{130i,t-1} + \omega_{13} ox^a_{130i,t-2} + \widetilde{\varepsilon}_{1i,t+1}$				
−0.1687	0.4067	0.0874	0.1021	0.7158
(0.0647)	(0.0014)	(0.0008)	(0.0214)	
Panel B: Autoregressive property of operating assets with different lags				
Model: $\widetilde{oa}_{i,t+1} = \omega_{10} + \omega_{11} oa_{i,t} + \widetilde{\varepsilon}_{1i,t+1}$				
0.0903	0.2432			0.7941
(0.0714)	(0.0015)			
Model: $\widetilde{oa}_{i,t+1} = \omega_{10} + \omega_{11} oa_{i,t} + \omega_{12} oa_{i,t-1} + \widetilde{\varepsilon}_{1i,t+1}$				
0.3092	0.2417	0.1274		0.8004
(0.1150)	(0.0023)	(0.3543)		
Model: $\widetilde{oa}_{i,t+1} = \omega_{10} + \omega_{11} oa_{i,t} + \omega_{12} oa_{i,t-1} + \omega_{13} oa_{i,t-2} + \widetilde{\varepsilon}_{1i,t+1}$				
0.0979	0.1767	0.0601	0.0148	0.8111
(0.0691)	(0.0056)	(0.1245)	(0.2234)	
Panel C1: Autoregressive property of abnormal operating earnings with other information variables				
Model: $\widetilde{ox}^a_{i,t+1} = \omega_{10} + \omega_{11} ox^a_{i,t} + \omega_{12} oa_{i,t} + \widetilde{\varepsilon}_{1i,t+1}$				
0.0699	0.2895	0.4023		0.6433
(0.1186)	(<0.0001)	(0.0049)		
Model: $\widetilde{ox}^a_{i,t+1} = \omega_{10} + \omega_{11} ox^a_{i,t} + \omega_{12} oa_{i,t} + \omega_{13} v_{1i,t} + \widetilde{\varepsilon}_{1i,t+1}$				
0.0708	0.2886	0.2744	0.1125	0.7236
(0.1161)	(<0.0001)	(0.0151)	(0.0278)	
Panel C2: Autoregressive property of abnormal comprehensive operating earnings with other information variables				
Model: $\widetilde{ox}^a_{130i,t+1} = \omega_{10} + \omega_{11} ox^a_{130i,t} + \omega_{12} oa_{i,t} + \widetilde{\varepsilon}_{1i,t+1}$				

(*continued*)

Table 87.3 (continued)

Estimated regression coefficients				
ω_{10}	ω_{11}	ω_{12}	ω_{13}	Adj. R^2
0.0422	0.3115	0.5104		0.7144
(0.0848)	(<0.0001)	(0.0001)		
Model: $\widetilde{ox}^a_{130i,t+1} = \omega_{10} + \omega_{11} ox^a_{130i,t} + \omega_{12} oa_{i,t} + \omega_{13} v_{1i,t} + \widetilde{\varepsilon}_{1i,t+1}$				
0.0214	0.2871	0.3481	0.3978	0.8412
(0.0214)	(<0.0001)	(<0.0001)	(0.0008)	
Panel D: Autoregressive property of operating assets with other information variables				
Model: $\widetilde{oa}_{i,t+1} = \omega_{10} + \omega_{11} oa_{i,t} + \omega_{12} v_{2i,t} + \widetilde{\varepsilon}_{1i,t+1}$				
−0.0602	0.2204	0.3052		0.8321
(<0.0001)	(<0.0001)	(0.0420)		
Panel E: Autoregressive property of growth of expected operating earnings with different lags				
Model: $\widetilde{v}_{1i,t+1} = \omega_{10} + \omega_{11} v_{1i,t} + \widetilde{\varepsilon}_{1i,t+1}$				
−0.0436	−0.0587			0.3125
(0.6591)	(<0.0001)			
Model: $\widetilde{v}_{1i,t+1} = \omega_{10} + \omega_{11} v_{1i,t} + \omega_{12} v_{1i,t-1} + \widetilde{\varepsilon}_{1i,t+1}$				
−0.0422	−0.0674	0.1324		0.3016
(0.6673)	(<0.0001)	(0.0745)		
Model: $\widetilde{v}_{1i,t+1} = \omega_{10} + \omega_{11} v_{1i,t} + \omega_{12} v_{1i,t-1} + \omega_{13} v_{1i,t-2} + \widetilde{\varepsilon}_{1i,t+1}$				
−0.0394	−0.0778	0.1238	0.0747	0.3214
(0.6882)	(<0.0001)	(0.1023)	(0.0621)	
Panel F: Autoregressive property of expected growth of operating assets with different lags				
Model: $\widetilde{v}_{2i,t+1} = \omega_{10} + \omega_{11} v_{2i,t} + \widetilde{\varepsilon}_{2i,t+1}$				
0.6493	0.6685			0.4874
(0.0004)	(<0.0001)			
Model: $\widetilde{v}_{2i,t+1} = \omega_{10} + \omega_{11} v_{2i,t} + \omega_{12} v_{2i,t-1} + \widetilde{\varepsilon}_{2i,t+1}$				
0.3393	0.1599	0.2912		0.4905
(0.0356)	(<0.0001)	(0.0647)		
Model: $\widetilde{v}_{2i,t+1} = \omega_{10} + \omega_{11} v_{2i,t} + \omega_{12} v_{2i,t-1} + \omega_{13} v_{2i,t-2} + \widetilde{\varepsilon}_{2i,t+1}$				
0.2717	0.1180	0.1568	0.2004	0.4951
(0.0771)	(<0.0001)	(0.0487)	(0.1311)	

Panel GMM methodology proposed by Arellano and Bond (1991) is used to examine the autoregressive property of earnings and other value-relevant information variable in the FO Model. The estimated coefficients ω_{ij} and the adjusted R^2 from the regression are provided in each panel. $ox^a_{i,t}$ is the abnormal operating earnings per share for firm i at year t. $ox^a_{130i,t}$ is the abnormal comprehensive operating earnings per share for firm i at year t. $oa_{i,t}$ is operating asset scaled by total assets for firm i at year t. $v_{1i,t}$ is the growth in expected operating earnings, measured as the difference between the median consensus analyst earnings forecast for next year and the expected net interest revenue (product of end-of-current-year financial liabilities and the interest on debt). $v_{2i,t}$ is the expected growth in operating assets, measured by the analyst earnings forecasts of long-term earnings growth rate. The p-values associated with each statistics are reported in the parenthesis below. Our sample contains all available data in Compustat, CRSP, and I/B/E/S from 1980 to 2008

Given that the Ohlson Model and FO Model are based on clean surplus relation, we also use comprehensive (operating) earnings in estimating the linear information dynamics and deriving implied valuation function. In Panel A2 of Table 87.2, we first investigate the autoregressive property of abnormal comprehensive earnings. Similar to the results found in Panel A1, we find strong statistical and economic magnitude of lag one comprehensive earnings, i.e., ranging from 0.7411 to 0.8012. The higher-ordered lags do not provide additional information given the adjusted R^2 does not improve after including additional lagged comprehensive earnings. The analyst earnings forecasts also retain its value relevance in the presence of the comprehensive earnings. Panel B2 of Table 87.2 shows that including analyst earnings forecasts improves the adjusted R^2 in the regression, from 0.4121 to 0.4363, with only lagged comprehensive earnings in Panel B1. These results confirm our previous finding that lagged one comprehensive earnings and analyst earnings forecasts are sufficient to estimate the earnings dynamics in the Ohlson Model. More interestingly, using comprehensive earning, instead of the bottom-line earnings, also seems to provide explanatory power given the high adjusted R^2 in the regression. Finally, Panel C showed the extended autoregressive process for the analyst earnings forecasts. We note that the AR(1) process again is sufficient for the linear information dynamics given that the further lagged variables do not contribute to the overall explanatory power of the model. In summary, we found that AR(1) process is sufficient for both the aerial correlation in abnormal earnings and analyst earnings forecasts. Incorporating the analyst earnings forecasts into the linear information dynamics indeed improved the ability of the model to explain the abnormal earnings process in addition to its own serial dependence.

Table 87.3 summarizes the autoregressive behavior of the abnormal operating earnings, operating assets, expected growth of operating earnings, and expected growth of operating assets in the linear information dynamics of the FO Model. We found that the AR(1) process is sufficient for all four variables given that the adjusted R^2 is approximately the same after more lagged variables are considered. In Panel A1, e.g., when the lag 2 abnormal earnings are added, the estimated coefficient for lag 1 is still statistically significant. Although the lag 2 abnormal earnings are also statistically significant, lag 1 abnormal earnings account for most of the serial dependence of the abnormal earnings process. Moreover, when additional lag 3 abnormal earnings are considered, the significance of the lag 1 abnormal earnings and adjusted R^2 is not affected. Similar results can be found for operating assets, expected growth in operating earnings, and expected growth in operating assets. Thus, we note that the AR(1) process is sufficient for the aforementioned variables in the linear information dynamics of the FO Model. In Panel A2 of Table 87.3, we investigate the dynamics of comprehensive operating earnings. Similar to the operating earnings dynamics in Panel A1, we find that lag 1 comprehensive earnings are sufficient in explaining the dynamics and additional lagged variables do not provide more information.

Panel C1 in Table 87.3 showed that incorporating more value-relevant variables in the linear information dynamics increases the ability of the model to explain the variation of abnormal operating earnings. With only lagged abnormal earnings and operating assets as the independent variables in the model, both of the variables are statistically significant with estimated coefficients of 0.2895 and 0.4023, respectively. When the expected growth of operating earnings is incorporated, lagged abnormal operating earnings and operating assets are still statistically significant and moreover the adjusted R^2 increased from 0.6433 to 0.7236. This indicates that the additional value-relevant variable, the expected growth in appearing earnings, indeed increased the explanatory power of the linear information dynamics. Moreover, the results in Panel C2 indicate that under the single equation estimation, incorporating operating assets and analyst earnings forecasts into the comprehensive operating earnings dynamics provides more information given the higher adjusted R^2 in the regression. In Panel D, we examine the addition of value-relevant variable, the expected growth in operating assets, in the autoregressive property of the operating assets. Incorporating the expected growth in operating assets as the additional value-relevant variable increased the adjusted R^2 from 0.7941 to 0.8321. Therefore, for the linear information dynamics of the operating assets, the expected growth in operating assets further improves the explanatory of the model. In summary, we find that similar to the Ohlson Model, the value-relevant variables incorporated in the FO Model provide additional information beyond the accounting variables. We next examine how different specification of these linear information dynamics affects the implied pricing function in evaluating the stock prices.

87.4.3 Estimation of Linear Information Dynamics

We start our empirical analysis by estimating the linear information dynamics using both the single equation estimation and simultaneous equation estimation in the Ohlson Model and FO Model.[8] We then use these estimated coefficients along with the observed inputs in the implied pricing functions to compute the theoretical price of the shares. Our conjecture is that given there exist feedback relations between the accounting variables and the value-relevant information variables, the simultaneous equation estimation more accurately estimates the linear information dynamics, and thus, the resulting pricing function produces smaller pricing errors than those under the single equation estimation.

Panel A1 in Table 87.4 provides the estimated coefficients from both the single equation specification and the simultaneous equation specification of the Ohlson Model information dynamics with the analyst earnings forecasts. The single

[8]The single equation estimation is conducted by the panel GMM estimator as in Table 87.2A and B. Since our system of simultaneous equation specification of information dynamics involves endogenous regressors from other equations, we use the more efficient error-component three-stage least square (3SLS) estimator proposed by Baltagi (1981) to conduct the estimation. Essentially, the 3SLS is a combination of the two-stage least square (2SLS) estimator and the seemingly unrelated regression (SUR) estimator. 3SLS considers both the simultaneous equation bias and the cross equation correlation of the errors.

Table 87.4 Estimation of the linear information dynamics: single equation and simultaneous equation estimation

Panel A1: Ohlson Model with analyst earnings forecasts				
Est. coeff.	Single equation		Simultaneous equations	
	$\widetilde{x}^a_{i,t+1} = \omega_{10} + \omega_{11} x^a_{i,t} + \omega_{12} v_{i,t} + \widetilde{\varepsilon}_{1i,t+1}$ $\widetilde{v}_{i,t+1} = \omega_{20} + \omega_{22} v_{i,t} + \widetilde{\varepsilon}_{2i,t+1}$		$\widetilde{x}^a_{i,t+1} = \omega_{10} + \omega_{11} x^a_{i,t} + \omega_{12} v_{i,t} + \widetilde{\varepsilon}_{1i,t+1}$ $\widetilde{v}_{i,t+1} = \omega_{20} + \omega_{21} x^a_{i,t} + \omega_{22} v_{i,t} + \widetilde{\varepsilon}_{2i,t+1}$	
$\boldsymbol{\omega_{10}}$	0.0524	(<0.0001)	0.1431	(<0.0001)
$\boldsymbol{\omega_{11}}$	0.6054	(<0.0001)	0.4005	(<0.0001)
$\boldsymbol{\omega_{12}}$	0.3052	(0.0002)	0.3341	(0.0449)
$\boldsymbol{\omega_{20}}$	0.1754	(<0.0001)	0.1301	(0.0088)
$\boldsymbol{\omega_{21}}$			0.3791	(<0.0001)
$\boldsymbol{\omega_{22}}$	0.4185	(<0.0001)	0.2866	(0.0156)
Panel A2: Ohlson Model (x_{130}) with analyst earnings forecasts				
Est. coeff.	Single equation		Simultaneous equations	
	$\widetilde{x}^a_{130i,t+1} = \omega_{10} + \omega_{11} x^a_{130i,t} + \omega_{12} v_{i,t} + \widetilde{\varepsilon}_{1i,t+1}$ $\widetilde{v}_{i,t+1} = \omega_{20} + \omega_{22} v_{i,t} + \widetilde{\varepsilon}_{2i,t+1}$		$\widetilde{x}^a_{130i,t+1} = \omega_{10} + \omega_{11} x^a_{130i,t} + \omega_{12} v_{i,t} + \widetilde{\varepsilon}_{1i,t+1}$ $\widetilde{v}_{i,t+1} = \omega_{20} + \omega_{21} x^a_{130i,t} + \omega_{22} v_{i,t} + \widetilde{\varepsilon}_{2i,t+1}$	
$\boldsymbol{\omega_{10}}$	0.0492	(0.0003)	0.1534	(0.0103)
$\boldsymbol{\omega_{11}}$	0.6231	(<0.0001)	0.4138	(0.0021)
$\boldsymbol{\omega_{12}}$	0.2878	(<0.0001)	0.3687	(0.0009)
$\boldsymbol{\omega_{20}}$	0.1977	(0.0008)	0.1544	(<0.0001)
$\boldsymbol{\omega_{21}}$			0.4369	(<0.0001)
$\boldsymbol{\omega_{22}}$	0.5013	(<0.0001)	0.2920	(0.0006)
Panel B1: FO Model with analyst earnings forecasts				
Est. coeff.	Single equation		Simultaneous equations	
	$\widetilde{ox}^a_{i,t+1} = \omega_{10} + \omega_{11} ox^a_{i,t} + \omega_{12} oa_{i,t} + \omega_{13} v_{1i,t} + \widetilde{\varepsilon}_{1i,t+1}$ $\widetilde{oa}_{i,t+1} = \omega_{20} + \omega_{22} oa_{i,t} + \omega_{24} v_{2i,t} + \widetilde{\varepsilon}_{2i,t+1}$ $\widetilde{v}_{1i,t+1} = \omega_{30} + \omega_{33} v_{1i,t} + \widetilde{\varepsilon}_{3i,t+1}$ $\widetilde{v}_{2i,t+1} = \omega_{40} + \omega_{44} v_{1i,t} + \widetilde{\varepsilon}_{4i,t+1}$		$\widetilde{ox}^a_{i,t+1} = \omega_{10} + \omega_{11} ox^a_{i,t} + \omega_{12} oa_{i,t} + \omega_{13} v_{1i,t} + \omega_{14} v_{2i,t} + \widetilde{\varepsilon}_{1i,t+1}$ $\widetilde{oa}_{i,t+1} = \omega_{20} + \omega_{21} ox^a_{i,t} + \omega_{22} oa_{i,t} + \omega_{23} v_{1i,t} + \omega_{24} v_{2i,t} + \widetilde{\varepsilon}_{2i,t+1}$ $\widetilde{v}_{1i,t+1} = \omega_{30} + \omega_{31} ox^a_{i,t} + \omega_{32} oa_{i,t} + \omega_{33} v_{1i,t} + \omega_{34} v_{2i,t} + \widetilde{\varepsilon}_{3i,t+1}$ $\widetilde{v}_{2i,t+1} = \omega_{40} + \omega_{41} ox^a_{i,t} + \omega_{42} oa_{i,t} + \omega_{43} v_{1i,t} + \omega_{44} v_{1i,t} + \widetilde{\varepsilon}_{4i,t+1}$	
$\boldsymbol{\omega_{10}}$	0.0708	(0.1161)	0.0222	(0.0015)

(*continued*)

Table 87.4 (continued)

$\boldsymbol{\omega_{11}}$	0.2886	(<0.0001)	0.4241	(<0.0001)
$\boldsymbol{\omega_{12}}$	0.2744	(0.0151)	0.1287	(0.0003)
$\boldsymbol{\omega_{13}}$	0.1125	(0.0278)	0.0204	(0.0234)
$\boldsymbol{\omega_{14}}$			0.2267	(0.0459)
$\boldsymbol{\omega_{20}}$	0.0602	(<0.0001)	0.1269	(<0.0001)
$\boldsymbol{\omega_{21}}$			0.1197	(0.0018)
$\boldsymbol{\omega_{22}}$	0.2204	(<0.0001)	0.1066	(<0.0001)
$\boldsymbol{\omega_{23}}$			0.0803	(0.0486)
$\boldsymbol{\omega_{24}}$	0.3052	(0.0402)	0.1872	(0.0339)
$\boldsymbol{\omega_{30}}$	−0.0436	(0.6591)	−0.0382	(0.7187)
$\boldsymbol{\omega_{31}}$			0.2725	(0.0016)
$\boldsymbol{\omega_{32}}$			0.1762	(0.0432)
$\boldsymbol{\omega_{33}}$	−0.0587	(<0.0001)	−0.0143	(0.0287)
$\boldsymbol{\omega_{34}}$			0.0508	(<0.0001)
$\boldsymbol{\omega_{40}}$	0.6493	(0.0004)	0.6244	(0.0008)
$\boldsymbol{\omega_{41}}$			0.0331	(0.7453)
$\boldsymbol{\omega_{42}}$			0.3895	(0.0108)
$\boldsymbol{\omega_{43}}$			−0.2142	(<0.0001)
$\boldsymbol{\omega_{44}}$	0.6685	(<0.0001)	0.6947	(<0.0001)
Panel B2: FO Model (ox_{130}) with analyst earnings forecasts				
Est. coeff.	Single equation		Simultaneous equations	
	$\widetilde{ox}^{a}_{130i,t+1} = \omega_{10} + \omega_{11} ox^{a}_{130i,t} + \omega_{12} oa_{i,t} + \omega_{13} v_{1i,t} + \widetilde{\varepsilon}_{1i,t+1}$ $\widetilde{oa}_{i,t+1} = \omega_{20} + \omega_{22} oa_{i,t} + \omega_{24} v_{2i,t} + \widetilde{\varepsilon}_{2i,t+1}$ $\widetilde{v}_{1i,t+1} = \omega_{30} + \omega_{33} v_{1i,t} + \widetilde{\varepsilon}_{3i,t+1}$ $\widetilde{v}_{2i,t+1} = \omega_{40} + \omega_{44} v_{1i,t} + \widetilde{\varepsilon}_{4i,t+1}$		$\widetilde{ox}^{a}_{130i,t+1} = \omega_{10} + \omega_{11} ox^{a}_{130i,t} + \omega_{12} oa_{i,t} + \omega_{13} v_{1i,t} + \omega_{14} v_{2i,t} + \widetilde{\varepsilon}_{1i,t+1}$ $\widetilde{oa}_{i,t+1} = \omega_{20} + \omega_{21} ox^{a}_{130i,t} + \omega_{22} oa_{i,t} + \omega_{23} v_{1i,t} + \omega_{24} v_{2i,t} + \widetilde{\varepsilon}_{2i,t+1}$ $\widetilde{v}_{1i,t+1} = \omega_{30} + \omega_{31} ox^{a}_{130i,t} + \omega_{32} oa_{i,t} + \omega_{33} v_{1i,t} + \omega_{34} v_{2i,t} + \widetilde{\varepsilon}_{3i,t+1}$ $\widetilde{v}_{2i,t+1} = \omega_{40} + \omega_{41} ox^{a}_{130i,t} + \omega_{42} oa_{i,t} + \omega_{43} v_{1i,t} + \omega_{44} v_{1i,t} + \widetilde{\varepsilon}_{4i,t+1}$	
$\boldsymbol{\omega_{10}}$	0.0345	(0.0845)	0.0301	(0.0012)
$\boldsymbol{\omega_{11}}$	0.3654	(0.0003)	0.5011	(<0.0001)

$\boldsymbol{\omega_{12}}$	0.3054	(0.0021)	0.1512	(<0.0001)
$\boldsymbol{\omega_{13}}$	0.0788	(0.0122)	0.0113	(0.0212)
$\boldsymbol{\omega_{14}}$			0.2548	(0.0341)
$\boldsymbol{\omega_{20}}$	0.0987	(<0.0001)	0.1054	(<0.0001)
$\boldsymbol{\omega_{21}}$			0.1364	(0.0002)
$\boldsymbol{\omega_{22}}$	0.2541	(<0.0001)	0.1278	(<0.0001)
$\boldsymbol{\omega_{23}}$			0.0811	(0.0387)
$\boldsymbol{\omega_{24}}$	0.3654	(<0.0001)	0.2013	(0.0498)
$\boldsymbol{\omega_{30}}$	0.0024	(0.0874)	−0.0541	(0.2141)
$\boldsymbol{\omega_{31}}$			0.2931	(0.0004)
$\boldsymbol{\omega_{32}}$			0.2348	(0.0632)
$\boldsymbol{\omega_{33}}$	−0.1021	(0.1654)	−0.0354	(<0.0001)
$\boldsymbol{\omega_{34}}$			0.1087	(<0.0001)
$\boldsymbol{\omega_{40}}$	0.7113	(<0.0001)	0.7146	(0.0009)
$\boldsymbol{\omega_{41}}$			0.0663	(0.2227)
$\boldsymbol{\omega_{42}}$			0.4873	(<0.0001)
$\boldsymbol{\omega_{43}}$			−0.3561	(<0.0001)
$\boldsymbol{\omega_{44}}$	0.6987	(<0.0001)	0.7328	(<0.0001)

In both the Ohlson and FO Model, panel GMM methodology proposed by Arellano and Bond (1991) is used to estimate the single equation estimation of the linear information dynamics. The simultaneous equation estimation of the linear information dynamics is estimated by the error-component three-stage least square (3SLS) estimator proposed by Baltagi (1981). $x^a_{i,t}$ is the abnormal earnings per share for firm i at year t. $x^a_{130i,t}$ is the abnormal comprehensive earnings per share for firm i at time t. $v_{1i,t}$ in the Ohlson Model is the consensus analyst earnings per share forecasts for year t+1 at year t. $ox^a_{i,t}$ is the abnormal operating earnings per share for firm i at year t. $ox^a_{130i,t}$ is the abnormal comprehensive operating earnings per share for firm i at year t. $oa_{i,t}$ is operating asset scaled by total assets for firm i at year t. $v_{1i,t}$ in the FO Model is the growth in expected operating earnings, measured as the difference between the median consensus analyst earnings forecast for next year and the expected net interest revenue (product of end-of-current-year financial liabilities and the interest on debt). $v_{2i,t}$ is the expected growth in operating assets, measured by the analyst earnings forecasts of long-term earnings growth rate. The p-values associated with each statistics are reported in the parenthesis to the right of the estimated coefficients. Our sample contains all available data in Compustat, CRSP, and I/B/E/S from 1980 to 2008 equation specification is essentially those shown in Panel B and C1 in Table 87.2 by using the panel GMM estimator. Following Tsay et al. (2008), we conjecture that there is a feedback relation between the current-period abnormal earnings x_{t+1}^{a} and current-period analyst forecasts for the next-period earnings $v_{i,t+1}$. Compared to the single equation specification, the simultaneous equation specification also estimates the feedback effect from the abnormal earnings to the analyst earnings forecasts. By jointly estimating the coefficient, we found that the coefficients ω_{21} are statistically significant (0.3791), indicating that the abnormal earnings indeed affect the analyst earnings forecasts for next period. In Panel A2 of Table 87.4, we estimate the linear information dynamics in the Ohlson Model by using the comprehensive earnings dynamics. The estimated ω_{21} is 0.4369 which is higher than 0.3791 in Panel A1 in which bottom-line earnings are used as the earnings measure in the information dynamics. This result suggests that abnormal comprehensive earnings provide a stronger feedback effect to the analyst earnings forecasts while the other estimated coefficients retain their statistical and economic magnitude. The statistical significant ω_{21} suggests that the single equation specification in traditional Ohlson Model linear information dynamics is not correctly identified and the simultaneous equation estimation of the linear information dynamics might yield more price-relevant information in forecasting future stock price.

Panel B1 in Table 87.4 compares the single equation and simultaneous equation specification of the linear information dynamic with bottom-line earnings in the FO Model. In the single equation specification, all the coefficients associated with the accounting variables and the value-relevant information variables are statistically significant. This indicates that the linear information dynamics in the FO Model indeed possess empirical content to capture the variation in abnormal operating earnings and operating assets. We further extended the single equation specification to simultaneous equation specification to examine whether there exist feedback relations between the accounting variables and the value-relevant information variables. For example, the coefficient ω_{31} which measures how current-period abnormal operating earnings affect the expected growth in abnormal operating earnings is statistically significant of 0.2725. Moreover, the coefficient ω_{42}, which measures how current-period operating asset affects the next-period expected growth in operating asset, is also statistically significant of 0.3895. These results suggest that there exist feedback relations of abnormal operating earnings and operating assets with their expected growth in the future periods. Estimating the linear information dynamics by the simultaneous equation approach improves the information content provided of these variables in computing the implied value of the shares. In Panel B2 of Table 87.4, we use comprehensive operating earnings to estimate the linear information dynamics in the FO Model. Compared to the results in Panel B1, we find stronger feedback effects given the larger estimated coefficients ω_{31} (0.2931) and ω_{42} (0.4873) under the estimation with abnormal comprehensive operating earnings. In the next section, we will further examine the pricing errors of the implied valuation function by employing these estimated coefficients with the observed inputs.

87.4.4 Prediction Errors of Stock Price Forecasts

Table 87.5 provides the summary of prediction errors of stock prices from the Ohlson Model and FO Model using different estimation methods for linear information dynamics (single equation vs. simultaneous equation estimation) and different earnings measures (bottom-line earnings vs. comprehensive earnings). More specifically, we estimated the coefficients in the linear information dynamics ω_{ij} and used them with the observed inputs abnormal earnings and analyst earnings forecasts to compute the theoretical price of the share at end of each year t given the implied valuation function for the Ohlson Model in Eq. 87.14 and for the FO Model in Eq. 87.18. We then measure how these implied values of the share differ from the observed current market price per share, i.e., the prediction errors for the stock prices. The prediction errors are represented by the mean forecast errors, which are calculated as the observed market price per share minus the implied price from the model divided by the market price per share at end of each period t.

In Panel A of Table 87.5, we first observe that simultaneous equation estimation of the linear information dynamics indeed improves the future stock forecast accuracy by producing significantly smaller prediction errors than those generated by the single equation estimation. The prediction error difference, $\Delta_{Simul-Single}$, is significantly negative for both abnormal earnings and abnormal comprehensive earnings at -0.0514 and -0.0305, respectively. We then discuss whether using abnormal comprehensive earnings can improve the stock price forecast ability of each individual model. Under both single equation and simultaneous equation estimation specifications, we calculate the forecast error differences between using abnormal earnings and comprehensive abnormal earnings as the earnings measures in the information dynamics, i.e., $\Delta_{x^a_{130}-x^a}$. We are expected to observe smaller forecast errors when using abnormal comprehensive earnings because of its consistency with the clean surplus relation which is used in deriving the implied valuation function in the Ohlson Model. The result in Panel A of Table 87.5 shows that the forecast error differences $\Delta_{x^a_{130}-x^a}$ are statistically significant at -0.0334 (-0.0225) under single equation (simultaneous equations) estimation. Our results suggest that under both estimation specifications of the linear information dynamics, abnormal comprehensive earnings outperform the abnormal earnings in terms of predicting future stock prices by generating smaller average forecast errors. Similar improvement in prediction accuracy can also be found in Panel B of Table 87.5 where comprehensive operating earnings are used in estimating the linear information dynamics in the FO Model. The result in Panel B of Table 87.5 suggests that the forecast error differences $\Delta_{ox^a_{130}-ox^a}$ are statistically significant at -0.0244 (-0.0186) under single equation (simultaneous equations) estimation. This indicates that abnormal comprehensive operating earnings provide more value-relevant information than abnormal operating earnings in estimating linear information dynamics and computing the 1-year-ahead model-implied stock prices. In sum, the empirical results we have shown in this chapter further demonstrate that comprehensive (operating) earnings can also produce more accurate future stock price forecasts in the residual valuation models.

Table 87.5 Prediction errors of 1-year-ahead stock prices of the Ohlson and Feltham-Ohlson Model

Panel A: Ohlson Model			
	Single equation	Simultaneous equations	$\Delta_{Simul\text{-}Single}$
	$\widetilde{x}^a_{i,t+1} = \omega_{10} + \omega_{11}x^a_{i,t} + \omega_{12}v_{i,t} + \widetilde{\varepsilon}_{1i,t+1}$ $\widetilde{v}_{i,t+1} = \omega_{20} + \omega_{22}v_{i,t} + \widetilde{\varepsilon}_{2i,t+1}$	$\widetilde{x}^a_{i,t+1} = \omega_{10} + \omega_{11}x^a_{i,t} + \omega_{12}v_{i,t} + \widetilde{\varepsilon}_{1i,t+1}$ $\widetilde{v}_{i,t+1} = \omega_{20} + \omega_{21}x^a_{i,t} + \omega_{22}v_{i,t} + \widetilde{\varepsilon}_{2i,t+1}$	
Abnormal earnings (x^a)	0.5078	0.4564	−0.0514
	(<0.0001)	(<0.0001)	(0.0007)
Abnormal comprehensive earnings (x^a_{130})	0.4744	0.4339	−0.0305
	(<0.0001)	(<0.0001)	(<0.0001)
$\Delta_{x^a_{130}-x^a}$	−0.0334	−0.0225	
	(<0.0001)	(0.0005)	
Panel B: FO Model			
	Single equation	Simultaneous equations	$\Delta_{Simul\text{-}Single}$
	$\widetilde{ox}^a_{130i,t+1} = \omega_{10} + \omega_{11}ox^a_{130i,t} + \omega_{12}oa_{i,t} + \omega_{13}v_{1i,t} + \widetilde{\varepsilon}_{1i,t+1}$ $\widetilde{oa}_{i,t+1} = \omega_{20} + \omega_{22}oa_{i,t} + \omega_{24}v_{2i,t} + \widetilde{\varepsilon}_{2i,t+1}$ $\widetilde{v}_{1i,t+1} = \omega_{30} + \omega_{33}v_{1i,t} + \widetilde{\varepsilon}_{3i,t+1}$ $\widetilde{v}_{2i,t+1} = \omega_{40} + \omega_{44}v_{1i,t} + \widetilde{\varepsilon}_{4i,t+1}$	$\widetilde{ox}^a_{130i,t+1} = \omega_{10} + \omega_{11}ox^a_{130i,t} + \omega_{12}oa_{i,t} + \omega_{13}v_{1i,t} + \omega_{14}v_{2i,t} + \widetilde{\varepsilon}_{1i,t+1}$ $\widetilde{oa}_{i,t+1} = \omega_{20} + \omega_{21}ox^a_{130i,t} + \omega_{22}oa_{i,t} + \omega_{23}v_{1i,t} + \omega_{24}v_{2i,t} + \widetilde{\varepsilon}_{2i,t+1}$ $\widetilde{v}_{1i,t+1} = \omega_{30} + \omega_{31}ox^a_{130i,t} + \omega_{32}oa_{i,t} + \omega_{33}v_{1i,t} + \omega_{34}v_{2i,t} + \widetilde{\varepsilon}_{3i,t+1}$ $\widetilde{v}_{2i,t+1} = \omega_{40} + \omega_{41}ox^a_{130i,t} + \omega_{42}oa_{i,t} + \omega_{43}v_{1i,t} + \omega_{44}v_{1i,t} + \widetilde{\varepsilon}_{4i,t+1}$	

Abnormal operating earnings (ox^a)	0.4059	0.3745	−0.0314
	(<0.0001)	(<0.0001)	(<0.0001)
Abnormal comprehensive operating earnings (ox^a_{130})	0.3815	0.3559	−0.0256
	(<0.0001)	(<0.0001)	(<0.0001)
$\Delta_{ox^a_{130}-ox^a}$	−0.0244	−0.0186	
	(0.0138)	(0.0159)	
$\Delta_{\text{FO-Ohlson}}$	−0.1019	−0.0818	
	(<0.0001)	(<0.0001)	

In both the Ohlson and FO Model, panel GMM methodology proposed by Arellano and Bond (1991) is used for single equation estimation of the linear information dynamics. The simultaneous equation estimation of the linear information dynamics is estimated by the error-component three-stage least square (3SLS) estimator proposed by Baltagi (1981). The estimated coefficients ω_{ij} are used with the observed inputs abnormal earnings and analyst earnings forecasts to compute the estimated price of the share at end of each year t. The mean forecast errors ($M.F.E.$) are the average of the difference between observed market price and model estimated price divided by the market price per share at end of each period t. $\Delta_{FO\text{-}Ohlson}$ represents the difference of forecast errors between the Ohlson Model and FO Model. $\Delta_{Simul\text{-}Single}$ represents the difference of forecast errors between simultaneous equation estimation and single equation estimation of the linear information dynamics in a given valuation model. $\Delta_{x^a_{130}-x^a}$ represents the Ohlson Model forecast error difference between using earnings and comprehensive earnings as the earnings measure. $\Delta_{ox_{130}-ox}$ represents the FO Model forecast error difference between using operating earnings and comprehensive operating earnings as the earnings measure. The p-values associated with each estimated coefficients are reported in the parenthesis below. Our sample contains all available data in Compustat, CRSP, and I/B/E/S from 1980 to 2008

Table 87.6 Prediction errors of stock price forecast of the Warren-Shelton Model

Number-of-years-ahead forecasts	M.F.E.	$\Delta_{WS\ -\ Ohlson}$	$\Delta_{WS\ -\ FO}$	$\Delta_{WS-Ohlson(x_{130})}$	$\Delta_{WS-FO(ox_{130})}$
1	0.3418	−0.1146	−0.0327	−0.0921	−0.0141
	(<0.0001)	(<0.0001)	(0.0050)	(<0.0001)	(0.0123)
2	0.4132	−0.0751	−0.0157	−0.0051	−0.0513
	(0.0003)	(0.0451)	(0.0122)	(0.0741)	(0.0523)
3	0.5981	0.0354	−0.0084	−0.0003	−0.0024
	(<0.0001)	(0.3214)	(0.0645)	(0.1522)	(0.0841)
4	0.7016	−0.0641	0.0123	0.0874	0.0845
	(0.0005)	(0.0354)	(0.2147)	(0.0987)	(0.1137)
5	0.9142	0.0687	−0.0011	0.0783	0.0746
	(<0.0001)	(0.3329)	(0.2457)	(0.0874)	(0.1068)

Table 87.6 summarizes the mean forecast errors (*M.F.E.*) for 1–5-year-ahead stock price forecasts of the WS Model. From 1980 to 2002, financial data of each firm are used as the base year information in the WS Model to forecast the future-period stock prices in the next 5 years. For each firm at year *t*, the sales growth rate is estimated by the linear regression model using all the past sales available in Compustat. We then calculated the mean forecast errors of the stock prices by each number-of-years-ahead forecast errors across all rolling periods. $\Delta_{WS\ -\ Ohlson}$ represents the difference of forecast errors between the WS Model and Ohlson Model with its linear information dynamics estimated by simultaneous equation method. $\Delta_{WS-Ohlson(x_{130})}$ represents the difference of forecast errors between the WS Model and Ohlson Model with comprehensive earnings as its earnings measure and its linear information dynamics estimated by simultaneous equation method. $\Delta_{WS\ -\ FO}$ represents the difference of forecast errors between the WS Model and FO Model with its linear information dynamics estimated by simultaneous equation method. $\Delta_{WS-FO(ox_{130})}$ represents the difference of forecast errors between the WS Model and FO Model with comprehensive operating earnings as its earnings measure and its linear information dynamics estimated by simultaneous equation method. The p-values associated with each estimated coefficients are reported in the parenthesis below. Our sample contains all available data in Compustat, CRSP, and I/B/E/S from 1980 to 2008

Table 87.6 summarizes the mean forecast errors for 1- to 5-year forecasts of the WS Model. At each year from 1980 to 2002, financial data of each firm are used as the base year information in the WS Model to forecast the future-period stock prices in the next 5 years. In particular, for each firm at year *t*, the sales growth rate is estimated by the linear regression model using all the past sales available in Compustat. We then calculated the mean forecast errors of future stock prices by each number-of-years-ahead forecast errors across all rolling periods. We note that the mean forecast errors mostly increase monotonically with number-of-years-ahead forecasts. We compare WS Model forecasts to Ohlson/FO Model forecasts with both bottom-line earnings and comprehensive earnings. We calculate the mean forecast error difference between the WS Model and Ohlson Model with bottom-line earnings ($\Delta_{WS\ -\ Ohlson}$) and comprehensive earnings ($\Delta_{WS-Ohlson(x_{130})}$). We also calculate the mean forecast error difference between the WS Model and FO Model with operating earnings ($\Delta_{WS\ -\ FO}$) and comprehensive operating earnings ($\Delta_{WS-FO(ox_{130})}$). Our results suggest that the WS Model produces more accurate stock prices than the Ohlson/FO Model in the shorter term. The average

1-year-ahead stock price forecast error is 0.3418 which is significantly lower than those forecasted by the Ohlson and FO Model using either bottom-line earnings or comprehensive earnings. The mean forecast error differences are all significantly negative for the 1-year forecast in Table 87.6. However, in the longer term, the model produces relatively less accurate forecasts. Our findings suggest that the WS Model considering the interrelations between investment activity, financing activity, dividend policy, and the production decision of the firm provides better forecast accuracy in terms of the stock prices than the residual income valuation models discussed in the previous sections. Nonetheless, the forecast accuracy crucially depends on the model inputs such as sales growth rate, current assets as a percent of sales, fixed asset as a percent of sales, dividend payout ratio, and the leverage ratio. Therefore, we next examined how the sensitivity of these model inputs affects the resulting forecasts.

Table 87.7 summarizes the sensitivity analysis of several model inputs in the WS Model, i.e., sales growth rate, total assets as a percent of sales, dividend payout ratio, and leverage ratio. In Panel A, we showed how sensitive the stock price forecasts are to the changes in the sales growth rate. It is expected that higher sales growth rate leads to higher stock prices because of higher future earnings. When the sales growth rate changed from its median to its third quartile, the 1-year-ahead mean forecast errors reduced from 0.4238 to 0.3954. Since the WS Model underestimates the stock prices as shown in Table 87.6, increases in sales growth rate result in higher stock prices and thus smaller pricing errors. When the sales growth rate decreases from its median to the first quartile, the implied stock prices decrease and pricing errors increase accordingly. Similar patterns of the changes of the pricing errors can be observed in longer years ahead forecasts. Furthermore, the sensitivity analysis regarding the total assets as a percent of sales is shown in Panel B. The stock prices are expected to decrease in assets as a percent of sales because more equity funds were required to support asset requirement and thus larger the pricing errors. The base case inputs for the total assets as a percent of the sales are the median value of each firm's available historical data. We examined the sensitivity of the stock price forecast to the first and third quartile value of the firm's total asset as a percent of sales. The mean forecast errors of 1-year-ahead stock prices increased (decreased) to 0.4756 (0.4048) from 0.4238 if the third (first) quartile of total assets as a percent of sales is used. Similar patterns are observed in 2- and 3-year-ahead forecasts but results for longer-year-ahead forecasts are not obvious.

Panel C summarizes the sensitivity of stock price forecasts to the firm's payout ratio. We expect the pricing errors to decrease in the dividend payout ratio because as more dividends were paid to shareholders, the firm relies more on new issues of common stocks for the financing requirement which leads to potential decline in the stock prices. The mean forecast errors increased (decreased) to 0.4451 (0.4198) if the third (first) quartile of the firm's historical dividend payout ratio is used as the model input. Panel D summarized the sensitivity of stock prices to the changes in the firm's leverage. We expect the stock prices to increase in leverage and thus smaller pricing errors. The mean forecast errors decrease (increase) to 0.4187 (0.4408) if the third (first) quartile of the historical leverage ratio is used as the

Table 87.7 Sensitivity analysis of stock price forecast errors of the Warren-Shelton Model

	Q1		Median		Q3	
Number-of-years-ahead forecasts	M.F.E.	V/P	M.F.E.	V/P	M.F.E.	V/P
Panel A: Sales growth rate						
1	0.6101	0.5436	0.4238	0.6381	0.2954	0.6741
2	0.7187	0.4981	0.5002	0.5402	0.365	0.5632
3	0.8354	0.4026	0.5841	0.4783	0.4899	0.5271
4	0.9163	0.3156	0.7136	0.4297	0.5153	0.5012
5	1.0125	0.2103	0.9362	0.3882	0.7041	0.4132
Panel B: Total asset as a percent of sales						
1	0.4048	0.6541	0.4238	0.6381	0.4756	0.6018
2	0.5015	0.5564	0.5011	0.5344	0.2321	0.4987
3	0.6136	0.4541	0.5841	0.4783	0.6328	0.4655
4	0.6654	0.4465	0.7136	0.4297	0.7843	0.4415
5	0.8972	0.4125	0.9362	0.3882	0.9637	0.4213
Panel C: Dividend payout ratio						
1	0.4198	0.6608	0.4238	0.6381	0.4451	0.6231
2	0.4936	0.5412	0.4981	0.5138	0.5215	0.5155
3	0.5787	0.5121	0.5841	0.4783	0.6123	0.4338
4	0.6897	0.4511	0.7136	0.4297	0.7122	0.4136
5	0.8987	0.4015	0.9362	0.3882	0.9345	0.3788
Panel D: Leverage ratio						
1	0.4408	0.5389	0.4238	0.6381	0.4187	0.6658
2	0.5321	0.5128	205144	0.5312	0.5108	0.5974
3	0.6087	0.4569	0.5841	0.4783	0.5941	0.4568
4	0.7136	0.4158	0.7136	0.4297	0.6543	0.4412
5	0.9302	0.4111	0.9362	0.3882	0.8741	0.4123

Table 87.7 summarizes the sensitivity analysis of the model inputs: sales growth rate, total assets as a percent of sales, dividend payout ratio, and leverage ratio in the WS Model. We calculate the mean forecast errors (*M.F.E.*) of stock price forecasts from WS Model by using the first quartile, median, and the third quartile of the respective model inputs. The p-values associated with each estimated coefficients are reported in the parenthesis below. Our sample contains all available data in Compustat, CRSP, and I/B/E/S from 1980 to 2008

model input in forecasting future stock prices. The longer-year-ahead forecasts are again not as apparent as the shorter-term forecasts suggesting the model's lack of ability of accurate forecast in the long run. In summary, our results suggest that the stock price forecasts are most sensitive to the changes in sales growth rate given the largest change in the mean forecast errors. This is not a surprising result since the starting point of the WS Model is the sales growth estimate. The sensitivity of the stock price forecasts to the other exogenously given variables such as the total assets as a percent of sales, dividend payout ratio, and the leverage ratio does not have a significant impact on the forecasts as the sales growth rate. Overall, the results are consistent with the model conjecture that stock price forecasts are influenced by the aforementioned inputs and sales growth rate is the most important factor in producing the accurate forecasts from the model.

87.4.5 Forecast Combination

In the previous section, the implied valuation functions of the Ohlson Model and FO Model were investigated for their ability in forecasting future stock prices. In this section, we examine how the stock price forecasts from these models can be combined with the Warren-Shelton Model to further improve their forecast ability in terms of prediction errors. More specifically, we employ the forecast combination proposed by Bates and Granger (1969), Granger and Newbold (1973), Granger and Ramanathan (1984), and Diebold and Pauly (1987) to examine the combined predictability of future stock prices from different primary individual forecasts.

We use the different weighting schemes in equations (W1) through (W4) in Sect. 87.2.3 to construct the weighting matrix W in the WLS estimation. Moreover, we employ the linear and quadratic deterministic time-varying parameters model to produce time-varying weights. Similar to the regression model in Eq. 87.12, the estimator can be written as

$$\text{Linear}:\quad {}_t\hat{y}_{t+1}=\left(\hat{p}_0^0+\hat{p}_1^0(t+1)\right)+\left(\hat{p}_0^1+\hat{p}_1^1(t+1)\right){}_tf_{t+1}^{RI}+\left(\hat{p}_0^2+\hat{p}_1^2(t+1)\right){}_tf_{t+1}^{WS}$$

$$\begin{aligned}\text{Quadratic}:\quad {}_t\hat{y}_{t+1}=&\left(\hat{p}_0^0+\hat{p}_1^0(t+1)+\hat{p}_2^0(t+1)^2\right)+\left(\hat{p}_0^1+\hat{p}_1^1(t+1)+\hat{p}_2^1(t+1)^2\right){}_tf_{t+1}^{RI}\\&+\left(\hat{p}_0^2+\hat{p}_1^2(t+1)+\hat{p}_2^2(t+1)^2\right){}_tf_{t+1}^{WS}\end{aligned}$$

where ${}_tf_{t+1}^{RI}$ and ${}_tf_{t+1}^{WS}$ are the 1-year stock price forecast from the residual income valuation models (Ohlson/FO Models) and the WS Model, respectively. We consider the following estimators for forecasting future stock prices:

M1. WLS, geometric weights, linear deterministic time-varying parameters
M2. WLS, geometric weights, quadratic deterministic time-varying parameters
M3. WLS, t^{λ} weights, linear deterministic time-varying parameters
M4. WLS, t^{λ} weights, linear quadratic time-varying parameters
M5. OLS (simple unrestricted regression-based combination)
M6. Variance-covariance combination

Panel A1 (A2) in Table 87.8 provides the prediction errors of 1-year-ahead stock prices from the forecast methods M1 through M6 combining the Ohlson Model forecasts using bottom-line (comprehensive) earnings and WS Model forecasts. We also compare the forecast errors of these combined forecasts to those individual forecasts under the Ohlson Model and WS Model alone. If the combined forecast methods indeed improve the model's ability to provide better accuracy, then the combined methods are expected to produce smaller pricing errors. Our results suggest that method M3 yields the best 1-year-ahead stock price forecast in terms of smallest mean forecast errors. Method M3 employs the t-lambda (t^{λ}) weighting specification with the linear deterministic time-varying parameters in the WLS estimator to generate optimal weights for each individual forecast from the Ohlson Model and WS Model. The M3 forecast combination method produces a mean

Table 87.8 Prediction errors of 1-year-ahead stock prices of combined forecast methods

Forecast methods	M.F.E.	$\Delta_{M_i-M_7}$	$\Delta_{M_i-M_8}$	M.F.E.	$\Delta_{M_i-M_7}$	$\Delta_{M_i-M_8}$
	Panel A1: Ohlson and WS combined			Panel A2: Ohlson(x_{130}) and WS combined		
M1: WLS, λ^t	0.3115	−0.1449	−0.0303	0.2981	−0.1583	−0.0437
(*Lin TVPM*)	(0.0002)	(0.0001)	(<0.0001)	(<0.0001)	(0.0150)	(0.0005)
M2: WLS, λ^t	0.3321	−0.1243	−0.0097	0.3425	−0.1139	0.0007
(*Qd TVPM*)	(0.0003)	(0.0764)	(0.0987)	(<0.0001)	(0.0254)	(0.1325)
M3: WLS, t^λ	0.2841	−0.1723	−0.0577	0.2779	−0.1785	−0.0639
(*Lin TVPM*)	(0.0231)	(0.0429)	(0.0008)	(<0.0001)	(0.0110)	(<0.0001)
M4: WLS, t^λ	0.3014	−0.1550	−0.0404	0.2845	−0.1719	−0.0573
(*Qd TVPM*)	(<0.0001)	(0.0745)	(0.0321)	(<0.0001)	(0.0005)	(0.0152)
M5: OLS	0.3614	−0.0950	0.0196	0.3554	−0.101	0.0136
	(0.0001)	(0.0712)	(0.1123)	(<0.0001)	(0.0656)	(0.0327)
M6: Var-Cov	0.3841	−0.0723	0.0423	0.3551	−0.1013	0.0133
	(0.0001)	(0.0432)	(0.0611)	(<0.0001)	(0.0841)	(0.0674)
M7: f^{Ohlson} alone	0.4564			0.4564		
	(<0.0001)			(<0.0001)		
M8: f^{WS} alone	0.3418			0.3418		
	(<0.0001)			(<0.0001)		
	Panel B1: FO and WS combined			Panel B2: FO(ox_{130}) and WS combined		
M1: WLS, λ^t	0.3152	−0.0593	−0.0266	0.3057	−0.0584	−0.0361
(*Lin TVPM*)	(0.0005)	(0.0421)	(0.0742)	(0.0063)	(0.0205)	(<0.0001)
M2: WLS, λ^t	0.3444	−0.0301	0.0026	0.2874	−0.0767	−0.0544
(*Qd TVPM*)	(<0.0001)	(0.0611)	(0.0887)	(<0.0001)	(0.0050)	(0.0026)
M3: WLS, t^λ	0.2411	−0.1334	−0.1007	0.2331	−0.1310	−0.1087

(*Lin TVPM*)	(0.0008)	(0.0013)	(0.0447)	(<0.0001)	(0.0085)	(0.0154)
M4: WLS, t^{λ}	0.2284	−0.1461	−0.1134	0.2045	−0.1596	−0.1373
(*Qd TVPM*)	(0.0011)	(0.0003)	(0.0011)	(<0.0001)	(0.0035)	(0.0005)
M5: OLS	0.4125	0.0380	0.0707	0.3845	0.0204	0.0427
	(0.0006)	(0.0844)	(0.0334)	(0.0001)	(0.0321)	(0.0085)
M6: Var-Cov	0.3841	0.0096	0.0423	0.4012	0.0371	0.0594
	(<0.0001)	(0.1002)	(0.1245)	(<0.0001)	(0.0509)	(<0.0001)
M7: f^{Ohlson} alone	0.3641			0.3641		
	(0.0002)			(0.0002)		
M8: f^{WS} alone	0.3418			0.3418		
	(<0.0001)			(<0.0001)		

Table 87.8 provides the mean forecast errors (*M.F.E.*) of 1-year-ahead stock price forecast from methods (M1) through (M6) combining the Ohlson Model and WS Model (Panel A1 and A2) and FO Model and WS Model (Panel B1 and B2) forecasts. The linear (*LIN TVPM*) and quadratic (*Qd TVPM*) time-varying parameters model are

Linear: $_t\hat{y}_{t+1} = \left(\hat{p}_0^0 + \hat{p}_1^0(t+1)\right) + \left(\hat{p}_0^1 + \hat{p}_1^1(t+1)\right)_t f_{t+1}^{RI} + \left(\hat{p}_0^2 + \hat{p}_1^2(t+1)\right)_t f_{t+1}^{WS}$

Quadratic: $_t\hat{y}_{t+1} = \left(\hat{p}_0^0 + \hat{p}_1^0(t+1) + \hat{p}_2^0(t+1)^2\right) + \left(\hat{p}_0^1 + \hat{p}_1^1(t+1) + \hat{p}_2^1(t+1)^2\right)_t f_{t+1}^{RI} + \left(\hat{p}_0^2 + \hat{p}_1^2(t+1) + \hat{p}_2^2(t+1)^2\right)_t f_{t+1}^{WS}$

M7 and M8 are the individual forecasts f^{Ohlson} (f^{FO}) and f^{WS} from the Ohlson (FO) Model and WS Model, respectively. $\Delta_{M_i-M_7}$ represents the prediction error differences between the forecast combination method M_i and the individual forecast from Ohlson Model. $\Delta_{M_i-M_8}$ represents the prediction error differences between the forecast combination method M_i and the individual forecast from WS Model. The mean forecast errors (*M.F.E.*) are the average of the difference between observed market price and model estimated price divided by the market price per share at end of each period t. The p-values associated with each estimated coefficients are reported in the parenthesis below. Our sample contains all available data in Compustat, CRSP, and I/B/E/S from 1980 to 2008 forecast error of 0.2841 (0.2779) in Panel A1 (A2) which is significantly lower than the individual forecast from either the Ohlson Model or the WS Model. For the Ohlson Model using comprehensive earnings, the mean forecast error differences between the M3 method and the Ohlson Model (WS Model) $\Delta_{M_i - M_7}$ ($\Delta_{M_i - M_8}$) are significant at -0.1785 (-0.0639), suggesting that forecast combination is indeed lower that the prediction error of individual forecasts. These results show that the WLS estimator with t-lambda (t^{λ}) weighting specification and linear deterministic time-varying parameters generates the lower prediction errors for stock price forecasts than other forecast combination methods and the individual forecasts. Moreover, we find that in general, the forecast combination methods M1 through M4 generate smaller pricing errors than the individual forecasts from the Ohlson Model and WS Model. The OLS (M5) and variance-covariance method (M6) on the other hand do not improve each model's ability in forecasting future stock prices given the insignificant prediction error differences.

In a similar fashion, Panel B1 (B2) in Table 87.8 provides the prediction errors of forecast combination methods M1 through M6 by combining the FO Model forecast using (comprehensive) operating earnings and WS Model forecasts. Compared to the individual forecasts from the FO Model and WS Model, methods M1 through M4 again produce significantly lower prediction errors while M5 and M6 fail to obtain improvement in forecasting 1-year-ahead stock prices in the sample. The M4 forecast combination method produces lowest mean forecast error, i.e., 0.2284 (0.2045) in Panel B1 (B2) among all different forecast combination methods. The mean forecast error generated by M4 is also significantly lower than that produced by either the FO Model or WS Model. For example, in Panel B2, the mean forecast error differences between the M4 method forecast and the FO Model (WS Model) forecast, or $\Delta_{M_i - M_7}$ ($\Delta_{M_i - M_8}$), are significant at -0.1596 (-0.1373). This suggests that forecast combination indeed is lower than the prediction error of individual forecasts. Overall, our findings demonstrate that the appealing features of geometric weighting schemes and time-varying parameters in forecast combination provide superior accuracy in predicting future stock prices. The WLS estimator with geometric weighting schemes and time-varying parameters places more weights on the better forecast technique over time. Given the different structural designs of the residual income valuation models and the WS Model, each of them could provide superior forecast than the other under specific market condition.

87.5 Summary

This chapter investigates the stock price forecast ability of three alternative valuation models, namely, the Ohlson (1995) Model, Feltham and Ohlson (1995) Model, and the Warren and Shelton (1971) Model. In this chapter, we have developed five research hypotheses to test whether different earnings measures, estimation techniques, and combined forecast methods can improve these models' ability in predicting future stock prices. In the first hypothesis, we test whether the FO Model can produce smaller prediction errors for future stock prices than the

Ohlson Model. The second hypothesis examines whether simultaneous equation estimation of the linear information dynamics in the Ohlson and FO Model can generate smaller prediction errors for future stock prices than those produced by single equation estimation. In the third hypothesis, we investigate whether the WS Model can generate more accurate future stock price forecasts than both the Ohlson and FO Model. The fourth hypothesis examines whether combination of individual forecasts from the Ohlson Model, FO Model, and WS Model can produce more accurate future stock price forecast than each individual model. Finally, the fifth hypothesis tests whether the use of comprehensive income, instead of net income, can generate more accurate future stock price forecasts in these valuation models.

We first use simultaneous equation estimation approach to estimate the information dynamics for the Ohlson Model and FO Model and to forecast future stock prices. Our empirical results suggest that the simultaneous equation estimation of the information dynamics improves the ability of the Ohlson Model and FO Model in capturing the dynamic of the abnormal earnings process. The predictability of the 1-year-ahead stock prices is also more accurate under the simultaneous equation estimation in terms of smaller prediction errors. We then use the WS Model to predict stock price per share and find that the WS Model can generate smaller future stock price prediction errors than those predicted by the Ohlson Model and FO Model. These findings indicate a better stock price forecast ability of the WS Model in determining future stock prices. The superior accuracy compared to the Ohlson Model and FO Model is due to the incorporation of both operation and financing decisions of the firms. We also combine these different stock price forecasts by using various time-varying parameters models proposed by Granger and Newbold (1973) and Diebold and Pauly (1987) to examine whether forecast combination provides better prediction accuracy. The combined forecasting methods generally produce more accurate stock price forecasts than those made by individual models.

Previous literature found supporting evidence that comprehensive income can provide price-relevant information beyond bottom-line net income (Cheng et al. 1993; Dhaliwal et al. 1999; O'Hanlon and Pope 1999). Given that the Ohlson Model and FO Model are based on the clean surplus relation, we further investigate the price relevance of comprehensive (operating) income in these valuation models. Our results suggest that using comprehensive (operating) income in the Ohlson (FO) Model can further reduce the prediction errors of future stock price forecasts under both single equation and simultaneous equation estimation of linear information dynamics. Moreover, this superior predictability also leads to smaller prediction errors in the combined forecasting in which Ohlson (FO) Model forecasts are combined with WS Model forecasts. Evidence shown in our study demonstrates that comprehensive (operating) income indeed provides incremental price-relevant information beyond bottom-line net income.

In sum, we investigate the empirical validity in terms of stock price forecast accuracy of alternative equity valuation models. By employing the simultaneous equation estimation and combined forecasting methods, we find that these models can produce higher estimate accuracy in predicting future stock prices. We also found that using comprehensive income can further reduce the prediction errors generated by these valuation models. Our findings contribute to the literature in equity valuation models as well as the setting of accounting standard on reporting comprehensive financial performance.

References

Abarbanell, J. S., & Bushee, B. J. (1997). Fundamental analysis, future earnings, and stock prices. *Journal of Accounting Research, 35*, 1–24.

Anderson, T. W., & Hsiao, C. (1981). Estimation of dynamic models with error components. *Journal of the American Statistical Association, 76*, 589–606.

Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies, 58*, 277–297.

Baltagi, B. H. (1981). Simultaneous equations with error components. *Journal of Econometrics, 17*, 189–200.

Bates, J. M., & Granger, C. W. J. (1969). The combination of forecasts. *Operations Research Quarterly, 20*, 451–468.

Begley, J., & Feltham, G. A. (2002). The relations between market value, earnings forecasts, and reported earnings. *Contemporary Accounting Research, 19*, 1–48.

Beresford, D., Johnson, L., & Reither, C. (1996). Is a second income statement needed? *Journal of Accountancy, 176*, 69–72.

Biddle, G., & Choi, J. H. (2006). Is comprehensive income useful? *Journal of Contemporary Accounting & Economics, 2*(1), 1–32.

Black, F. (1993). Choosing accounting rules. *Accounting Horizons, 7*, 1–17.

Brief, R., & Peasnell, K. (Eds.). (1996). *Clean surplus: A link between accounting and finance*. New York: Garland Publishing.

Brimble, M. & Hodgson, A. (2007). On the intertemporal value relevance of conventional financial accounting in Australia. *Accounting & Finance, 4*, 599–622.

Cahan, S., Courtenay, S., Gronewoller, P., & Upton, D. (2000). Value relevance of mandated comprehensive income disclosures. *Journal of Business, Finance, and Accounting, 27*, 1273–1301.

Callen, J. L., & Morel, M. (2001). Linear accounting valuation when abnormal earnings are AR(2). *Review of Quantitative Finance and Accounting, 16*, 191–203.

Callen, J. L., & Segal, D. (2005). Empirical tests of the Feltham-Ohlson (1995) model. *Review of Accounting Studies, 10*, 409–429.

Chambers, D., Thomas, J. L., Shakespeare, C., & Sougiannis, T. (2007). An evaluation of SFAS No. 130 comprehensive income disclosures. *Review of Accounting Studies, 12*, 557–593.

Cheng, C., Liu, C., & Schaefer, T. (1996). Earnings permanence and the incremental information content of cash flows from operations. *Journal of Accounting Research, 34*, 173–181.

Chong, Y. Y., & Hendry, D. F. (1986). Econometric evaluation of linear macroeconomic models. *Review of Economic Studies, 53*, 671–690.

Dechow, P. M., Hutton, A. P., & Sloan, R. G. (1999). An empirical assessment of the residual income valuation model. *Journal of Accounting and Economics, 26*, 1–34.

Dhaliwal, D., Subramanyam, K., & Trezevant, R. (1999). Is comprehensive income superior to net income as a measure of firm performance? *Journal of Accounting and Economics, 26*, 43–67.

Diebold, F. X., & Lopez, J. A. (1996). Forecast evaluation and combination. In G. S. Maddala & C. R. Rao (Eds.), *Handbook of statistics* (pp. 241–268). Amsterdam: North-Holland.

Diebold, F. X., & Pauly, P. (1987). Structural change and the combination of forecasts. *Journal of Forecasting, 6*, 21–40.

Fair, R. C., & Shiller, R. J. (1989). The informational content of ex ante forecasts. *Review of Economics and Statistics, 71*, 325–331.

Fair, R. C., & Shiller, R. J. (1990). Comparing information in forecasts from econometric models. *American Economic Review, 80*, 375–389.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance, 25*, 383–417.

Fama, E. F. (1991). Efficient markets II. *Journal of Finance, 46*, 1575–1617.

Fama, E. F., & French, K. R. (1988). Permanent and temporary components of stock prices. *Journal of Political Economy, 96*, 246–273.

Feltham, G. A., & Ohlson, J. A. (1995). Valuation and clean surplus accounting for operating and financial activities. *Contemporary Accounting Research, 11*(2), 689–731.

Granger, C. W. J., & Newbold, P. (1973). Some comments on the evaluation of economic forecasts. *Applied Economics, 5*, 35–47.

Granger, C. W. J., & Ramanathan, R. (1984). Improved methods of forecasting. *Journal of Forecasting, 3*, 197–204.

Hansen, L. P., & Hodrick, R. J. (1980). Forward exchange rates as optimal predictors of future spot rates: An econometric investigation. *Journal of Political Economy, 88*, 829–853.

Holthausen, R., & Watts, R. (2001). The relevance of the value relevance literature for financial accounting setting. *Journal of Accounting and Economics, 31*, 3–75.

Johnson, L., Reither, C., & Swieringa, R. (1995). Toward reporting comprehensive income. *Accounting Horizons, 9*, 128–137.

Lehman, B. (1993). Earnings, dividend policy, and present value relations: Building blocks of dividend policy invariant cash flows. *Review of Quantitative Finance and Accounting, 3*, 263–282.

LeRoy, S. F., & Porter, R. D. (1981). The present value relation: Tests based on implied variance bounds. *Econometrica, 49*, 555–574.

Lin, S. (2006). Testing the information set perspective of UK financial reporting standard No. 3: Reporting financial performance. *Journal of Business Finance & Accounting, 33*(7 & 8), 1110–1141.

Linsmeier, T., Gribble, J., Jennings, R., Lang, M., Penman, S., Petroni, K., Shores, D., Smith, J., & Warfield, T. (1997a). Response to FASB exposure draft, "Proposed statement of financial accounting standards – Reporting comprehensive income". *Accounting Horizons, 11*, 117–119.

Linsmeier, T., Gribble, J., Jennings, R., Lang, M., Penman, S., Petroni, K., Shores, D., Smith, J., & Warfield, T. (1997b). An issues paper on comprehensive income. *Accounting Horizons, 11*, 120–126.

Liu, J., & Ohlson, J. A. (2000). The Feltham–Ohlson (1995) Model: Empirical implications. *Journal of Accounting, Auditing and Finance, 15*, 321–331.

Liu, J., Nissim, D., & Thomas, J. (2002). Equity valuation using multiples. *Journal of Accounting Research, 40*(1), 135–172.

Myers, J. N. (1999). Implementing residual income valuation with linear information dynamics. *Accounting Reviews, 74*, 1–28.

Nissim, D. & Penman, S. H. (2001). Ratio analysis and equity valuation: From research to practice. *Review of Accounting Studies, 6*, 109–154.

O'Hanlon, J., & Pope, P. (1999). The value-relevance of U.K. dirty surplus accounting flows. *British Accounting Review, 31*, 459–482.

Ohlson, J. A. (1995). Earnings, book values, and dividends in equity valuation. *Contemporary Accounting Research, 11*(2), 661–687.

Ohlson, J. A. (2001). Earnings, book values, and dividends in equity valuation: An empirical perspective. *Contemporary Accounting Research, 18*(1), 107–120.

Ohlson, J. A., & Zhang, X. J. (1998). Accrual accounting and equity valuation. *Journal of Accounting Research, 36*(3), 85–111.

Penman, S. H. (2000). *Financial statement analysis and security valuation*. New York: McGraw-Hill.

Robinson, L. (1991). The time has come to report comprehensive income. *Accounting Horizons, 5*, 107–112.

Shiller, R. J. (1979). The volatility of long term interest rates and expectations models of the term structure. *Journal of Political Economy, 87*, 1190–1219.

Smith, P., & Reither, C. (1996). Comprehensive income and the effect of reporting it. *Financial Analysts Journal, 52*, 14–19.

Tsay, R. S., Lin, Y. M., & Wang, H. W. (2008). Residual income, value-relevant information and equity valuation: A simultaneous equations approach. *Review of Quantitative Finance and Accounting, 31*, 331–358.

Warren, J. M., & Shelton, J. P. (1971). A simultaneous equation approach to financial planning. *Journal of Finance, 26*(5), 1123–1142.

Yee, K. K. (2008). A Bayesian framework for combining valuation estimates. *Review of Quantitative Finance and Accounting, 30*, 339–354.

Time Series Models to Predict the Net Asset Value (NAV) of an Asset Allocation Mutual Fund VWELX

88

Kenneth D. Lawrence, Gary Kleinman, and Sheila M. Lawrence

Contents

Abstract

This research examines the use of various forms of time series models to predict the total NAV of an asset allocation mutual fund. In particular, the mutual fund case used is the Vanguard Wellington Fund. This fund maintains a balance between relatively conservative stocks and bonds. The period of the study on

K.D. Lawrence (✉)
New Jersey Institute of Technology, Newark, NJ, USA
e-mail: carpetfour@yahoo.com

G. Kleinman
Montclair State University, Montclair, NJ, USA
e-mail: gklnjhn@gmail.com

S.M. Lawrence
Rutgers, The State University of New Jersey, New Brunswick, NJ, USA
e-mail: lawrencesm5@aol.com

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_88,

which the prediction of the total NAV is based is the 24-month period of 2010 and 2011, and the forecasting period is the first 3 months of 2012. Forecasting the total NAV of a massive conservative allocation fund, composed of an extremely large number of investments, requires a method that produces accurate result. Achieving this accuracy has no necessary relationship to the complexity of the variety of the methods typically present in many financial forecasting studies.

Various types of methods and models were used to predict the total NAV of the Vanguard Wellington Fund. The first set of model structures included simple exponential smoothing, double exponential smoothing, and Winter's method of smoothing. The second set of predictive models used represented trend models. They were developed using regression estimation. They included linear trend model, quadratic trend model, and an exponential model. The third type of method used was a moving-average method. The fourth set of models incorporated the Box-Jenkins method, including an autoregressive model, a moving-average model, and an unbounded autoregressive and moving-average method.

Keywords
NAV of a mutual fund • Asset allocation fund • Combination of forecasts • Single exponential smoothing • Double exponential smoothing • Winter's method • Linear trend model • Quadratic trend model • Exponential trend model • Moving-average method • Autoregressive model • Moving-average model • Unbounded autoregressive moving-average model

88.1 Introduction

The purpose of this study is to develop a predictive time series model for the total NAV of a massive balanced asset allocation mutual fund during a period of time when there is not a massive decline in the economy. The historical period chosen was the 24-month period beginning January of 201 and running through December 2011. The forecasting period is the first 3 months of 2012.

The forecast of mutual funds with vast numbers of investments is certainly not the same as forecasting a single investment or a group of like investments. Forecasting net asset values of an investment structure consisting of a massive asset allocation of stocks in various industry groups, various types of bond investments, as well as both domestic and international investments presents a specialized type of financial forecasting problems.

88.1.1 NAV of a Mutual Fund

A mutual fund is an investment vehicle that operates as an investment pool. Initial investors put up prearranged amounts and issued shares of the mutual fund, with these shares representing their ownership interest. After the initial issuance, more investors can buy into the mutual fund by buying shares at the current net asset value, or total NAV. This may be more or less than the original total NAV, depending on how the investments of the mutual fund have performed.

NAV is simply the current total value of the assets of the fund minus any liabilities or management fees divided by the number of shares. For most mutual funds, NAV is constantly calculated during market hours, from the moment of the first investment is made. Thus, investors can know exactly how their mutual fund is performing at any given time, and investors can buy or sell mutual funds at any time.

NAV mutual fund investors follow NAV closely, and some even use an investment strategy of investing in mutual funds, based on the trends in NAV over time. The study of NAV history is a common practice among sophisticated investors and is considered by many to be one of the best metrics of mutual fund performance.

Mutual funds are variable. Even the best mutual funds have occasionally declined. Many sector-based mutual funds are very cyclical. The NAV managers make big savings up and down across multiyear economic cycles. Many investors study these patterns and make investment decisions on the basis of AV histories and others just look to invest in funds or managers with good track records in crating NAVs.

88.1.2 Asset Allocation Fund

On a historical basis, the performance of stocks has typically outperformed most other investments. However, stock investments carry more risk than many other investment types. Many tout the success of asset allocation in the investment process. Asset allocation investments cut across such investment classes (stocks, bonds, and cash) and across international boundaries.

The asset allocation strategy seems to balance risk and reward by apportioning portfolio assets according to individual goals, risk tolerance, and investment horizon. The three main asset classes, equities, fixed income, and cash and equivalents, all have different levels of risk and return, so each will behave differently over time.

There are no simple methods that can find the correct allocation for every individual. However, most financial experts agree that asset allocator is a key and crucial decision for investing. The selection of an individual security is secondary to how each investor allocates their investment in stocks, bonds, and case and equivalent.

The asset allocation mutual funds (life cycle or target date funds) that attempt to provide investors with portfolio structures that address the investor’s age, risk profile land investment objective, via an appropriate apportionment of asset classes.

88.1.3 Vanguard Wellington Fund (VWELX)

Vanguard Wellington Funds are one of the few funds that survived the stock market crash of 1929. Vanguard Wellington was created just months before that collapse of the stock market.

As of June 2012, VWELX has assets totaling over $30 billion. Roughly two-thirds of the portfolio consists of stocks that typically pay dividends. The fund can invest up to 25 % of its portfolio in non-US securities.

Typically the fund maintains a balance between relatively conservative stocks and bonds. Thus, it is designed to weather bear markets, but it may not perform as well as more aggressive funds during market rallies. Such a balanced fund is a middle-of-the-road investment that seeks to provide some combination of income, capital appreciation, and the conservation of capital by investing in a max of stocks and bonds. The fund invests in a mix of undervalued and dividend-paying stocks and mostly investment-grade bonds.

88.2 Computations: Models and Results

The forecasting models used in this study were implemented through the use of Minitab 16 time series procedures. The data used in the forecasting process can be viewed in Fig. 88.1. In the period of January 2010 through December 2011, the total NAV of VWELX goes from a bi less than 30 billion to a slightly more than 30 billion. The nature of these modeling structures can be detailed in the appendix of the paper (Atsulaki and Valavanis 2009; Botlen and Busse 2004; Francis and Ghijsels 1999; Frances et al. 2005; Lendasse et al. 2000; Pas and Lin 2005; Taylor and Snyder 2012; Wang et al. 2012).

The first set of models used is the regression-type model:

1. Linear growth model
2. Quadratic growth model
3. Exponential growth model
4. An S-curve model

The second type of model used was a moving-average model. The third set of models used consists of exponential smoothing models:

1. Single exponential smoothing
2. Double exponential smoothing
3. Winter's multiplicative model
4. Winter's additive model

The fourth set of models consists of an autoregressive model:

1. Autoregressive nonseasonal model

In forecasting the first 3 months of 2012, the following can be seen:

Best forecasting model for period (25):

1. Quadratic growth model
2. Autoregressive nonseasonal model
3. Winter's additive model

Best forecasting model for period (26):

1. Linear growth model
2. Exponential growth model
3. S-curve model

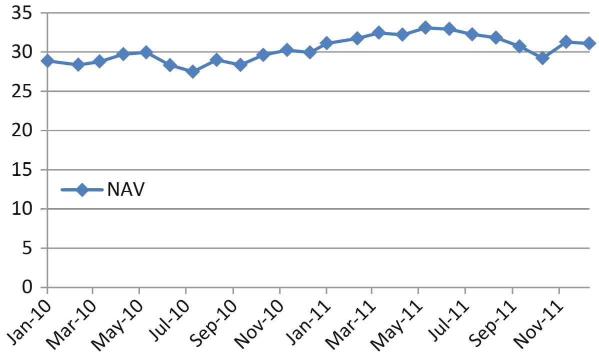


Fig. 88.1 NAV of VWELX

Best forecasting model for period (27):

1. Exponential growth model
2. Autoregressive nonseasonal model
3. S-curve model

It appears that the best model for forecasting in all three periods is the exponential growth model, followed closely by the autoregressive nonseasonal model. In evaluating the fit of the exponential growth model, the results, in terms of MAPE, MAD, and MSD, are well based.

While this model is quite simple in its performance on predicting the NAV in months 25, 26, and 27, the January–March 2012 is the best of all the models.

The forecasting period of the model structure is somewhat limited in nature. Months prior to January 2010 were a time of great economic turmoil (the worst since the Great Depression almost 70 years ago), and thus, these months were not used. In addition to further study of time series of NAV using the first 3 months of 2012 in the forecasting process of the second 3 months of 2012, we suggest using other forecasting studies to predict the total NAV of VWELX. We also suggest the development of macroeconomic econometric models to predict the total NAV of VWELX. Furthermore, the use of combination of forecasts is another topic for consideration in this forecasting process.

88.2.1 Combination of Forecasts

Besides just selecting one forecast when several sets of forecasts are available, the possibility exists to form a composite set of forecasts. This very basic idea has proven to be very effective over a wide range of management forecasting applications (see Bates and Granger 1969; Clemen 1984; Granger 1989).

If one has J forecasts over a given quantity S in a previous time period, a combined forecast is a weighted average of individual forecasts:

$$D = \sum_{I=1}^{J} w_i = 1 \text{ and } w_i = 0, j = 1, \ldots, k$$

The simplest weighting would be to consider all forecasting methods equal. If, however, a historical record of past forecasting results was available, then an appropriate combination of weights would be based upon their performance. The weights could vary over time. A possible methodology based on various historical forecasting evaluations could, also, be determined (see, e.g., Lawrence and Reeves 1981, 1982; Lawrence et al 2009). Such a process will be the topic of future research (Tables 88.1 and 88.2).

Appendix 1: A Forecasting Method Used in NAV Forecasting

Simple Exponential Smoothing

Simple exponential smoothing provides an exponentially weighted moving average of all previously observed values. The model is often appropriate for data with no predictable upward or downward trend. Exponential smoothing continually revises an estimate in the light of more recent experiences. This method is based on averaging past values of a series in an exponentially decreasing manner. The exponentially smoothing equation is

$$\hat{Y}_{t+1} = \alpha Y_t (1 - \alpha) \hat{Y}_t$$

where
$\hat{Y}_{t+1}$: new smoothed value or the forecasting value for the next period
α: smoothing constant ($0 < \alpha < 1$)
Y_t: actual value of the series in period t
$\hat{Y}_t$: old smoothed value of the forecast for period t

Simple exponential smoothing, as moving averages, uses only past values of a time series to forecast future values of the same series and is properly employed when there is no trend or seasonality present in the data. With exponential smoothing, the forecast value at any time is a weighted average of all the available previous values: the weights decline geometrically as you go back in time. Moving-average forecasting gives equal weights to the past values included in each average; exponential smoothing gives more weight in the recent observations and less to the older observations. The weights are made to decline geometrically with the age of the observation to confirm to the argument that the most recent observations contain the most relevant information so that they should be accorded proportionately more influence than older observations.

Exponential smoothing proceeds as do moving averages by smoothing past values of the series; the calculations for producing exponentially smoothed forecasts can be expressed as an equation. The weight of the most recent observation is

Table 88.1 A comparison of the accuracy forecasting models

#	Forecasting model	Actual			Estimate			Error			Total error
		NAV (25)	NAV (26)	NAV (27)	NAV (25)	NAV (26)	NAV (27)	NAV (25)	NAV (26)	NAV (27)	
1	Linear growth	31.34	32.37	33.21	32.31	32.46	32.68	−0.86	−0.09	−1.56	2.51
2	Quadratic growth	31.34	32.37	33.21	31.32	31.23	31.13	0.02	1.14	2.08	3.24
3	Exponential growth model	31.34	32.37	33.21	32.33	32.50	32.66	−0.99	−0.29	−0.33	1.61
4	S-curve	31.34	32.37	33.21	31.75	31.79	31.82	−0.41	0.58	1.39	2.38
5	Moving average	31.34	32.37	33.21	30.52	30.52	30.52	0.92	1.85	2.69	5.46
6	Single exponential smoothing	31.34	32.37	33.21	28.96	28.96	28.96	0.26	1.29	2.13	3.68
7	Double exponential smoothing	31.34	32.37	33.21	28.72	29.95	29.92	0.18	1.24	2.05	3.37
8	Winter’s multiplicative	31.34	32.37	33.21	31.13	31.78	31.28	0.21	0.69	1.93	2.83
9	Winter’s additive	31.34	32.37	33.21	31.15	31.78	31.27	0.19	0.59	1.94	2.72
10	Autoregressive nonseasonal	31.34	32.37	33.21	31.42	31.70	31.94	−0.08	0.67	1.27	2.02

Table 88.2 Forecasting models and fit measures

#	Type	Model	MAPE	MAD	MSD	MS
1	Linear growth	$Y_t = 28.393 + 0.156 * t$	3.18356	0.96921	1.37241	
2	Quadratic growth	$Y_t = 27.400 + .386\ t + 0.00917\ t^{**2}$	3.08086	0.92772	1.21882	
3	Exponential growth model	$Y_t = 28.4085\ t * (1.00519)^{**t}$	3.19579	0.97401	1.38708	
4	S-curve	$Y_t = 10^{*\ *\ 3}/(31.1135 + 5.73251) * (0.897029 * *^t)$	3.26895	0.98605	1.31431	
5	Moving average	None	2.01333	0.60863	0.49273	
6	Single exponential smoothing	$\alpha = 0.2$	3.37824	1.03623	1.63207	
7	Double exponential smoothing	$\alpha = 0.2$	3.59835	1.09078	1.83576	
8	Winter's multiplicative	$\alpha = 0.02, \gamma = 0.2, \Delta = 0.2$	3.51305	1.05655	1.76160	
9	Winter's additive	$\alpha = 0.02, \gamma = 0.2, \Delta = 0.2$	3.44271	1.03609	1.75187	
10	Autoregressive nonseasonal	$Y_t = .040625 + .7057AR(1) + 1.1172MA(1)$				0.7735

assigned by multiplying the observed value by α, the next most recent observation by $(1-\alpha)\alpha$, the next observation by $(1-\alpha)^2\alpha$, and so on. The number chosen for α is called the smoothing constant.

The value of the smoothing constant α must be between 0 and 1. The value of the smoothing constant cannot be equal to 0 or 1; if it is, the entire idea of exponential smoothing is negated. If a value close to 1 is chosen, recent values of the time series are weighted heavily relative to those of the distant past when the smoothed values are calculated. Likewise, if the value of α is chosen close to 0, then the values of the time series in the distant past are given weights comparable to those given the recent values. The rate at which the weights decrease can be seen from their values for an α of 0.1.

Regardless of the smoothing constant chosen, the weights will eventually sum to 1. Whether the sum of the weights converges on 1 quickly or slowly depends on the smoothing constant chosen. If a smoothing constant of 0.9 is chosen, then the sum of the weights will approach 1 much more rapidly than when the smoothing constant is 0.1.

In choosing α, select values close to 0 if the series has a great deal of random variation; select values close to 1 if the forecast values depend strongly on recent changes in the actual values. The root-mean-squared error is often used as the criterion for assigning an appropriate smoothing constant; the smoothing constant with the smallest root-mean-squared error would be selected as the model likely to produce the smallest error in generating additional forecasts. In practice, relatively small values of α generally work best when simple exponential smoothing is the most appropriate model.

Double Exponential Smoothing

A variation of simple exponential smoothing that includes trends is referred to as double exponential smoothing. This technique is appropriate only when the data vary around an average or have a step or gradual change. This double exponential model is composed of two elements:

$$\hat{Y}_{t+1} = S_t + T_t$$

where
S_t: smoothed error
T_t: current trend estimate

$S_t = Y_t + \alpha_1 \left(Y_t - \hat{Y}_e\right)$

$T_t = T_{t-1} {}_{+}\alpha_2 \left(\hat{Y}_t - \hat{Y}_{t-1-} - T_{t-1}\right)$
α_1: smoothing constant
α_2: smoothing constant

A method similar to double moving averages is double exponential smoothing. This method accounts for trend and retains the advantage of requiring less data than moving averages, an attribute of all exponential smoothing methods.

Double exponential smoothing uses a single coefficient, alpha, for both smoothing operations. As in double moving averages, this method computes the difference between single and double smoothed values as a measure of trend. It then adds this value t the single smoothed value together with adjustment for the current trend.

As is true for simple exponential smoothing, double smoothing requires starting values to initialize the formulae. The advantages of double smoothing are as follows:

1. It models the trends and level of a time series.
2. It is computationally more efficient than double moving averages.
3. It requires less data than double moving averages. Because one parameter is used, parameter optimization is simple.

Although parameter optimization is simple, there is some loss of flexibility because the best smoothing constants for the level and trend may not be equal. The exponential smoothing model is not a full model; it does not model the seasonality of a series. Often series have seasonality. Thus, it is not recommended unless the data is first deseasonalized.

Winter's Method

Winter's method is similar to those of other linear exponential smoothing methods but has the advantage of being capable of dealing with seasonal data in addition to data that have a trend. The three basic smoothing equations of Winter's method are as follows:

$$S_t = \alpha + [X_t/I_{t-L}] + (1-\alpha)(S_{t-1} + T_{t-1})$$

$$T_t = \beta(S_{t-}S_{t-1}) + (1-\beta)T_{t-1}$$

$$I_t = \tau[X_t/S_L]\tau + (1-\tau)I_{t-L}$$

where
S_t: smoothed value of deseasonalized service
T_t: smoothed value of trend
I_t: smoothed value of seasonal factor
L: length of seasonality (four quarters or a year)

Trend Models

The expected value of a series may change as time passes since the influence of factors changes. If the movement is long term and only in one direction, up or down, it is called a trend.

With trended series, one assumes that there is some functional relationship between the expected value of the series and the time variable. This function usually involves some unknown parameters β_0, β_1, ... that must be estimated from the series history.

A time series whose average value changes over time is called nonstationary. Trend is a particular kind of nonstationary in which the pattern of change has two main properties:

1. It is of long duration compared to the forecast horizon.
2. It is predominantly in one direction only, either upward or downward.

A time series y_1, y_2, ... is said to have trend if the expected value of y_i changes over time so that

$$E(y_t) = f(\beta_0, \beta_1, \ldots, t)\, t = 1, 2, \ldots$$

where f (β_0, β_1,. . ., t) is predominantly either an increasing or a decreasing function of t over the forecasting horizon (the time for which forecasts are to be constructed).

A model for trended time series is as follows:
y_t = the actual value of the series at time t
T_t = the expected value (trend) of the series at time t
= f(β_0, β_1, . . .;t)
ε_t = a random variable representing irregular fluctuations around the trend
where f(β_0, β_1,. . .; t) is an increasing or decreasing function describing the trend pattern.

The general trend model is $y_t = T_i + \varepsilon\ t = 1, 2, \ldots$. This model describes a consistently growing or declining series with variation from trend (T_t) caused by random influences (ε_t). In most cases, it is assumed that the ε_t's form a stationary random series with $E(\varepsilon_t) = 0$ for all t.

There may be some confusion about whether a trend is only a part of a very long cycle. However, practically speaking, a portion of a long cycle may be modeled as trend if the cycle has not changed direction during the time period of interest to the forecaster and if no such change is anticipated. Under these circumstances, it is probably only academic whether the movement is truly a trend of a cycle. If a cycle is modeled as trend and it happens to turn over during the forecast horizon, then large forecast errors may result.

Among the major sources of trend are the following:

1. Population changes
2. Technological changes
3. Changes in social custom
4. Inflation/deflation
5. Changing environmental conditions
6. Changing market acceptance

Linear Trend Model

These models indicate the increases or decreases like a straight line and are represented by the equation

$$\hat{Y}_t = b_0 + b_1 t$$

where

Y_t is the predicted value of the trend at time t. The values of b_0 and b_1 are found by the method of least squares.

Quadratic Trend Model

Another trend model often used is a quadratic trend model:

$$\hat{Y}_t = b_0 + b_1 t_1 + \ldots + b_2 t^2$$

where the values of b_0, b_1, and b_2 are found by the method of least squares.

Exponential Trend Model

Furthermore, when the time series starts slowly and then appears to be increasing at an increasing rate such that the percentage difference from observed is constant, an exponential trend can be fitted. The exponential trend is given by

$$\hat{Y}_t = b_0 b_1^t$$

The coefficient is related to the growth rate. If the exponential trend b_1 is fit to annual data, the annual growth rate is estimated to be 100 $(b_1 - 1)$%.

Moving-Average Method

In the method of simple averages, the mean of the data set is to forecast the future. However, if the forecaster is more concerned with more recent observations, a constant number of data points can be specified to compute the mean for the most recent set of observations. This process is referred to as a moving average. As each new data set observation becomes available, a new mean is computed by adding the newest value and dropping the oldest. This moving average is used to forecast the next period.

A moving average of adding K MA(K) is given by
$\hat{Y}_{t+1}$: the actual value at period t
Y_{e+1}: forecasted value for the next period
K: the number of terms in the moving average

Each observation in the process is given an equal weight. The rate of response to changes in the series depends on the number of period in K. The moving-average model does not handle trend or seasonality factor.

To compute the weighted average of a set of data, each first observation is multiplied by a weight representing its relative importance; then the produces are summed, and finally this sum is divided by the sum of the weights. The resulting weights then sum to 1.0 ad are called normalized weights, which are often expressed as percentages. They give quite a clear picture of the proportionate influence of each observation on the value of the average. In forecasting formulae,
w_0 = the weight given to the most recent observation
w_1 = the weight given to the one-period-old observation
w_2 = the weight given to the two-period-old observation
...
w_i = the weight given to the i-period-old observation

In smoothing schemes with normalized weights, then the one-step-ahead forecast is

$$\hat{Y}_{t+1} = \beta_0(t) = w_0 y_{t+\ wt} y_{t-1} + \ldots + w_{k-1} y_{t-k+1}\, \mathbf{k} \le \mathbf{t} = \sum_{0}^{k-1} w_i y_{t-1}$$

and the p-step-ahead forecast is

$$\hat{Y}_{t+p} = \hat{Y}_{t+1} \quad \mathbf{p} \ge \mathbf{t}$$

Because the weights move toward one step each time a new observation is obtained, the average computed is called a moving average.

With the updating procedure, the weighting scheme changes at each step. That is, the weights at time t are 1/t,1/t, 1/t,. . ., 1/t and at time t + 1 become 1/(t + 1), 1/(t + 1), . . ., 1/(t + 1). This method is referred to as updating the mean. Since all previous observations enter equally into the average, updating the mean provides a maximal amount of smoothing and, thus, a minimal amount of tracking.

Another concept used in interpreting moving averages is the center of the average. The center is that point in time which the average most represents. The center is not necessarily the idle of the interval; it depends on the weighting scheme. However, the average used in updating the mean is centered at (t + 1)/2, the middle of the interval from 1 to t, because of the symmetry of the weights.

Varying amounts of smoothing and tracking can be obtained by altering the weighting scheme. A common way to reduce the extreme smoothing of updated mean forecasts is to compute the mean of only the k most recent observations. The result is called a moving average of length k. Varying the value of k varies the amount of smoothing, with smaller values of producing less smoothing (more tracking). For any particular series, the forecaster may select k according to the amount of smoothing desired, or k may be chosen to optimized one of the standard criteria (RMSE, MAD, etc.) mentioned earlier in this section. In actually, this method discards all observations that are more than k units old, and the weight to each of the k most recent observations is equal and constant over time. There must be k observations to begin the forecasting, and to compute new forecasts we must retain all k of the most recent observations from step to step.

Box-Jenkins (ARIMA) Model

ARIMA models are a class of linear models that are capable of representing stationary as well as nonstationary time series. ARIMA models make use of information in the series itself to generate forecasts.

Box-Jenkins method is a forecasting methodology since it does not assume any particular pattern in the historical data of the series. To begin with the selection of an ARIMS model is based on the examination of the time series, and it is autocorrelated for several time lags. In specifics, the pattern of known autocorrelation is associated with a particular ARIMA model structure.

The Box-Jenkins method of forecasting is different from most methods. The technique does not assume any particular pattern in their historical data of the series to be forecast. It uses an iterative approach of identifying a possible useful model from a general class of models. The chosen model is then checked against the historical data to see whether it accurately describes the series. The model fits well if the residuals between the forecasting model and the historical data points are small, randomly distributed, and independent. If the specified model is not satisfactory, then the process is repeated by using another model designed to improve on the original one.

A general class of Box-Jenkins models for a stationary time series is the ARIMA, or autoregressive integrated moving average, models. Stationary time series is one shoes average value is not changing over time. This group of models includes the autoregressive models with only autoregressive terms with both autoregressive and moving-average terms. The Box-Jenkins methodology allows the analyst to select the model that best fits the data.

Selection of an appropriate model can be made by comparing the distributions of autocorrelation coefficients of the time series being fitted with the theoretical distributions for the various models.

There are three basic stages in the modeling process: identification, fitting, and diagnostic checking.

The objective of identification is to select the forecast model that seems most appropriate to the time series under study. The data are used to generate a series of sample autocorrelation functions. These are now compared to certain theoretical autocorrelation functions from known forecast models to seek the best match between the sample and theoretical results. With this, the forecast model is identified and selected. The principle of parsimony is applied in that the model with the smallest number of coefficients that is suitable for the item is the selected model.

Upon selecting the model, fitting is carried on where the coefficients are estimated. The estimates are found so that they yield the fit of the past demands which procedures the minimum sum of squared residual errors.

Using the fitted results, the residual errors are examined to determine the adequacy of the fit. A good fit will yield residual errors that are randomly distributed with mean zero and a common variance. The check is made via the autocorrelation function of the residual errors.

If the diagnostic check fails, then the three stages are repeated until a model is found which gives appropriate results.

Autoregressive Models

The first type of model that will be examined is an autoregressive model. A pth order autoregressive model takes the following form:

$$Y_t : a_0 + a_1 Y_{t-1} + a_2 Y_{t-2} + \ldots + a_p Y_{t-p} + e_t$$

where

$Y_{e:}$ response variable at time t

Y_{t-1}, Y_{t-2}, . . . + Y_{t-p}: response variable at various time lags

a_0, $a_{1,}$ a_2, $a_{3,\ldots}$, a_p: coefficient to be estimated

e_t: error term at time t

The given model has the form of a regression model with lagged values of the dependent variable in the independent variable. This model is referred to as an autoregressive model.

Moving-Average Model

A moving-average model is given by the gth-order moving-average model, which takes the following form:

$$Y_t = \mu + e_t - e_1 t_{t-1} - b_2 t_{t-1} - \ldots - b_g t_{t-g}$$

where

Y_t: response variable at time t

μ: constant mean of the process

$b_{1,}$ b_2, $b_{3,\ldots}$, b_g: coefficients to be estimated

e_t: error term; variables not explained by the model

e_{t-1}, t_{t-2} . . . $-e_{t-g}$: errors in previous time periods

A moving-average model provides a forecast of Yt based on a linear combination of a finite number of past errors.

Unbounded Autoregressive and Moving-Average Models

A model that has both autoregressive terms can be combined with a model having moving-average terms. It is referred to as a mixed autoregressive moving-average form.

The form of the model is

$$Y_t : a_0 + a_1 Y_{t-1} + a_2 Y_{t-2} + \ldots + a_p Y_{t-p} + \varepsilon_t + \left[-b_1 \varepsilon_{t-1} - b_2 \varepsilon_{t-2} \ldots - b_g t_{t-g}\right]$$

References

Armstrong, J. S. (1989). Combining forecasts: The end of the beginning or the beginning of the end? *International Journal of Forecasting, 5*(4), 585–592.

Atsulakis, G. S., & Valavanis, K. P. (2009). Surveying stock market forecasting techniques, Part II, soft computing methods. *Export Systems Applications, 3*, 5932–41.

Bates, J. M., & Granger, C. W. J. (1969). The combination of forecasts. *Operations Research Quarterly, 20*, 451–468.

Botlen, N., & Busse, J. (2004). Short-term persistence in mutual fund performance. *Review of Financial Studies, 8*(2), 569–597.

Box, G. E., Jenkins, G. W., & Reinsel, G. (1994). *Time series analysis, forecasting and control* (3rd ed.). Englewood Cliffs: Prentice Hall.

Clemen, R. T. (1984). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasts, 5*, 559–583.

Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting, 5*, 559–581.

Diebold, F. X. (1989). Forecast combination and encompassing: Reconciling two divergent literatures. *International Journal of Forecasting, 5*, 589–592.

Diebold, F. X., & Lopez, J. (1996). Forecast evaluation and combination. In G. S. Maddala & C. S. Rao (Eds.), *Handbook of statistics* (pp. 241–268). Amsterdam: North-Holland.

Diebold, F. X., & Pauly, P. (1990). The use of price information in forecast combination. *International Journal of Forecasting, 6*, 503–508.

Diebold, F. X., & Rudebusch, G. D. (1991). Forecasting output with the composite leading index: An ex ante analysis. *Journal of the American Statistical Association, 86*, 603–610.

Elliott, G., & Timmermann, A. (2002). *Optimal forecast combination under regime switching*. Manuscript, Department of Economics, University of California, San Diego.

Frances, P. H., Marches, M., & Murray, A. (2005). A hybrid genetic-neural-architecture for stock index forecasting. *Information Science, 170*(1), 3–37.

Francis, P., & Ghijsels, H. (1999). Additive outliers GARCH and forecasting volatility. *International Journal of Forecasting, 15*(1), 1–9.

Granger, C. W. J. (1989). Combining forecasts twenty years later. *Journal of Forecasting, 8*, 167–173.

Granger, C. W. J., & Newhold, P. (1986). *Forecasting economic time series* (2nd ed.). San Diego: Academic.

Lawrence, K. D., & Reeves, G. (1981). Consensus time series forecasting. In J. Morse (Ed.), *Lecture notes in economics and mathematical systems* (Vol. 190, pp. 199–204). Berlin: Springer.

Lawrence, K. D., & Reeves, G. (1982). Combining multiple forecasts given multiple objectives. *Journal of Forecasting, 1*, 271–279.

Lawrence, K. D., Klimberg, R. K., & Lawrence, S. M. (2009). *Fundamentals of forecasting using excel*. New York: Industrial Press.

Lendasse, A., DeBodt, E., Wentz, & Verlegsen, M. (2000). Nonlinear financial time series forecasting-application to the Belgian 20 stock market index. *European Journal of Economical and Social Systems, 14*(1), 81–91.
Mahoud, E. (1989). Combining forecasts: Some managerial issues. *International Journal of Forecasting, 5*(4), 599–600.
Pas, P. F., & Lin, C. S. (2005). A hybrid ARIMA and support vector machines model on stock price forecasting. *OMEGA, 33*, 497–505.
Stock, J. H., & Watson, M. W. (1998). A dynamic factor model framework for forecast combination. Manuscript, Kennedy School, Harvard University and Woodrow Wilson School, Princeton University.
Taylor, J. W., & Snyder, R. D. (2012). Forecasting intraday time series with multiple seasonal cycles using parsimonious seasonal exponential smoothing. *OMEGA, 40*, 748–757.
Wang, J. J., Wang, J., Zhang, Z., & Guo, S. (2012). Stock index forecasting based on a hybrid mode. *OMEGA, 40*, 758–766.
Winker, R. L. (1989). Combining forecasts: A philosophical basis and some current issues. *International Journal of Forecasts, 5*, 605–608.
Winkler, R. L., & Makridakis, S. (1983). The combination of forecasts. *Journal of the Royal Statistical Society, A, 146*, 150–157.

Discriminant Analysis and Factor Analysis: Theory and Method

89

Lie-Jane Kao, Cheng-Few Lee, and Tzu Tai

Contents

Abstract

Three multivariate techniques, namely, discriminant analysis, factor analysis, and principal component analysis, are discussed in detail. In addition, the stepwise discriminant analysis by Pinches and Mingo (1973) is improved using a goal

L.-J. Kao (✉)
Department of Finance and Banking, Kainan University, Taoyuan, ROC, Taiwan
e-mail: ljkao@mail.knu.edu.tw

C.-F. Lee
Department of Finance and Economics, Rutgers Business School, Rutgers, The State University of New Jersey, Piscataway, NJ, USA

Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
e-mail: lee@business.rutgers.edu; cflee@business.rutgers.edu

T. Tai
Department of Finance and Economics, Rutgers, The State University of New Jersey, Piscataway, NJ, USA
e-mail: tzutai@pegasus.rutgers.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_89,

programming technique. These methodologies are applied to determine useful financial ratios and the subsequent bond ratings. The analysis shows that the stepwise discriminant analysis fails to be an efficient solution as the hybrid approach using the goal programming technique outperforms it, which is a compromised solution for the maximization of the two objectives, namely, the maximization of the explanatory power and the maximization of discriminant power.

Keywords

Multivariate technique • Discriminant analysis • Factor analysis • Principal component analysis • Stepwise discriminant analysis • Goal programming • Bond ratings • Compromised solution • Explanatory power • Discriminant power

89.1 Introduction

Financial ratios are widely used in all financial analysis and planning. Banks use a firm's current and quick ratio to determine acceptability, for commercial loans; the leverage ratio is used as a proxy for a firm's capital measure in predicting bankruptcy and to analyze the impact of leverage on the market value of a firm. In financial analysis and planning determination, lenders or managers need to measure a customer's (either an individual's or a firm's) short-term or long-term financial position. The multivariate statistical techniques of factor analysis (or principal component analysis) and discriminant analysis had been used in such instances to identify important financial ratios and to construct one or several well-known financial indicator(s), namely, the financial z-score (Altman 1968; Altman and Eisenbeis 1968; Altman et al. 1977). Although the financial *z*-scores are a compromise between theory and practice, z-scores have been used extensively by practitioners and academicians in credit analysis and bankruptcy prediction (Altman 1983). In literature, factor analysis (or principal component analysis) is often used to extract a set of variables with significant explanatory power in a high-dimensional dataset. Together with discriminant analysis, objects are classified into several homogeneous groups by the financial indicator(s) extracted from the factor analysis (or principal component analysis). The credit-scoring model by Pinches and Mingo (1973) is an example of the aforementioned stepwise principal component-discriminant analysis approach. However, when used together, due to their conflicting objectives, the performance of the stepwise principal component-discriminant analysis might not be as satisfactory as it would be when used separately.

Instead of stepwise principal component-discriminant analysis, the methodology of factor -discriminant analysis in conjunction with the minimax goal programming is explored in this chapter. In Sect. 89.2, the theory and methodology of discriminant analysis will be explored. Section 89.3 will discuss the theory and methodology of factor (or principal component) analysis.

In Sect. 89.4, the performance of a hybrid multivariate credit-scoring model that involves a compromise maximization of the two conflicting objectives associated with the principal component and discriminant analysis is presented, and its performance using Pinches and Mingo (1973) data is illustrated. Finally, the results of this chapter are summarized in Sect. 89.5.

89.2 Discriminant Analysis

The purposes of discriminant analysis are twofold: (1) to test for mean group differences among the groups and (2) to construct a classification scheme based upon a set of m variables in order to assign previously unclassified observations to appropriate groups. For example, in a study of corporate bankruptcy, Altman (1968) used data from a sample of failed firms and a sample of existing firms and developed a classification rule that uses financial ratios to determine whether bankrupt firms had significantly different financial ratios prior to failure than did solvent firms.

89.2.1 Two-Group Discriminant Analysis

A linear two-group discriminant function can be defined as

$$Y_i = a_1 x_{1i} + \dots . + a_m x_{mi} \tag{89.1}$$

where Y_i is a binary variable used to indicate two alternative options, e.g., good or bad accounts in credit analysis, bankrupt or non-bankrupt firms in corporate bankruptcy analysis, and problem or nonproblem banks in banking analysis. x_{1i},.., x_{mi} are m explanatory variables. Two different methods can be used to estimate the coefficients of Eq. 89.1; they are the dummy regression method and the eigenvalue method. The dummy regression method is given in Sect. 89.2.1.1, while the eigenvalue method is given in Sect. 89.2.1.2.

It is important for readers to understand the relationship between the logic of two-group discriminant analysis and the regression technique to estimate related discriminant function parameters. Following Tatsuoka (1988), Johnston and Dinardo (1996), and Eisenbeis and Avery (1972), the basic equation of discriminant analysis as derived in Appendix 1 can be defined as

$$\begin{aligned} &(B - EC)A = 0, \\ &D' = \left[\overline{X}_{1,1} - \overline{X}_{1,2} \dots , \overline{X}_{m,1} - \overline{X}_{m,2}\right], \end{aligned} \tag{89.2}$$

where $B = DD'$ is the between-group variance; C is the within-group variance; A is the vector representing the coefficients such that

$|A - \lambda I| = 0$, where I is the identity matrix and λ is the eigenvalue; and E is the ratio of the weighted between-group variance to the pooled within-group variance. Multiplying by C^{-1} into both sides of Eq. 89.2, the characteristic equation associated with Eq. 89.2 is

$$\left(C^{-1}B - EI\right)A = 0. \tag{89.3}$$

In order to use the linear discriminant function for empirical analysis, one must estimate the coefficients of Eq. 89.2. To illustrate the computation of two-group

Table 89.1 Roster of liquidity and leverage ratios

Group 1 ($N_1 = 6$)			Group 2 ($N_2 = 8$)		
x_{1i}	x_{2i}	Y_i	x_{1i}	x_{2i}	Y_i
2.0	0.50	1	1.8	0.35	0
1.8	0.48	1	1.9	0.34	0
2.3	0.49	1	1.7	0.42	0
3.1	0.41	1	1.5	0.49	0
1.9	0.43	1	2.2	0.36	0
2.5	0.44	1	2.8	0.38	0
			1.6	0.55	0
			1.4	0.56	0

Two groups with response variable Y and two predictors X_1 and X_2

discriminant functions as a multiple regression equation and the eigenvalue method, we shall use a numerical example illustrated in Table 89.1, which shows the number scores of two groups on two predictor variables, i.e., the liquidity ratio X_1 and the leverage ratio X_2, and the response variable Y. All members of group 1 are assigned $Y = 1$, and all members of group 2 are given $Y = 0$. There are two alternative methods, the dummy regression and the eigenvalue method, to estimate the discriminant function. The two methods are stated as below.

89.2.1.1 Dummy Regression Method

If there are two explanatory variables, i.e., $m = 2$, Eq. 89.1 can be written as

$$y_i = a_1 x_{1i} + a_2 x_{2i} \tag{89.4}$$

where $y_i = Y_i - \overline{Y}$, $x_{1i} = a_1 X_{1i} + a_2 \overline{X}_1$, and $x_{2i} = X_{2i} - \overline{X}_2$. The equation system used to solve a_1 and a_2 can be defined as

$$Var(x_{1i})a_1 + Cov(x_{1i}, x_{2i})a_2 = Cov\left(x_{1i}, y_i\right) Cov(x_{1i}, x_{2i})a_1 + Var(x_{2i})a_2 = Cov\left(x_{2i}, y_i\right).$$

Following the data listed in Table 89.1, $Var(x_{1i})$, $Var(x_{2i})$, $Cov(x_{1i}, x_{2i})$, $Cov(x_{1i}, y_i)$, and $Cov(x_{2i}, y_i)$ are calculated as follows:

$$\begin{aligned} Var(x_{1i}) &= \frac{\sum X_{1i}^2}{n} - \left(\frac{\sum X_{1i}}{n}\right)^2 \\ &= \frac{32 + 29.19}{n} - \left(\frac{13.6 + 14.9}{n}\right)^2 \\ &= \frac{61.19}{14} - \left(\frac{28.5}{14}\right)^2 = 0.2267; \end{aligned}$$

$$
\begin{aligned}
Var(x_{2i}) &= \frac{\sum X_{2i}^2}{n} - \left(\frac{\sum X_{2i}}{n}\right)^2 \\
&= \frac{1.2671 + 1.5447}{14} - \left(\frac{2.75 + 3.45}{14}\right)^2 = 0.0048;
\end{aligned}
$$

$$
\begin{aligned}
Cov(x_{1i}, x_{2i}) &= \frac{12.424}{14} - (2.0357)(0.4428) \\
&= 0.8874 - 0.9014 \\
&= -0.014;
\end{aligned}
$$

$$
\begin{aligned}
Cov(x_{1i}, y_i) &= \frac{\sum X_{1i}Y_i}{n} - \left(\overline{X}_1\right)\left(\overline{Y}\right) \\
&= \frac{13.6}{14} - (2.0357)(0.4285) \\
&= 0.9714 - 0.8722 \\
&= 0.0992;
\end{aligned}
$$

$$
\begin{aligned}
Cov(x_{2i}, y_i) &= \frac{\sum X_{2i}Y_i}{n} - \left(\overline{X}_2\right)\left(\overline{Y}\right) \\
& \frac{2.75}{14} - (0.4428)(0.4288) \\
&= 0.0.1964 - 0.1897 \\
&= 0.00668.
\end{aligned}
$$

Using Cramer's rule, a_1 and a_2 can be calculated as:

$$
\begin{aligned}
a_1 &= \frac{\begin{vmatrix} 0.0992 & -0.0140 \\ 0.0066 & 0.0048 \end{vmatrix}}{\begin{vmatrix} 0.2267 & -0.0140 \\ -0.0140 & 0.0048 \end{vmatrix}} \\
&= \frac{0.00047616 + 0.0000924}{0.0010886 - 0.000196} = 0.63697
\end{aligned}
$$

$$
\begin{aligned}
a_2 &= \frac{\begin{vmatrix} 0.2267 & 0.00992 \\ -0.0140 & 0.0066 \end{vmatrix}}{\begin{vmatrix} 0.2267 & -0.0140 \\ -0.0140 & 0.0048 \end{vmatrix}} \\
&= \frac{0.00149 + 0.00138}{0.00108 - 0.00019} = 3.2359
\end{aligned}
$$

Normalizing the regression coefficient by dividing a_2 into a_1, we obtain

$$\begin{bmatrix} a_1/a_2 \\ 1 \end{bmatrix} = \begin{bmatrix} \frac{0.63697}{3.2359} \\ 1 \end{bmatrix} = \begin{bmatrix} 0.1968 \\ 1 \end{bmatrix}. \tag{89.5}$$

89.2.1.2 Eigenvalue Method

For the eigenvalue method, the elements of the matrices C and B in Eq. 89.2 can be calculated as:

$$C_{11} = 32 - \frac{(13.6)^2}{6} + 29.19 - \frac{(14.9)^2}{8} = 2.612;$$

$$\begin{aligned} C_{22} &= 1.2671 - \frac{(2.75)^2}{6} + 1.5447 - \frac{(3.45)^2}{8} \\ &= 0.0636 \end{aligned}$$

$$\begin{aligned} C_{12} &= 6.179 - \frac{13.6 \times 2.75}{6} + 6.245 - \frac{14.9 * 3.45}{8} \\ &= -0.2350; \end{aligned}$$

$$\begin{aligned} B_{11} &= 6\left[\frac{13.6}{6} - 2.0357\right]^2 + 8\left[\frac{14.9}{8} - 2.0357\right]^2 \\ &= 0.5601; \end{aligned}$$

$$\begin{aligned} B_{22} &= 6\left[\frac{2.75}{6} - 0.4428\right]^2 + 8\left[\frac{3.45}{8} - 0.4428\right]^2 \\ &= 0.0025 \end{aligned}$$

$$\begin{aligned} B_{12} &= 6\left(\frac{13.6}{6} - 2.0357\right)\left(\frac{2.75}{6} - 0.4428\right) \\ &\quad + 8\left(\frac{14.9}{8} - 2.0357\right)\left(\frac{3.45}{8} - 0.4428\right) \\ &= 0.03753. \end{aligned}$$

The matrices C and B can now be formulated as

$$C = \begin{bmatrix} 2.612 & -0.2350 \\ -0.2350 & 0.0636 \end{bmatrix} \quad B = \begin{bmatrix} 0.5601 & 0.03753 \\ 0.03753 & 0.0025 \end{bmatrix}.$$

Therefore, we can obtain

$$C^{-1}B = \begin{bmatrix} 0.4007 & 0.0268 \\ 0.0268 & 0.1384 \end{bmatrix}.$$

Now, the characteristic equation $|C^{-1}B - EI| = 0$ can be solved as $E^2 - 0.5391E = 0$, whose single nonzero root is $E_1 = 0.5391$. Now, the adjoint of $C^{-1}B - E_1 I$ is

$$\begin{bmatrix} -0.4007 & -0.0268 \\ -0.2071 & -0.1384 \end{bmatrix}$$

Hence, the eigenvector of $C^{-1}B$, with larger elements set equal to unity, is

$$A_1 = \begin{bmatrix} -0.4007/-2.071 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.1935 \\ 1 \end{bmatrix}.$$

When the vector of regression weights a_1 and a_2 obtained earlier is similarly rescaled, we arrive at the same solution as in Eq. 89.5, which was obtained by the dummy regression method in Sect. 89.2.1.1. Alternatively, the nonzero root, $E_1 = 0.5391$, can be substituted into Eq. 89.2, yielding:

$$\begin{bmatrix} -0.4007 - 0.5391 & -0.0268 \\ -2.071 & -0.1384 - 0.5391 \end{bmatrix} \begin{bmatrix} a_1 \\ a_2 \end{bmatrix} = 0.$$

Equation 89.3 implies that

$$-0.138a_1 + 0.027a_2 = 0$$

$$2.071a_1 - 0.4007a_2 = 0.$$

By normalizing $a_1^2 + a_2^2 = 1$, one has the same result in Eq. 89.5.

89.2.2 K-Group Discriminant Analysis

The two-group discriminant analysis theory can be readily generalized to the K-group case. Assume there are K samples of size N for each of the K populations with means $\boldsymbol{\mu}_1, \ldots, \boldsymbol{\mu}_K$ and weights $w_1, \ldots, w_K$, respectively, and a common variance-covariance matrix Σ. A set of r linear combinations of the p explanatory variables $X = (X_1, \ldots, X_p)'$, or discriminant functions $D_1 = \theta_1' X, \ldots, D_r = \theta_r' X$, $r \leq \min(K - 1, p)$, where the coefficient vectors $\theta_1 = \Sigma^{-1/2}\xi_1, \ldots,$

$\theta_r = \Sigma^{-1/2}\xi_r$, where ξ_j is subject to the constraints $\xi_i'\xi_i = 1$ and $\xi_i'\xi_j = 0$ for all $j < i$, are determined sequentially as follows.

Given the first i-1 coefficient vectors $\theta_1, \ldots, \theta_{i-1}$, $i \leq r$, the ith coefficient vector $\boldsymbol{\theta}_i$ is obtained by choosing ξ_i that maximizes the ratio comparing the variability between the groups to that within the groups, which can be formulated in the object function

$$g_1(\xi_i) = \xi_i' A \xi_i \tag{89.6}$$

where the $p \times p$ positive-definite matrix

$$A = \Sigma^{-\frac{1}{2}}\left(\sum_{j=1}^{K} w_j\left(\boldsymbol{\mu}_j - \overline{\boldsymbol{\mu}}\right)\left(\boldsymbol{\mu}_j - \overline{\boldsymbol{\mu}}\right)'\right)\Sigma^{-\frac{1}{2}} \tag{89.7}$$

Here, $\overline{\boldsymbol{\mu}} = \sum_{i=1}^{k} w_i \boldsymbol{\mu}_i$ is the overall mean and Σ is the common $p \times p$ variance-covariance matrix of the K populations. By maximizing the objective function g, the K population means of the r discriminant functions $D_1 = \theta_1' X, \ldots, D_r = \theta_r' X$, $r \leq \min(K-1, p)$ are separated as much as possible so that the chance of misclassification of objects is reduced and thereby the discriminant power is maximized.

89.3 Factor Analysis and Principal Component Analysis

89.3.1 Factor Analysis

Factor analysis has been applied in marketing, finance, and accounting. For example, Johnson and Dinardo (1996) used factor analysis to extract eight factors from 61 financial ratios. Anderson (2003), Tatsuoka (1988), and Churchill and Iacobucci (2004) argued that factor analysis is a popular "analyses of interdependence" technique, which is concerned with the overall relationships among the set of p explanatory variables that characterize the objects. In more specific, the model can be formulated as

$$Y = \boldsymbol{\mu} + \beta \boldsymbol{f} + U \tag{89.8}$$

Where $\boldsymbol{f}$ is the $m \times 1$ vector of unobservable factors, $\boldsymbol{\mu}$ is a fixed vector of means, and U is a vector of unobservable specific errors. The $p \times m$ matrix β consists of factor loadings ($m < p$), which are parameters to be estimated. To estimate model (89.8), it is always assumed

$$\begin{aligned}
&E(\boldsymbol{f}) = 0,\\
&E(U) = 0, E(\boldsymbol{f}'\boldsymbol{f}) = M,\\
&E\left(UU'\right) = \Sigma,\\
&E(\boldsymbol{f}U') = 0
\end{aligned}$$

where Σ is a diagonal matrix. From (89.8), the covariance matrix of the observations Y is

$$E(Y-\boldsymbol{\mu})(Y-\boldsymbol{\mu})^{'}=\beta M\beta^{'}+\Sigma. \tag{89.9}$$

Two alternative approaches exist in estimating (89.8); they are the principal component method and the maximum-likelihood method. The principal component method for extracting factors or calculating the coefficient matrix to meet the foregoing statistical assumption can be found in both Johnson and Dinardo (1996) and Tatsuoka (1988). The maximum-likelihood method can be found in Lawley (1940), Lawley and Maxwell (1963), and Joreskog (1967). In addition to model (89.8), factor scores of the latent factors $f_1, \backslash \ldots, f_{\mathrm{m}}$ are also of the primary interest and are estimated.

89.3.2 Principal Component Analysis

In many empirical studies, the number of variables under consideration is too large to handle. A way of reducing the number of variables to be treated is to discard the linear combinations that have small variances and to study only those with large variances. For example, the first principal component is the normalized linear combination (i.e., the sum of squares of the coefficients being one) with maximum variance. In essence, principal component analysis is for dimensionality reduction, i.e., for transforming a number p of possibly correlated observable variables

$$\boldsymbol{X}=\left(X_1,\ldots,X_p\right)^{'}$$

into a smaller number u, $u < p$, of uncorrelated variables $Y_1,\ldots, Y_u$ called principal components to account for as much variability, or the explanatory power, as possible. The principal components $Y_1=\boldsymbol{\varphi}_1^{'}X,\ \ldots\ ,\ Y_u=\boldsymbol{\varphi}_u^{'}X$ are determined sequentially that given the first i-1 principal components, $i\leq u$, the ith principal component Y_i is determined by choosing $\boldsymbol{\varphi}_i$ that maximizes the variance of Y_i, i.e.,

$$var(Y_i)=\boldsymbol{\varphi}_i^{'}\Xi\boldsymbol{\varphi}_i$$

subject to the constraints $\boldsymbol{\varphi}_i^{'}\boldsymbol{\varphi}_j=0$ for all $j < i$ and $\boldsymbol{\varphi}_i^{'}\boldsymbol{\varphi}_i=1$, where Ξ is the overall variance-covariance matrix

$$\Xi=\Sigma+\sum_{i=1}^{k}w_i(\boldsymbol{\mu}_i-\overline{\boldsymbol{\mu}})(\boldsymbol{\mu}_i-\overline{\boldsymbol{\mu}})^{'}.$$

To be compatible with discriminant analysis, the constraints $\boldsymbol{\varphi}_i^{'}\boldsymbol{\varphi}_j=0$ for all $j < i$ and $\boldsymbol{\varphi}_i^{'}\boldsymbol{\varphi}_j=1$ are relaxed and let $\boldsymbol{\varphi}_i=\Sigma^{-\ 1/2}\xi_i,\ i\leq u$, where the vector

ξ_i satisfies the constraints that $\xi_i'\xi_i = 1$ and $\xi_i'\xi_j = 0$ for all $j < i$. As the norm $\boldsymbol{\varphi}_i'\boldsymbol{\varphi}_i = \xi_i'\Sigma^{-1}\xi_i$ is not restricted to one, we consider an objective function $g_2$that represents the normalized explanatory power of Y_i as follows

$$g_2(\xi_i) = \frac{\xi_i'\Sigma^{-\frac{1}{2}}\Xi\Sigma^{-\frac{1}{2}}\xi_i}{\xi_i'\Sigma^{-1}\xi_i}. \tag{89.10}$$

89.4 Multivariate Analysis of Industrial Bond Ratings

Pinches and Mingo developed a credit-scoring model based on the stepwise discriminant analysis procedure, in which a set of six key financial ratios with significant variability or explanatory power is chosen first from 35 candidate financial variables using factor analysis first. After the selection process, discriminant analysis is implemented to combine and weight the six key financial variables to produce credit scores with maximal discriminatory power that discriminate obligors among different benchmark ratings ex ante as much as possible. However, the object function of the factor analysis, similar to that of the principal component analysis g_1 in (89.10), and the object function g_2 of the K-group discriminant analysis in (89.6) contradict with each other. For this reason, the extracted factors with significant explanatory power might have very low discriminatory power. Or it might occur that extracted factors with insignificant explanatory power but have very high discriminatory power. Some compromised approach needs to be developed to overcome the problem. A goal programming technique, namely, the minimax goal programming technique by Ignizio and Romero (2003), is considered to be used in conjunction with the principal component and K-group discriminant analysis to find a compromise solution between the two objectives g_1 and g_2. The description of the minimax goal programming technique is given in Appendix 2.

Here, the rating data from Pinches and Mingo (1973) in which a total of 132 industrial corporate bonds rated B, Ba, Baa, A, Aa based on Moody's ratings from January 1, 1967, to December 31, 1968, are adopted for analysis. Four different approaches are compared with one another, they are (1) principal component analysis, (2) K-group discriminant analysis, (3) stepwise discriminant analysis by Pinches and Mingo, and (4) hybrid of principal component and discriminant analysis using minimax goal programming technique. The two object functions g_1 and g_2, or the discriminant power and explanatory power, of the four approaches are listed in Tables 89.2 and 89.3.

89.5 Summary

In this chapter, method and theory of the three multivariate techniques, namely, discriminant analysis, factor analysis, and principal component analysis, are discussed in detail. In addition, the stepwise principal component-discriminant analysis by Pinches and Mingo (1973) is improved using a goal programming technique. These methodologies are applied to determine useful financial ratios and the subsequent bond ratings.

Table 89.2 Classification using hybrid of principal component and discriminant analysis

	Actual rating					
Predicted rating	Aa	A	Baa	Ba	B	Total
In-sample						
Aa	10	2	0	0	0	12
A	1	14	1	0	0	16
Baa	0	1	12	5	0	18
Ba	0	0	2	19	0	21
B	0	0	0	0	17	17
Total	11	17	15	24	17	84
Out-sample						
Aa	2	3	0	0	0	5
A	1	5	2	0	0	8
Baa	0	1	6	1	0	8
Ba	0	0	2	17	0	19
B	0	0	0	2	6	8
Total	3	9	10	20	6	48

Table 89.3 Object functions of four approaches

	Model			
Criteria	Principal component analysis	K-group discriminant	Stepwise discriminant analysis	Hybrid discriminant analysis
In-sample				
g_1	54.80 %	4.01 %	10.55 %	11.79 %
g_2	48.81 %	88.10 %	63.10 %	85.71 %
Out-sample				
g_1	46.70 %	4.82 %	10.37 %	11.65 %
g_2	58.33 %	70.83 %	70.83 %	75.00 %

The analysis shows that the stepwise discriminant analysis fails to be an efficient solution as the hybrid approach using the goal programming technique outperforms it, which is a compromised solution for the maximization of the two objectives, namely, the maximization of the explanatory power and the maximization of discriminant power.

Appendix 1: Derivation of the Discriminant Function

Data composed of two samples of size N_1 and N_2 for two-group discriminant analysis must meet the following assumptions: (1) that the groups being investigated are discrete and identifiable, (2) that each observation in each group can be described by a set of measurements on m characteristics or variables, and (3) that these m variables have a multivariate normal distribution in each population.

In vector notation, the nth observation can be represented as an $m \times 1$ column vector of the form

$$X_n^{'} = (X_{1n}, X_{2n}, \cdots, X_{mn}),$$

where $n = 1, \ldots\ldots, N_1$ or $n = 1, \ldots\ldots\ldots, N_2$. Under these assumptions, the linear discriminant function can be defined as Eq. 89.11

$$Y_i = a_1 X_{1i} + a X_{2i} + \cdots + a_m X_{mi}. \tag{89.11}$$

The a_i's were chosen to maximize the ratio of the weighted between-group variance to the pooled within-group variance. Ladd (1966) has proposed a discriminant criterion E, as defined as Eq. 89.12, to determine the coefficients $a_1, a_2, \ldots, a_m$:

$$E = \frac{A^{'}DD^{'}A}{A^{'}CA} = \frac{A^{'}BA}{A^{'}CA} \tag{89.12}$$

where
$A' = [a_1, a_2, \cdots, a_m]$;
$D^{'} = \left[\overline{X}_{1,1} - \overline{X}_{1,2}, \overline{X}_{2,1} - \overline{X}_{2,2}, \overline{X}_{m,1} - \overline{X}_{m,2}\right]$;
$C =$ within-group variance matrix and
$DD' =$ between-group variance matrix.

There are two alternative methods (a) and (b) that can be used to derive the basic equation of discriminant analysis as defined in Eq. 89.13 in the following:

(a) Subsequent vector derivation method
Symbolically, we may find the derivative of E with the respect to the column vector of A and equate the result to the $m \times 1$ vector (see Tatsuoka (1971), p. 160–161). The vector equation thus obtained is

$$\frac{\partial E}{\partial A} = \frac{2\left[BA\left(A^{'}CA\right) - \left(A^{'}BA\right)CA\right]}{\left(A^{'}CA\right)^2} = 0.$$

Dividing both numerator and denominator of the middle member by $A'CA$ and using the definition of E in Eq. 89.12, this equation reduces to

$$\frac{2[BA - ECA]}{A^{'}CA} = 0,$$

which is equivalent to

$$(B - EC)A = 0. \tag{89.13}$$

(b) Long-hand method

If $m = 2$, then Eq. 89.12 can be rewritten as

$$E = \frac{b_{11}a_1^2 + b_{22}a_2^2 + 2b_{12}a_1a_2}{c_{11}a_1^2 + c_{22}a_2^2 + 2c_{12}a_1a_2}$$

where b_{11}, b_{22}, and b_{12} are elements of B and c_{11}, c_{21}, and c_{12} are elements of C in Eq. 89.2. Taking the partial derivative of E with respect to a_1, we obtain:

$$\begin{aligned}\frac{\partial E}{\partial a_1} &= \left[(2b_{11}a_1 + 2b_{12}a_2)\left(C_{11}a_1^2 + C_{22}a_2^2 + 2C_{12}a_1a_2\right.\right.\\ &\quad - \left(b_{11}a_1^2 + b_{22}a_2^2 + 2b_{12}a_1a_2\right)\left.\left(2C_{11}\mathrm{a}_1 + 2C_{12}\mathrm{a}_2{}^2\right)\right]\\ &\quad \times \left(C_{11}\mathrm{a}_1^2 + C_{22}\mathrm{a}_2^2 + 2C_{12}\mathrm{a}_1\mathrm{a}_2\right)^{-2}\\ &= 2[b_{11}a_1 + b_{12}a_2) - E(C_{11}a_1 + C_{12}a_2)]\\ &\quad \times \left(C_{11}\mathrm{a}_1^2 + C_{22}\mathrm{a}_2^2 + 2C_{12}\mathrm{a}_1\mathrm{a}_2\right)^{-1}.\end{aligned}$$

Setting this equation equal to zero and simplifying, we get $b_{11}a_1 + b_{12}a_2 = E(C_{11}a_1 + C_{12}a_2)$. Similarly, upon equating $\partial E/\partial a_2$ to zero and simplifying, we get $b_{21}a_1 + b_{22}a_2 = E(C_{21}a_1 + C_{22}a_2)$. Using matrix notation, we have

$$\begin{bmatrix} b_{11} & b_{12} \\ b_{21} & b_{22} \end{bmatrix} A = E \begin{bmatrix} C_{11} & C_{12} \\ C_{21} & C_{22} \end{bmatrix} A.$$

Therefore, we obtain Eq. 89.13, which can be used to formulate the characteristic equation for solving A as follows. If A^* maximizes the function shown in Eq. 89.11, then so does any $A^{**} = \kappa A^*$, where κ is a scalar. Substituting A^{**} for A^* simply multiplies both the numerator and denominator by κ^2. Because the coefficients themselves are not unique, there are several methods of calculating the discriminant function. Johnston and Dinardo (1996) show that the vector A^* that maximizes the ratio E is proportional to the vector A (i.e., $A = \kappa A^*$), which maximizes $G = A'DD'A$ subject to the constraint that $L = A'CA$, where L is an arbitrary constant. Let λ be a Lagrange multiplier and define

$$F = A'DD'A - \lambda\left(A'CA - L\right).$$

Setting the derivative of F with respect to A equal to zero yields

$$\frac{\partial F}{\partial A} = 0 = 2DD'a - 2\lambda CA \tag{89.14}$$

Where $D'A$ is a scalar, say H. Hence, Eq. 89.14 can be rewritten as (λ/H) $CA = D$, and thus

$$A\left(\frac{\lambda}{H}\right) = C^{-1}D = A_1.$$

It can be seen that A_1 is a solution to Eq. 89.13 since

$$\left(DD^{'} - \lambda C\right)A_1 = \left(DD^{'} - \lambda C\right)A\left(\frac{\lambda}{H}\right) = 0.$$

Alternatively, adding the term EBA to both sides of Eq. 89.13, one has $(1 + E)BA = E(B + C)A$ or

$$\left(B - E^{'}S\right)A = 0, \tag{89.15}$$

where $E' = E/(1 + E)$ and $S = (B + C)$. Now Eq. 89.15 can be used as an alternative of Eq. 89.13, and an alternative objective function can be defined as

$$E^{'} = \frac{A'DD'A}{A^{'}SA},$$

where $S = [S_{ij}] = \textit{the}\ m \times m$ matrix of S_{ij}; here S_{ij} is the sum of the cross products of $X_i - \overline{X}$ and $X_j - \overline{X}$. Following the same procedure mentioned above, we have

$$A\left(\frac{\lambda}{H}\right) = S^{-1}D = A_2.$$

Ladd (1966) has shown that A_2 is proportional to A_1. These results imply that the parameters of a two-group discriminant function can be estimated by using the related data of S and D. Let $Y = 1$ for observations in group 1 and $Y = 0$ for those in group 2; then, using the multiple regression technique, Ladd (1966) showed that the regression coefficient vector $A = S^{-1}MD$, where $M = [N_1N_2]/[N_1 + N_2]$, and N_1 and N_2 are total observations in group 1 and group 2, respectively. It follows straightforwardly showing that the parameters obtained from the eigenvalue method differ from those of the dummy regression method only by a constant M.

Appendix 2: Minimax Goal Programming Technique

The goal programming (GP) technique is a multi-objective optimization approach. Several classes of goal programming are developed, including weighted goal programming, lexicographic goal programming, and minimax goal programming. In this study, as a balance between discriminant power as well as explanatory power is desired, the minimax goal programming technique is introduced as follows. In the standard GP formulation, each of the objectives is given aspiration levels. Unwanted deviations from this aspiration levels are minimized in an achievement function. For the minimax goal programming, the maximal deviation δ from amongst the set of goals is minimized. There are numerous forms of minimax goal programming; here we adopt the one by Ignizio and Romero (2003) as follows:

$$\begin{aligned}
&Achievement\ \ function : Min\ \ \delta \\
&\quad s.t. \\
&\qquad \frac{g_1^{max} - g_1(\xi)}{\left(g_1^{max} - g_1^{min}\right)} \le \delta \\
&\qquad \frac{g_2^{max} - g_2(\xi)}{\left(g_2^{max} - g_2^{min}\right)} \le \delta \\
&Constraints : \xi' \xi = 1 \ \text{ and } \ \delta \ge 0
\end{aligned}$$

where ξ is the decision variable and the four values $g_1^{max} = g_1(\xi_1^*)$, $g_2^{min} = g_2(\xi_1^*)$, $g_2^{max} = g_2(\xi_2^*)$, and $g_1^{min} = g_1(\xi_2^*)$ and ξ_1^* and ξ_2^* are the optimal solutions for objectives g_1 and g_2, respectively.

References

Altman, E. I. (1968). Financial ratios, discriminant analysis, and the prediction of corporate bankruptcy. *Journal of Finance, 23*, 589–609.

Altman, E. I. (1983). Predicting corporate bankruptcy: The Z-Score model. In: *Corporate financial distress: A complete guide to predicting, avoiding and dealing with bankruptcy*. New York: Wiley.

Altman, E. I., & Eisenbeis, R. A. (1968). Financial applications of discriminant analysis: A clarification. *Journal of Financial and Quantitative Analysis, 13*, 185–195.

Altman, E. I., Haldeman, R., & Narayanan, P. (1977). ZETA analysis, a new model for bankruptcy classification. *Journal of Banking and Finance, 1*, 29–54.

Anderson, T. W. (2003). *An introduction to multivariate statistical analysis* (3rd ed.). New York: Wiley.

Churchill, G. A., & Iacobucci, D. (2004). *Marketing research: Methodological foundations* (9th ed.). Mason: South-Western College Publishing.

Eisenbeis, R. A., & Avery, R. B. (1972). *Discriminant analysis and classification procedure*. Lexington: Lexington Books.

Ignizio, J. P., & Romero, C. (2003). Goal programming. *Encyclopedia of Information Systems, 2*, 489–500.

Johnston, J., & Dinardo, J. (1996). *Econometrics methods* (4th ed.). New York: McGraw-Hill.

Joreskog, K. G. (1967). Some contributions to maximum-likelihood factor analysis. *Psychometrika, 32*, 443–482.

Ladd, G. W. (1966). Linear probability functions and discriminant functions. *Econometrica, 34*, 873–885.

Lawley, D. N. (1940). The estimation of factor loadings by the method of maximum likelihood. *Proceedings of the Royal Society of Edinburg, Series A, 62*, 64–82.

Lawley, D. N., & Maxwell, A. E. (1963). *Factor analysis as a statistical method*. London: Butterworth.

Pinches, G. E., & Mingo, K. A. (1973). A multivariate analysis of industrial bond ratings. *J Fin, 28*, 1–18.

Tatsuoka, M. M. (1988). *Multivariate analysis: Techniques for educational and psychological research* (2nd ed.). New York: Wiley.

Implied Volatility: Theory and Empirical Method

90

Cheng-Few Lee and Tzu Tai

Contents

Abstract

The estimation of the implied volatility is the one of most important topics in option pricing research. The main purpose of this chapter is to review the different theoretical methods used to estimate implied standard deviation and to show how the implied volatility can be estimated in empirical work. The OLS method for estimating implied standard deviation is first introduced, and the formulas derived by applying a Taylor series expansion method to Black–Scholes option pricing model are also described. Three approaches of estimating

C.-F. Lee (✉)
Department of Finance and Economics, Rutgers Business School, Rutgers, The State University of New Jersey, Piscataway, NJ, USA

Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
e-mail: lee@business.rutgers.edu; cflee@business.rutgers.edu

T. Tai
Department of Finance and Economics, Rutgers, The State University of New Jersey, Piscataway, NJ, USA
e-mail: tzutai@pegasus.rutgers.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_90,

implied volatility are derived from one, two, and three options, respectively. Regarding to these formulas with the remainder terms, the accuracy of these formulas depends on how an underlying asset is close to the present value of exercise price in an option. The formula utilizing three options for estimating implied volatility is more accurate rather than other two approaches.

In empirical work, we use call options on S&P 500 index futures in 2010 and 2011 to illustrate how MATLAB can be used to deal with the issue of convergence in estimating implied volatility of future options. The results show that the time series of implied volatility significantly violates the assumption of constant volatility in Black–Scholes option pricing model. The volatility parameter in the option pricing model fluctuates over time and therefore should be estimated by the time series and cross-sectional model.

Keywords

Implied volatility • Implied standard deviation (ISD) • Option pricing model • MATLAB • Taylor series expansion • Ordinary least squares (OLS) • Black–Scholes model • Options on S&P 500 index futures

90.1 Introduction

The volatility of the return of the underlying asset is the important factor in option pricing model (see Merton 1973; Black and Scholes 1973). However, the standard deviation of the underlying asset return cannot be observed directly (See Merton 1976; Macbeth et al. 1979; Chance 1986). The estimation of the implied volatility of the underlying asset in option framework becomes the one of most important topics in option pricing research. There are two main methods developed in the finance literature to estimate the standard deviation of the underlying asset in option framework: (1) the historical standard deviation and (2) the implied standard deviation (called ISD hereafter) derived from the Black–Scholes' option pricing model framework (See Hull 2011).

Garman and Klass (1980) study the historical standard deviation by using open, high, low prices, and closed prices' data to estimate the standard deviation. To support the use of historical standard deviation for implied standard deviation in option pricing model requires that the underlying asset's rate of return is stationary over the option's life, which contradicts the time-varying standard deviation documented by Schwert (1989).

Since the Black–Scholes' option pricing model is a nonlinear equation, an explicit analytic solution for the ISD is not available in the literature (except for at-the-money call), and numerical methods are used to approximate the ISD (see Latane and Rendleman 1976; Beckers 1981; Manaster and Koehler 1982; Brenner and Subrahmanyam 1988; Lai et al. 1992; Chance 1996; Hallerback 2004; Corrado and Miller 1996, 2004; Li 2005). The derivation and use of the ISD for an option as originated by Latane and Rendleman (1976) have become a widely used methodology for variance estimation. By applying the Newton-Raphson method, Manaster and Koehler (1982) provide an iterative algorithm for the ISD. Brenner and Subrahmanyam (1988) applied Taylor series expansion at zero base to the cumulative normal function in pricing Black–Scholes option pricing model up to the first-order term and set the underlying asset price to equal the present value of exercise price to solve the ISD. Lai et al. (1992) derive a closed-form solution for the ISD in terms of the delta $\partial C/\partial S$ and $\partial C/\partial E$. Following the same approach as Brenner and Subrahmanyam, Corrado and Miller (1996, 2004) utilize the cumulative normal function at zero to the first-order term to derive a quadratic equation of the ISD. Then the ISD can be obtained by solving the quadratic equation. Hallerback (2004) also derives an improved formula, which is similar to Corrado and Miller's formula (1996), to compute the ISD. Later, Li (2005) bases on Brenner and Subrahmanyam's approach to expand the expression to third-order term and solve for the ISD with a cubic equation. Since Li includes third-order term in the Taylor expansion on the cumulative normal distribution in his derivation, Li claims that his formula of ISD provides a consistently more accurate estimate of the true ISD than that of Brenner and Subrahmanyam's formula.

However, the fact that there are as many estimated ISD of an underlying asset as the number of different exercise price in options violates the constant variance assumption used in deriving the Black–Scholes' option pricing model. Chance (1996) assumes different exercise prices result in different ISDs, which violate the constant variance assumption used in deriving the Black–Scholes' option pricing model. Under the existence of a call at-the-money assumption, Chance uses Brenner and Subrahmanyam's formula to calculate the at-the-money's ISD. Chance then applies Taylor series expansion to the difference of the call options in terms of the first and the second-order terms. The drawback of Chance's method is the constraint of the use only for at-the-money option price. In other words, if the underlying asset price deviates from the present value of the exercise price and the call option price is not available (or unobservable) in the market, then Chance's formula for the ISD may not apply. Later, Ang et al. (2009, 2012) relax the constraint in Chance's method and develop three formulas which depend on a Taylor series expansion and utilize single, two, and three options, respectively, to estimate implied volatility.

The purpose of this chapter is to review the different theoretical methods used to estimate ISD and to show how the implied volatility can be estimated in empirical work. We use the data from options on S&P 500 index futures in 2010 and 2011 to illustrate how MATLAB can be used to deal with the issue of convergence in estimating implied volatility of options on index futures. This chapter is organized as follows. In Sect. 90.2, we review the OLS method used in estimation of the ISD in Black–Scholes' option pricing model and expand this method to estimate the implied volatility of the underlying asset for options on the index futures. Then, in Sect. 90.3, we introduce the formulas of implied volatility developed by Ang et al. (2012) which apply a Taylor series expansion to the Black–Scholes option pricing model. The process and results of empirical work on estimating ISD for options on S&P 500 index futures are shown in Sect. 90.4. Finally, Sect. 90.5 represents the conclusions of this study.

90.2 Estimating the Implied Standard Deviation with OLS Method

Black and Scholes (1973) and Merton (1973) derive the European call option pricing model on a stock as follows:

$$\begin{aligned} C &= SN(d_1) - Ke^{-r\tau}N(d_2) \\ d_1 &= [1\text{n}(S/K) + (r+\sigma^2/2)\tau]/\sigma\sqrt{\tau} \\ d_2 &= d_1 - \sigma\sqrt{\tau}, \end{aligned} \tag{90.1}$$

where C is the call premium, S is the underlying stock price, K is the exercise price, r is the instantaneous risk-free rate, τ is the time to the maturity, σ is the standard deviation of the underlying asset rate of return on annual basis, and N(x) is the standard cumulative normal distribution function up to x.

The sensitivities, or first partial derivatives, of the call option formula in Eq. 90.1 with respect to the change of the volatility of the underlying stock can be derived as

$$\frac{\partial C}{\partial \sigma} = S\sqrt{\tau}N'(d_1) = \frac{S\sqrt{\tau}}{S\sqrt{2\pi}}e^{-d_1^2/2} \tag{90.2}$$

where N′(x) is the standard normal probability density function at value x. Equation 90.2 shows the positive relationship between the call option price and the volatility of the underlying stock. Since a call option has no downside risk (except for its cost), increasing risk of the underlying stock simply enlarges the probability that the option will end up in the money by expiration (hence, with a larger intrinsic value).

The OLS method for estimating implied standard deviation is first proposed by Whaley (1982). Although Whaley's original intent for this method was to improve upon the existing weighting techniques, his ordinary least squares (OLS) approach can also be used to derive the implied standard deviations (ISD) for call options. To begin the development of his method, Whaley applies a Taylor series expansion around some initial value of the standard deviation and omits higher-order terms. Mathematically, this is expressed as

$$C_{j,t}^M = C_{j,t}^T(\sigma_0) + \left(\frac{\partial C_{j,t}^T}{\partial \sigma}\bigg| \sigma = \sigma_0\right)(\sigma_s - \sigma_0) + e_{j,t}, \tag{90.3}$$

$$(j = 1, 2, \ldots, \text{J})$$

where $C_{j,t}^M$ denotes the market price for the option j at time t, $C_{j,t}^T$ is the theoretical model price estimated by Eq. 90.1 for the option j at time t based on an estimated value for the ISD (σ_0), σ_0 is the estimated ISD evaluated from some initialization value up to some minimum level of tolerance of error, σ_{s} denote the true or actual

ISD which we are looking for, and $e_{j,t}$ is the random disturbance term for option j at time t. By rearranging Eq. 90.3, we can obtain

$$\left[C^M_{j,t} - C^T_{j,t}(\sigma_0)\right] + \sigma_0\left(\frac{\partial C^T_{j,t}}{\partial\sigma}\bigg|\sigma = \sigma_0\right) = \sigma_s\left(\frac{\partial C^T_{j,t}}{\partial\sigma}\bigg|\sigma = \sigma_0\right) + e_{j,t} \qquad (90.4)$$
$$(j = 1, 2, \ldots, \mathrm{J})$$

since $C^M_{j,\,t}$ is observable in market $\left(\frac{\partial C^T_{j,t}}{\partial\sigma}\Big|\sigma = \sigma_0\right)$ and $C^T_{j,\,t}(\sigma_0)$ can be evaluated at any given value σ_0 by using Eqs. 90.1 and 90.2.

Whaley (1982) then applies OLS, which minimizes the sum of squared residuals, to achieve a single, weighted σ from the options on a particular stock. The actual estimation procedures begin from a linearization of the option pricing model around 0, and then OLS is applied to Eq. 90.4. The process thus proceeds in an iteration manner until the estimated ISD $\hat{\sigma}_s$ satisfies an acceptable tolerance of

$$\left|\frac{\hat{\sigma}_s - \sigma_0}{\sigma_0}\right| < Q, \qquad (90.5)$$

where Q is a small positive number where Whaley(1982) uses Q equal to 0.0001 as the acceptable tolerance of estimated error and $\hat{\sigma}_s$ is the estimate for the true ISD σ_s for the market option price. If the tolerance criterion is not satisfied, $\hat{\sigma}_s$ becomes the new initialization value and the OLS procedure is repeated.

The OLS method also can be applied to estimate the ISD for options on index future with the similar procedure of a Taylor series expansion (See Wolf 1982; Park et al. 1985; Ramaswamy et al. 1985; Brenner et al. 1985). The call options on index future derived by Black (1975, 1976) are given by

$$\begin{aligned} C^F_t &= e^{-r\tau}[F_t N(d_3) - KN(d_4)] \\ d_3 &= \left[1\mathrm{n}(F_t/K) + \left(\sigma_f^2/2\right)\tau\right]/\sigma_f\sqrt{\tau} \\ d_4 &= d_3 - \sigma_f\sqrt{\tau}, \end{aligned} \qquad (90.6)$$

where C^F_t is the model price for a call option on index future at time t, F_t is the underlying index future price at time t, K is the exercise price of the call option on index future, τ is the option's remaining time to maturity in terms of a year, r is the continuous annualized risk-free rate, σ_f^2 is the instantaneous variance of returns of the underlying index future contract over the remaining life of the option, and N(x) is the standard cumulative normal distribution function up to x.

There is similar procedure with Whaley's method to calculate the ISD for options on S&P 500 index futures. The ISD is obtained by first choosing an initial estimate, σ_0, and then using Eq. 90.7 to iterate towards the correct value as follows:

$$C^F{}_{t,j} - C^F{}_{t,j}(\sigma_0) = (\sigma_1 - \sigma_0)\left(\frac{\partial C^F_{j,t}}{\partial\sigma}\bigg|\sigma = \sigma_0\right), \qquad (90.7)$$

where $C^F{}_{t,j}$ denotes the market price of call option j at time t, $C^F{}_{t,j}$ (σ_0) is the theoretical price of call option j at time t given σ equal to $\sigma_{0,}$ σ_0 is the initialized estimate of the ISD, σ_1 = estimate of the true ISD from iteration, $\left(\frac{\partial C^F_{t,j}}{\partial \sigma}\middle| \sigma = \sigma_0\right)$ is the partial derivative of the call option on index future with respect to the standard deviation evaluated at a $\sigma_{0.}$ In the context of the Black (1976) option pricing model, the partial with respect to the standard deviation of the underlying index future can be expressed explicitly as

$$\frac{\partial C^F_{t,j}}{\partial \sigma} = F_t e^{-r\tau}\sqrt{\tau}N'(d_3) = F_t e^{-r\tau}\frac{\sqrt{\tau}}{\sqrt{2\pi}}e^{-d_3^2/2}, \tag{90.8}$$

where d_3 is defined as in Eq. 90.6 and N′(x) is the standard normal probability density function at value x. The partial derivative formula in Eq. 90.8 is also called Vega of a call option on index futures which is represented the rate of change of the value of a call option on index futures with respect to the volatility of the underlying index futures. The iteration proceeds by reinitializing σ_0 to equal σ_1 at each successive stage until an acceptable tolerance level in Eq. 90.5 is attained.

90.3 Estimating the Implied Standard Deviation with Taylor Series Expansion Method

In this section, we first introduce the exact closed-form solution in for the ISD under the condition that the underlying asset price equals the present value of the exercise price. Then we discuss Ang et al.'s (2012) alternative formulas to estimate the ISD by applying a Taylor series expansion to the Black–Scholes option pricing model under the relaxation of the previous restrictive condition.

When the underlying stock price equals the present value of the exercise price (i.e., $S = Ke^{-r\tau}$), the Eq. 90.1 can be reduced as follows:

$$\begin{aligned} C &= S[N(\sigma\sqrt{\tau}/2) - N(-\sigma\sqrt{\tau}/2)] \\ &= S[1 - 2N(-\sigma\sqrt{\tau}/2)] \\ &= S[2N(\sigma\sqrt{\tau}/2) - 1]. \end{aligned} \tag{90.9}$$

Based on the characteristics of existence and uniqueness of the inverse cumulative normal distribution, an exact closed-form solution for the ISD in Eq. 90.9 can be derived as

$$\sigma\sqrt{\tau} = 2N^{-1}[(S + C)/(2S)]. \tag{90.10}$$

Ang et al. (2012) apply Taylor's formula to the cumulative normal functions in Eq. 90.1 at base $\ln(S/Ke^{-r\tau})/(\sigma\sqrt{\tau})$ up to the second-order terms, then the

European call option in Eq. 90.1 can be rearranged as a quadratic equation of $\sigma\sqrt{\tau}$ plus the remainder term as follows[1]:

$$\begin{aligned}&\sigma^2\tau[4(S+Ke^{-r\tau})-(S-Ke^{-r\tau})\ln(S/Ke^{-r\tau})]-4\sigma\sqrt{\tau}\sqrt{2\pi}(2C-S+Ke^{-r\tau})\\&+8\ln(S/Ke^{-r\tau})\left[(S-Ke^{-r\tau})\left(1+(\ln(S/Ke^{-r\tau})/4)^2\right)-(S+Ke^{-r\tau})\ln(S/Ke^{-r\tau})/4\right]+\varepsilon=0.\end{aligned} \tag{90.11}$$

Dropping the remainder term ε, the ISD can be obtained by solving the root of quadratic equation function in Eq. 90.11. Since Ang et al. (2012) utilize four times of a Taylor series expansion method to derive the quadratic function of a European call option and the remainder terms are omitted, the ISD calculated by Eq. 90.11 is not an exact formula. Therefore, the effectiveness of using Eq. 90.11 to estimate the ISD depends on the deviation of the underlying stock price (S) from the present value of exercise price ($Ke^{-r\tau}$).

Moreover, Ang et al. (2012) derive the second alternative formula for estimating ISD by using two call options,C_1 and C_2, on the same underlying stock but at different exercise, K_1 and K_2, respectively (here we assume $K_1 < K_2$). By applying Taylor series expansion to Eq. 90.1 for two call options at K_2 for C_1 and at K_1 for C_2, respectively, we can obtain

$$C_1 = C_2 - e^{-r\tau}N\left(\ln(S/K_2e^{-r\tau})/\left(\sigma\sqrt{\tau}\right) - \sigma\sqrt{\tau}/2\right)(K_1-K_2)+\varepsilon_1, \tag{90.12}$$

$$C_2 = C_1 - e^{-r\tau}N\left(\ln(S/K_1e^{-r\tau})/\left(\sigma\sqrt{\tau}\right) - \sigma\sqrt{\tau}/2\right)(K_2-K_1)+\varepsilon_2. \tag{90.13}$$

Here ε_1 and ε_2 are the remainder terms of C_1 at K_2 and C_2 at K_1 from Eq. 90.1. Dividing both sides of Eqs. 90.12 and 90.13 by $e^{-r\tau}$ (K_2-K_1) and simple manipulations produce the same left-hand side of $(C_1–C_2)/e^{-r\tau}(K_2-K_1)$.

Then applying the inverse function of cumulative normal function on both sides and after using the Taylor's formula yields the following equations:

$$N^{-1}[(C_1-C_2)/e^{-r\tau}(K_2-K_1)] = \ln(S/K_1e^{-r\tau})/\left(\sigma\sqrt{\tau}\right) - \sigma\sqrt{\tau}/2+\eta_1, \tag{90.14}$$

$$N^{-1}[(C_1-C_2)/e^{-r\tau}(K_2-K_1)] = 1\mathrm{n}(S/K_2e^{-r\tau})/\left(\sigma\sqrt{\tau}\right) - \sigma\sqrt{\tau}/2+\eta_2, \tag{90.15}$$

where η_1 and η_2 are the remainder terms of Taylor's formulas derived from Eqs. 90.12 and 90.13, respectively. After combining Eqs. 90.14 and 90.15 and dropping the remainder terms (η_1 + η_2), the quadratic function of $\sigma\sqrt{\tau}$ can be shown as

[1]The details of the derivation of Eq. 90.10 can be found in Ang et al. (2012) paper.

$$\left(\sigma\sqrt{\tau}\right)^2 + 2N^{-1}[(C_1 - C_2)/e^{-r\tau}(K_2 - K_1)]\left(\sigma\sqrt{\tau}\right) - \ln\left[S^2/\left(e^{-2r\tau}K_1K_2\right)\right] = 0. \quad (90.16)$$

Thus, the ISD can be solved as

$$\sigma\sqrt{T} = \begin{cases} -N^{-1}[(C_1 - C_2)/e^{-r\tau}(K_2 - K_1)] + \sqrt{\zeta} & \text{when } \mathrm{S} > \mathrm{K}_1 \\ -N^{-1}[(C_1 - C_2)/e^{-r\tau}(K_2 - K_1)] - \sqrt{\zeta} & \text{when } \mathrm{S} \le \mathrm{K}_1 \le \mathrm{K}_2 \end{cases}$$
$$\zeta = \left[N^{-1}((C_1 - C_2)/e^{-r\tau}(K_2 - K_1))\right]^2 + \ln\left(S^2/e^{-2r\tau}K_1K_2\right). \quad (90.17)$$

It is clear that if stock price is less than the lower exercise K_1 (i.e., then both call options are out of the money), and if we had chosen the value with the plus sign of $\sqrt{\zeta}$ in Eq. 90.17, ISD calculated by Eq. 90.17 will be overstated. The advantage of this formula is that a sufficient condition to calculate ISD by Eq. 90.17 only requires that there existed any two consecutive call option values with different exercise prices. But, the accuracy of this formula will depend on the magnitude of the deviation between these two exercise prices.

Ang et al. (2012) further extend this approach to include a third option to derive the third formula. Similar to Eq. 90.16, if there is a third call option C_3 with the exercise price K_3, then the following Eq. 90.18 must hold for K_2, K_3 and C_2, C_3.

$$\left(\sigma\sqrt{\tau}\right)^2 + 2N^{-1}[(C_2 - C_3)/e^{-r\tau}(K_3 - K_2)]\left(\sigma\sqrt{\tau}\right) - \ln\left[S^2/\left(e^{-2r\tau}K_2K_3\right)\right] = 0. \quad (90.18)$$

Given the constant variance assumption in Black and Scholes option model, the following Eq. 90.19 is thus derived by subtracting Eq. 90.18 from Eq. 90.16 as follows:

$$\sigma\sqrt{\tau} = \ln(K_3/K_1)/\left[2\left(N^{-1}((C_1 - C_2)/e^{-r\tau}(K_2 - K_1)) - N^{-1}((C_2 - C_3)/e^{-r\tau}(K_3 - K_2))\right)\right]. \quad (90.19)$$

An advantage of using Eq. 90.19 rather than Eq. 90.17 to estimate the ISD is to circumvent the sign issue that appears in Eq. 90.17. However, a drawback of using Eq. 90.19 is that there must exist at least three instead of two call options for Eq. 90.17. Equation 90.19 provides a simple formula to calculate ISD because all option values and exercise price are given and the inverse function of the standard cumulative normal function also available in the Excel spreadsheet. Ang et al. (2012) state that this third formula in Eq. 90.19 is more accurate method for estimating ISD based on their simulation results.

90.4 Illustration of Estimating Implied Standard Deviation by MATLAB

The data for this study for estimating ISD include the call options on the S&P 500 index futures which are traded at the Chicago Mercantile Exchange (CME).[2] According to Eq. 90.6, we need the information of market call option price on S&P 500 index, the annualized risk-free rate, S&P 500 index futures price, exercise price, and maturity date on the contracts as input variables to calculate the ISD of call option on S&P 500 index futures. Daily closed-price data of S&P 500 index futures and options on S&P 500 index futures was gathered from Datastream for two periods of time: the options expired on March, June, and September, 2010; options expired on March, June, and September, 2011; and the S&P 500 index future from October 1, 2008, to November 4, 2011. The S&P 500 spot price is based on the closed price of S&P 500 index on Yahoo! Finance[3] during the same period of S&P 500 index future data. The risk-free rate used in Black model is based on 3-month Treasury bill from Federal Reserve Bank of St. Louis.[4] The selection of these futures option contracts is based on the length of trading days. The futures options expired on March, June, September, and December have over 1 year trading date (above 252 observations), and other options only have more or less 100 observations. Therefore, we only choose the futures options with longer trading period to investigate the distributional statistics of these ISD series. Studying two different time periods (2010 and 2011) of call options on S&P 500 index futures will allow the examination of ISD characteristics and movements over time as well as the effects of different market climates.

The tolerance level used is the same formula as shown in Eq. 90.5, and let the tolerance level Q equal to 0.000001 as follows:

$$\left| \frac{\sigma_1 - \sigma_0}{\sigma_0} \right| < .000001$$

This chapter utilized financial toolbox in MATLAB to calculate the implied volatility for futures option that the code of function is as follows[5]:

$$\text{Volatility} = \text{blsimpv}(\text{Price},\ \ \text{Strike},\ \ \text{Rate},\ \ \text{Time}, \text{Value}, \text{Limit},\ \ \text{Tolerance},\ \ \text{Class})$$

[2]Nowadays Chicago Mercantile Exchange (CME), Chicago Board of Trade (CBOT), New York Mercantile Exchange (NYMEX), and Commodity Exchange (COMEX) are merged and operate as designated contract markets (DCM) of the CME Group which is the world's leading and most diverse derivatives marketplace. Website of CME group: http://www.cmegroup.com/

[3]Website of Yahoo! Finance is as follows: http://finance.yahoo.com

[4]Website of Federal Reserve Bank of St. Louis: http://research.stlouisfed.org/

[5]The syntax and the code from m-file source of MATLAB for Implied Volatility Function of Futures Options are represented in Appendix 1. The detailed information of the function and example of calculating the implied volatility for futures option also can be referred on MathWorks website: http://www.mathworks.com/help/toolbox/finance/blkimpv.html

where the blsimpv is the function name in MATLAB; Price, Strike, Rate, Time, Value, Limit, Tolerance, and Class are input variables; Volatility is the annualized ISD (also called implied volatility). The advantages of this function are the allowance of the upper bound of implied volatility (Limit variable) and the adjustment of the implied volatility termination tolerance (Tolerance variable), in general, equal to 0.000001.

A summary of the ISD distributional statistics for S&P 500 index futures call options in 2010 and 2011 appears in Table 90.1. The most noteworthy feature from this table is the significantly different mean values of the ISD that occur for different exercise prices. The means and variability of the ISD in 2010 and 2011 appear to be inversely related to the exercise price. Comparing the mean ISDs across time periods, it is quite evident that the ISDs in 2011 are significantly smaller. Also, the time-to-maturity effect observed by Park and Sears (1985) can be identified. The September options in 2011 possess higher mean value of the ISD than those maturing in June and March with the same strike price.

The other statistical measures listed in Table 90.1 are the relative skewness and relative kurtosis of the ISD series, along with the studentized range. Skewness measures lopsidedness in the distribution and might be considered indicative of a series of large outliers at some point in the time series of the ISDs. Kurtosis measures the peakedness of the distribution relative to the normal and has been found to affect the stability of variance (see Lee and Wu 1985). The studentized range gives an overall indication as to whether the measured degrees of skewness and kurtosis have significantly deviated from the levels implied by a normality assumption for the ISD series.

Although an interpretation of the effects of skewness and kurtosis on the ISD series needs more accurate analysis, a few general observations are warranted at this point. Both 2010 and 2011 ISD's statistics present a very different view of normal distribution, certainly challenging any assumptions concerning normality in Black–Scholes option pricing model framework. Using significance tests on the results of Table 90.1 in accordance with Jarque–Bera test, the 2010 and 2011 skewness and kurtosis measures indicate a higher proportion of statistical significance. We also utilize simple back-of-the-envelope test based on the studentized range to identify whether the individual ISD series approximate a normal distribution. The studentized range larger than 4 in both 2010 and 2011 indicates that a normal distribution significantly understates the maximum magnitude of deviation in individual ISD series.

As a final point to this brief examination of the ISD skewness and kurtosis, note the statistics for MAR10 1075, MAR11 1200, and MAR11 1250 contracts. The relative size of these contract's skewness and kurtosis measures reflect the high degree of instability that its ISD exhibited during the last 10 days of the contract's life. Such instability is consistent across contracts. However, these distortions remain in the computed skewness and kurtosis measures only for these particular contracts to emphasize how a few large outliers can magnify the size of these statistics. For example, the evidence that S&P 500 future price jumped on

Table 90.1 Distributional statistics for the ISD series of call options on S&P 500 index futures

Option series[a]	Mean	Std. dev.	CV[b]	Skewness	Kurtosis	Studentized range[c]	Observations
Call futures options in 2010							
MAR10 1075 (C070WC)	0.230	0.032	0.141	2.908	14.898	10.336	251
JUN10 1050 (B243UE)	0.263	0.050	0.191	0.987	0.943	6.729	434
JUN10 1100 (B243UF)	0.247	0.047	0.189	0.718	−0.569	4.299	434
SEP10 1100 (C9210T)	0.216	0.024	0.111	0.928	1.539	6.092	259
SEP10 1200 (C9210U)	0.191	0.022	0.117	0.982	2.194	6.178	257
Call futures options in 2011							
MAR11 1200 (D039NR)	0.206	0.040	0.195	5.108	36.483	10.190	384
MAR11 1250 (D1843V)	0.188	0.027	0.145	3.739	25.527	10.636	324
MAR11 1300 (D039NT)	0.176	0.021	0.118	1.104	4.787	8.588	384
JUN11 1325 (B513XF)	0.165	0.016	0.095	−1.831	12.656	10.103	200
JUN11 1350 (A850CJ)	0.161	0.018	0.113	−0.228	1.856	8.653	234
SEP11 1250 (B9370T)	0.200	0.031	0.152	2.274	6.875	7.562	248
SEP11 1300 (B778PK)	0.185	0.024	0.131	2.279	6.861	7.399	253
SEP11 1350 (B9370V)	0.170	0.025	0.147	2.212	5.848	6.040	470

[a]Option series contain the name and code of futures options with information of the strike price and the expired month, for example, SEP11 1350 (B9370V) represents that the futures call option is expired on September 2011 with the strike price $1,350, and the parentheses is the code of this futures option in Datastream

[b]CV represents the coefficient of variation that is standard deviation of option series divided by their mean value

[c]Studentized range is the difference of the maximum and minimum of the observations divided by the standard deviation of the sample

January 18, 2010, and plunged on February 2, 2011, causes the ISD of these particular contracts sharply increasing on that dates. Thus, while still of interest, any skewness and kurtosis measures must be calculated and interpreted with caution.

One difficulty in discerning the correct value for the volatility parameter in the option pricing model is due to its fluctuation over time. Therefore, since an accurate estimate of this variable is essential for correctly pricing an option, it would seem that time series and cross-sectional analysis of this variable would be as important as the conventional study of security price movements. Moreover, by examining the ISD series of each call options on S&P 500 index futures over time as well as within different time sets, the unique relationships between the underlying stochastic process and the pricing influences of differing exercise prices, maturity dates, and market sentiment (and, indirectly, volume), might be revealed in a way that could be modeled more efficiently. Therefore, we should consider autoregressive–moving-average (ARMA) models or cross-sectional time series regression models to analyze the ISD series and forecast the price of call options on S&P 500 index futures by predicting the future ISD of these options.

90.5 Summary and Concluding Remarks

The research in estimation of the implied volatility becomes the one of most important topics in option pricing research because the standard deviation of the underlying asset return, which is the important factor in Black–Scholes' option pricing model, cannot be observed directly. The purpose of this chapter is to review the different theoretical methods used to estimate implied standard deviation and to show how the implied volatility can be estimated in empirical work. We review the OLS method and a Taylor series expansion method for estimating the ISD in previous literature. Three formulas for the estimation of the ISD by applying a Taylor series expansion method to Black–Scholes option pricing model can be derived from one, two, and three options, respectively. Regarding to these formulas with the remainder terms in a Taylor series expansion method, the accuracy of these formulas depends on how an underlying asset is close to the present value of exercise price in an option.

In empirical work, we illustrate how MATLAB can be used to deal with the issue of estimating implied volatility for call options on S&P 500 index futures in 2010 and 2011. The results show that the time series of implied volatility significantly violate the assumption of constant volatility in Black–Scholes option pricing model. The skewness and kurtosis measures reflect the instability and fluctuation of the ISD series over time. Therefore, in the future research in the ISD, we should consider autoregressive–moving-average (ARMA) models or cross-sectional time series regression models to analyze and predict the ISD series to forecast the future price of call options on S&P 500 index futures.

Appendix 1: The Syntax and Code for Implied Volatility Function of Futures Options in MATLAB

The function name of estimating implied volatility for European call options on index futures in this chapter are as below:

Syntax

$$\text{Volatility} = \text{blsimpv}(\text{Price}, \text{Strike}, \text{Rate}, \text{Time}, \text{Value}, \text{Limit}, \ldots \text{Tolerance}, \text{Class})$$

The input variables that can be a scalar, vector, or matrix in the function of estimating implied volatility are described in Table 90.2

The code from m-file source of MATLAB for implied volatility function of futures options is shown as below:

```
function volatility = blkimpv(F, X, r, T, value, varargin)
% BLKIMPV Implied volatility from Black's model for futures options.
% Compute the implied volatility of a futures price from the market
% value of European futures options using Black's model.
%
% Volatility = blkimpv(Price, Strike, Rate, Time, Value)
% Volatility = blkimpv(Price, Strike, Rate, Time, Value, Limit, ...
% Tolerance, Class)
%
% Optional Inputs: Limit, Tolerance, Class.
%
% Inputs:
% Price - Current price of the underlying asset (i.e., a futures contract).
%
```

Table 90.2 The description of input variables used in blsimpv function in MATLAB

Price	Current price of the underlying asset (a futures contract)
Strike	Exercise price of the futures option
Rate	Annualized, continuously compounded risk-free rate of return over the life of the option, expressed as a positive decimal number
Time	Time to expiration of the option, expressed in years
Value	Price of a European futures option from which the implied volatility of the underlying asset is derived
Limit (optional)	Positive scalar representing the upper bound of the implied volatility search interval. If Limit is empty or unspecified, the default = 10, or 1,000 % per annum
Tolerance (optional)	Implied volatility termination tolerance. A positive scalar. Default = 1e-6
Class (optional)	Option class (call or put) indicating the option type from which the implied volatility is derived. May be either a logical indicator or a cell array of characters. To specify call options, set Class = true or Class = {'call'}; to specify put options, set Class = false or Class = {'put'}. If Class is empty or unspecified, the default is a call option

```
% Strike - Strike (i.e., exercise) price of the futures option.
%
% Rate - Annualized continuously compounded risk-free rate of return
% over the life of the option, expressed as a positive decimal number.
%
% Time - Time to expiration of the option, expressed in years.
%
% Value - Price (i.e., value) of a European futures option from which
% the implied volatility is derived.
%
% Optional Inputs:
% Limit - Positive scalar representing the upper bound of the implied
% volatility search interval. If empty or missing, the default is 10,
% or 1000% per annum.
%
% Tolerance - Positive scalar implied volatility termination tolerance.
% If empty or missing, the default is 1e-6.
%
% Class - Option class (i.e., whether a call or put) indicating the
% option type from which the implied volatility is derived. This may
% be either a logical indicator or a cell array of characters. To
% specify call options, set Class = true or Class = {'call'}; to specify
% put options, set Class = false or Class = {'put'}. If empty or missing,
% the default is a call option.
%
% Output:
% Volatility - Implied volatility derived from European futures option
% prices, expressed as a decimal number. If no solution is found, a
% NaN (i.e., Not-a-Number) is returned.
```

```
%
% Example:
% Consider a European call futures option trading at
$1.1166, with an
% exercise prices of $20 that expires in 4 months. Assume
the current
% underlying futures price is also $20 and that the risk-
free rate is 9%
% per annum. Furthermore, assume we are interested in
implied volatilities
% no greater than 0.5 (i.e., 50% per annum). Under these
conditions, any
% of the following commands
%
% Volatility = blkimpv(20, 20, 0.09, 4/12, 1.1166, 0.5)
% Volatility = blkimpv(20, 20, 0.09, 4/12, 1.1166, 0.5,
[], {'Call'})
% Volatility = blkimpv(20, 20, 0.09, 4/12, 1.1166, 0.5,
[], true)
%
% return an implied volatility of 0.25, or 25%, per annum.
%
% Notes:
% (1) The input arguments Price, Strike, Rate, Time, Value,
and Class may be
% scalars, vectors, or matrices. If scalars, then that
value is used to
% compute the implied volatility from all options. If more
than one of
% these inputs is a vector or matrix, then the dimensions of
all
% non-scalar inputs must be the same.
% (2) Ensure that Rate and Time are expressed in consistent
units of time.
%
% See also BLKPRICE, BLSPRICE, BLSIMPV.
% Copyright 1995-2003 The MathWorks, Inc.
% $Revision: 1.4.2.2 $ $Date: 2004/01/08 03:06:15 $
% References:
% Hull, J.C., "Options, Futures, and Other Derivatives",
Prentice Hall,
% 5th edition, 2003, pp. 287-288.
% Black, F., "The Pricing of Commodity Contracts," Journal
of Financial
```

```
% Economics, March 3, 1976, pp. 167-79.
%
%
% Implement Black's model for European futures options as
a wrapper
% around a general Black-Scholes option model.
%
% In this context, Black's model is simply a special case of
a
% Black-Scholes model in which the futures/forward con-
tract is
% the underlying asset and the dividend yield = the risk-
free rate.
%
ifnargin< 5
error('Finance:blkimpv:TooFewInputs', ...
'Specify Price, Strike, Rate, Time, and Value.')
end
switchnargin
case 5
[limit, tol, optionClass] = deal([]);
case 6
[limit, tol, optionClass] = deal(varargin{1}, [], []);
case 7
[limit, tol, optionClass] = deal(varargin{1}, varargin
{2}, []);
case 8
[limit, tol, optionClass] = deal(varargin{1:3});
otherwise
error('Finance:blkimpv:TooManyInputs', 'Too many
inputs.')
end
try
volatility = blsimpv(F, X, r, T, value, limit, r, tol,
optionClass);
catch
errorStruct = lasterror;
errorStruct.identifier = strrep(errorStruct.identifier,
'blsimpv', 'blkimpv');
errorStruct.message = strrep(errorStruct.message,
'blsimpv', 'blkimpv');
rethrow(errorStruct);
end
```

References

Ang, J. S., Jou, G. D., & Lai, T. Y. (2009). Alternative formulas to compute implied standard deviation. *Review of Pacific Basin Financial Markets and Policies, 12*, 159–176.

Ang, J. S., Jou, G. D., Lai, T. Y. (2012). A comparison of formulas to compute implied standard deviation. In C. F. Lee, A. C. Lee, J. C. Lee (Eds.), *Encyclopedia of finance* (2nd ed.). Springer, NY. forthcoming.

Beckers, S. (1981). Standard deviation implied in option prices as predictors of future stock price variability. *Journal of Banking and Finance, 5*, 363–381.

Black, F. (1975). Fact and fantasy in the use of options. *Financial Analysts Journal, 31*(4), 36–41. +61–72.

Black, F. (1976). The pricing of commodity contracts. *Journal of Financial Economics, 3*, 167–179.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy, 81*, 637–654.

Brenner, M., & Subrahmanyam, M. G. (1988). A simple formula to compute the implied standard deviation. *Financial Analysts Journal, 44*, 80–83.

Brenner, M., Courtadon, G., & Subrahmanyam, M. (1985). Options on the spot and options on futures. *Journal of Finance, 40*(5), 1303–1317.

Chance, D. M. (1986). Empirical tests of the pricing of index call options. *Advances in Futures and Options Research, 1*, 141–166.

Chance, D. (1996). A generalized simple formula to compute the implied volatility. *The Financial Review, 31*, 859–867.

Corrado, C. J., & Miller, T. W. (1996). A note on a simple, accurate formula to compute implied standard deviations. *Journal of Banking and Finance, 20*, 595–603.

Corrado, C. J., & Miller, T. W. (2004). *Tweaking implied volatility* (Working paper). Auckland: Massey University-Albany.

Garman, M. S., & Klass, M. J. (1980). On the estimation of security price volatility from historical data. *Journal of Business, 53*, 67–78.

Hallerback, W. G. P. M. (2004). *An improved estimator for Black-Scholes-Merton implied volatility. ERIM Report Series*. Erasmus University.

Hull, J. C. (2011). *Options, futures, and other derivatives* (8th ed.). Englewood Cliffs: Prentice-Hall.

Lai, T. Y., Lee, C. F., & Tucker, A. L. (1992). An alternative method for obtaining the implied standard deviation. *Journal of Financial Engineering, 1*, 369–375.

Latane, H. A., & Rendleman, R. J. (1976). Standard deviation of stock price ratios implied by option premia. *Journal of Finance, 31*, 369–382.

Lee, C. F., & Wu, C. (1985). The impacts of kurtosis on risk stability. *Financial Review, 20*, 263–270.

Li, S. (2005). A new formula for computing implied volatility. *Applied Mathematics and Computation, 170*, 611–625.

Macbeth, J. D., & Merville, L. J. (1979). An empirical examination of the Black-Scholes call option pricing model. *Journal of Finance, 34*(5), 1173–1186.

Manaster, S., & Koehler, G. (1982). The calculation of implied variances from the Black–Scholes model: A note. *The Journal of Finance, 37*, 227–230.

Merton, R. C. (1973). The theory of rational option pricing. *Bell Journal of Economics and Management Science, 4*, 141–183.

Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics, 3*(1&2), 125–144.

Park, H. Y., & Sears, R. S. (1985). Changing volatility and the pricing of options on stock index futures. *Journal of Financial Research, 8*, 265–296.

Ramaswamy, K., & Sundaresan, S. M. (1985). The valuation of options on futures contracts. *Journal of Finance, 40*(5), 1319–1340.
Schwert, G. W. (1989). Why does stock market volatility change over times? *Journal of Finance, 44*, 1115–1154.
Whaley, R. (1982). Valuations of American call options on dividend paying stocks: Empirical test. *Journal of Financial Economics, 10*, 29–58.
Wolf, A. (1982). Fundamentals of commodity options on futures. *Journal of Futures Markets, 2*(4), 391–408.

Measuring Credit Risk in a Factor Copula Model

91

Jow-Ran Chang and An-Chi Chen

Contents

Abstract

In this chapter, we provide a new approach to estimate future credit risk on target portfolio based on the framework of CreditMetrics™ by J.P. Morgan. However, we adopt the perspective of factor copula and then bring the principal component analysis concept into factor structure to construct a more appropriate dependence structure among credits.

In order to examine the proposed method, we use real market data instead of virtual one. We also develop a tool for risk analysis which is convenient to use, especially for banking loan businesses. The results show the fact that people assume dependence structures are normally distributed will indeed lead to risk

J.-R. Chang (✉)
National Tsing Hua University, Hsinchu City, Taiwan
e-mail: jrchang@mx.nthu.edu.tw

A.-C. Chen
KGI Securities Co. Ltd., Taipei, Taiwan
e-mail: angel.chen@kgi.com.tw

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_91,

underestimation. On the other hand, our proposed method captures better features of risks and shows the fat-tail effects conspicuously even though assuming the factors are normally distributed.

Keywords
Credit risk • Credit VaR • Default correlation • Copula • Factor copula • Principal component analysis

91.1 Introduction

Credit risk is a risk that generally refers to counterparty failure to fulfill its contractual obligations. The history of financial institutions has shown that many banking association failures were due to credit risk. For the integrity and regularity, financial institutions attempt to quantify credit risk as well as market risk. Credit risk has great influence on all financial institutions as long as they have contractual agreements. The evolution of measuring credit risk has been progressed for a long time. Many credit risk measure models were published, such as CreditMetrics by J.P. Morgan and CreditRisk + by Credit Suisse. On the other side, New Basel Accords (Basel II Accords) which are the recommendation on banking laws and regulations construct a standard to promote greater stability in financial system. Basel II Accords allowed banks to estimate credit risk by using either a standardized model or an internal model approach, based on their own risk management system. The former approach is based on external credit ratings provided by external credit assessment institutions. It describes the weights, which fall into five categories for banks and sovereigns and four categories for corporations. The latter approach allows banks to use their internal estimation of creditworthiness, subject to regulatory. How to build a credit risk measurement model after banking has constructed internal customer credit rating? How to estimate their default probability and default correlations? This thesis attempts to implement a credit risk model tool which links to internal banking database and gives the relevant reports automatically. The developed model facilitates banks to boost their risk management capability.

The dispersion of the credit losses, however, critically depends on the correlations between default events. Several factors such as industry sectors and corporation sizes will affect correlations between every two default events. The CreditMetrics™ model (1997) issued from J.P. Morgan proposed a binomial normal distribution to describe the correlations (dependence structures). In order to describe the dependence structure between two default events in detail, we adopt copula function instead of binomial normal distribution to express the dependence structure.

When estimating credit portfolio losses, both the individual default rates of each firm and joint default probabilities across all firms need to be considered. These features are similar to the valuation process of collateralized debt obligation (CDO). A CDO is a way of creating securities with widely different risk characteristics from a portfolio of debt instrument. The estimating process is almost the same between our goal and CDO pricing. We focus on how to estimate risks. Most CDO pricing literature adopted copula functions to capture the default correlations. Li (2000) extended Sklar's issue (1959) that a copula function can be applied to solve financial problems of default correlation. Li (2000) pointed out that if the dependence structure were assumed to be normally distributed through binomial normal probability density function, the joint transformation probability would be consistent with the result from using a normal copula function. But this assumption is too strong. It has been discovered that most financial data have skew or fat-tail phenomenon. Bouye et al. (2000) and Embrechts et al. (1999) pointed out that the estimating VaR would be underestimated if the dependence structure were described by normal copula comparing to actual data. Hull and White (2004) combined factor analysis and copula functions as a factor copula concept to investigate reasonable spread of CDO. How to find a suitable correlation to describe the dependence structure between every two default events and to speed up the computational complexity is our main object. Bielecki et al. (2012) apply the Markov copula approach to model joint default between counterparty and the reference name in a CDS contract.

This chapter aims to:

1. Construct an efficient model to describe the dependence structure
2. Use this constructed model to analyze overall credit, marginal, and industrial risks
3. Build up an automatic tool for banking system to analyze its internal credit risks

91.2 Methodology

91.2.1 CreditMetrics

Gupton et al. (1997) adopt the main framework of CreditMetrics and calculate credit risks by using real commercial bank loans. The calculating dataset for this chapter is derived from a certain commercial bank in Taiwan. Although there may be some conditions which are different from the situations proposed by CreditMetrics, the calculating process by CreditMetrics can still be appropriately applied to this chapter. For instance, the number of rating degrees in CreditMetrics adopted in S&P's rating category is 7, i.e., AAA to C, but in this loan dataset, there are 9° instead. The following is the introduction to CreditMetrics model framework.

This model can be roughly divided into three components, i.e., value at risk due to credit, exposures, and correlations, respectively, as shown in Fig. 91.1. In this section, these three components and how does this model work out on credit risk valuation will be briefly introduced. Further details could be referred to CreditMetrics technical document.

91.2.1.1 Value at Risk Due to Credit

The process of valuing value at risk due to credit can be decomposed into three steps. For simplicity, we assumed there is only one stand-alone instrument which is

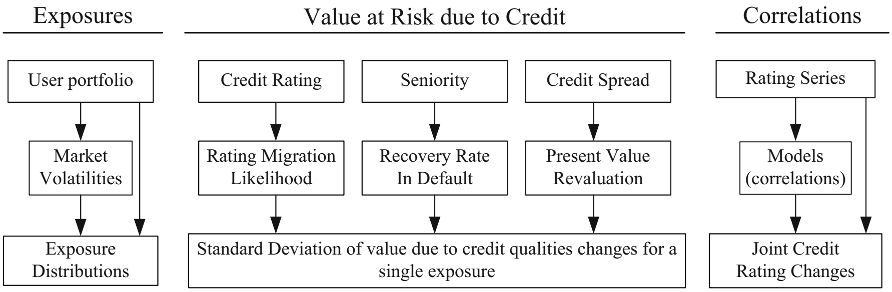


Fig. 91.1 Structure of CreditMetrics model

Table 91.1 One-year transition matrix

	Rating at year-end (%)							
Initial rating	AAA	AA	A	BBB	BB	B	CCC	D
AAA	90.81	8.33	0.68	0.06	0.12	0	0	0
AA	0.70	90.65	7.79	0.64	0.06	0.14	0.02	0
A	0.09	2.27	91.05	5.52	0.74	0.26	0.01	0.06
BBB	0.02	0.33	5.95	86.93	5.30	1.17	0.12	0.18
BB	0.03	0.14	0.67	7.73	80.53	8.84	1.00	1.06
B	0	0.11	0.24	0.91	6.48	83.46	4.07	5.20
CCC	0.22	0	0.22	1.30	2.38	11.24	64.86	19.79

Source: J.P. Morgan's CreditMetrics – technical document (1997)

a corporation bond. (The bond property is similar to loan as they both receive certain amount of cash flow every period and principal at the maturity.) This bond has 5-year maturity and pays an annual coupon at the rate of 5 % to express the calculation process if necessary. Some modifications to fit real situations will be considered later. In Step 1, CreditMetrics assumed that all risks of one portfolio are due to credit rating changes, no matter defaulting or rating migrating. It is significant to estimate not only the likelihood of default but also the chance of migration to move toward any possible credit quality state at the risk horizon. Therefore, a standard system that evaluated "rating changing" under a certain horizon of time is necessary. This information is represented more concisely in transition matrix. Transition matrix can be calculated by observing the historical pattern of rating change and default. They have been published by S&P and Moody's rating agencies or calculated by private banking internal rating systems. Besides, the transition matrix should be estimated for the same time interval (risk horizon) which can be defined by user demand, usually in 1-year period. Table 91.1 is an example to represent 1-year transition matrix.

In the transition matrix table, AAA level is the highest credit rating, and D is the lowest; D also represents that default occurs. According to the above transition matrix table, a company which stays in AA level at the beginning of the year has the probability of 0.64 % to go down to BBB level at the end of the year. By the same

Table 91.2 Recovery rates by seniority class

Class		Recovery rate of Taiwan debt business research using TEJ data	
		Mean (%)	Standard deviation (%)
Loan	Secured	55.38	35.26
	Unsecured	33.27	30.29
Corporation bond	Secured	67.99	26.13
	Unsecured	36.15	37.17

Source: Da-Bai Shen et al. (2003), Research of Taiwan recovery rate with TEJ Data Bank

way, a company which stays in CCC level at the beginning of the year has the probability of 2.38 % to go up to BB level at the end of the year. In this chapter, the transition matrix is to be seen as an external data.[1]

In Step 1, we describe the likelihood of migration to move to any possible quality states (AAA to CCC) at the risk horizon. Step 2 is valuation. The value at the risk horizon must be determined. According to different states, the valuation falls into two categories. First, in the event of a default, recovery rate of different seniority class is needed. Second, in the event of up (down) grades, the change in credit spread that results from the rating migration must be estimated, too.

In default category, Table 91.2 shows the recovery rates by seniority class which this chapter adopts to revaluate instruments. For instance, if the holding bond (5-year maturity that pays an annual coupon at the rate of 5 %) is unsecured and the default occurs, the recovery value will be estimated using its mean value which is 36.15 %.

In rating migration category, the action of revaluation is to determine the cash flows which result from holding the instrument (corporation bond position). Assuming a face value of $100, the bond pays $5 (an annual coupon at the rate of 5 %) each at the end of the next 4 years. Now, the calculating process to describe the value V of the bond assuming the bond upgrades to level A by the formula below:

$$V = 5 + \frac{5}{(1 + 3.72\%)} + \frac{5}{(1 + 4.32\%)^2} + \frac{5}{(1 + 4.93\%)^3} + \frac{105}{(1 + 5.32)^4} = 108.66$$

The discount rate in the above formula comes from the forward zero curves shown in Table 91.3, which is derived from CreditMetrics technical document. This chapter does not focus on how to calculate forward zero curves. It is also seen as an external input data.

Step 3 is to estimate the volatility of value due to credit quality changes for this stand-alone exposure (level A, corporation bond). From step 1 and step 2, the likelihood of all possible outcomes and distribution of values within each outcome are known. CreditMetrics used two measures to calculate the risk estimate:

[1]We do not focus on how to model probability of default (PD) but focus on how to establish the dependence structure. The 1-year transition matrix is a necessary input to our model.

Table 91.3 One-year forward zero curves by credit rating category

Category	Year 1	Year 2	Year 3	Year 4
AAA	3.60	4.17	4.73	5.12
AA	3.65	4.22	4.78	5.17
A	3.72	4.32	4.93	5.32
BBB	4.10	4.67	5.25	5.63
BB	5.55	6.02	6.78	7.27
B	6.05	7.02	8.03	8.52
CCC	15.05	15.02	14.03	13.52

Source: J.P. Morgan's CreditMetrics – technical document (1997)

One is standard deviation, and the other is percentile level. Besides these two measures, this chapter also embraces marginal VaR which denotes the increment VaR due to adding one new instrument in the portfolio.

91.2.1.2 Exposures

As discussed above, the instrument is limited to corporation bonds. CreditMetrics has allowed the following generic exposure types:

1. Non-interest bearing receivables
2. Bonds and loans
3. Commitments to lend
4. Financial letters of credit
5. Market-driven instruments (swap, forwards, etc.)

The exposure type this chapter aims at is loans. The credit risk calculation process of loans is similar to bonds as previous example. The only difference is that loans do not pay coupons. Instead, loans receive interests. But the CreditMetrics model can definitely fit the goal of this chapter to estimate credit risks on banking loan business.

91.2.1.3 Correlations

In most circumstances, there is usually more than one instrument in a target portfolio. Now, multiple exposures are taken into consideration. In order to extend the methodology to a portfolio of multiple exposures, estimating the contribution to risk brought by the effect of nonzero credit quality correlations is necessary. Thus, the estimation of joint likelihood in the credit quality co-movement is the next problem to be resolved. There are many academic papers which address the problems of estimating correlations within a credit portfolio. For example, Gollinger and Morgan (1993) used time series of default likelihood to correlate default likelihood, and Stevenson and Fadil (1995) correlated the default experience across 33 industry groups. On the other hand, CreditMetrics proposed a method to estimate default correlation. They have several assumptions:

(A) A firm's asset value is the process which drives its credit rating changes and default.
(B) The asset returns are normally distributed.
(C) Two asset returns are correlated and bivariate normally distributed, and multiple asset returns are correlated and multivariate normally distributed.

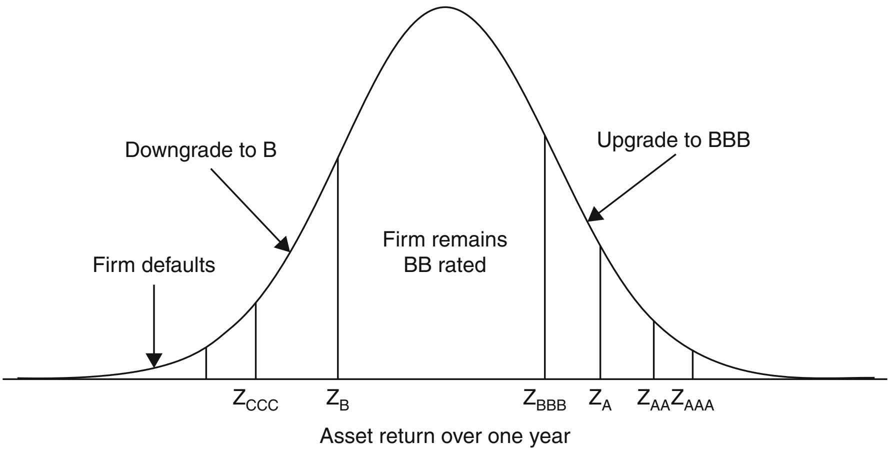


Fig. 91.2 Distribution of asset returns with rating change thresholds

According to assumption A, individual threshold of one firm can be calculated. For a two-exposure portfolio, which credit ratings are level B and level AA and standard deviations of returns are σ and σ′, respectively, it only remains to specify the correlation ρ between two asset returns. The covariance matrix for the bivariate normal distribution is

$$\Sigma = \begin{pmatrix} \sigma^2 & \rho\sigma\sigma' \\ \rho\sigma\sigma' & \sigma'2 \end{pmatrix}$$

Then the joint probability of co-movement that both two firms stay in the same credit rating can be described by the following formula:

$$\Pr\{Z_{BB} < R_1 < Z_B, Z'_{AAA} < R_2 < Z'_{AA}\} = \int_{Z_{BB}}^{Z_B} \int_{Z'_{AAA}}^{Z'_{AA}} f(r, r'; \Sigma)(dr')dr$$

where Z_{BB}, Z_B, Z'_{AAA}, Z'_{AA} are the thresholds. Figure 91.2 gives a concept of the probability calculation. These three assumptions regarding estimating the default correlation are too strong, especially assuming the multiple asset returns are multinormally distributed. In the next section, a better way of using copula to examine the default correlation is proposed.

91.2.2 Copula Function

Consider a portfolio consists of *m* credits. The marginal distribution of each individual credit risks (defaults occur) can be constructed by using either the historical approach or the market implicit approach (derived credit curve from market information). But the question is: how to describe the joint distribution or co-movement between these risks (default correlation)? In a sense, every joint distribution function for a vector of risk factors implicitly contains both a description of the marginal behavior of individual risk factors and a description of their dependence structure. The simplest assumption of dependence structure is mutual independence among the credit risks. However, the independent assumption of the credit risks is obviously not realistic. Undoubtedly, the default rate for a group of credits tends to be higher when the economy is in a recession and lower in a booming. This implies that each credit is subject to the same factors from macroeconomic environment and that there exists some form of dependence among the credits. The copula approach provides a way of isolating the description of the dependence structure. That is, the copula provides a solution to specify a joint distribution of risks, with given marginal distributions. Of course, this problem has no unique solution. There are many different techniques in statistics which can specify a joint distribution with given marginal distributions and a correlation structure. In the following section, the copula function is briefly introduced.

91.2.2.1 Copula Function

An m-dimensional copula is a distribution function on $[0,1]^m$ with standard uniform marginal distributions:

$$C(u) = C(u_1, u_2, \ldots, u_m) \tag{91.1}$$

C is called a copula function.

The copula function C is a mapping of the form C:$[0, 1]^m \rightarrow [0, 1]$, i.e., a mapping of the m-dimensional unit cube $[0, 1]^m$ such that every marginal distribution is uniform on the interval [0, 1]. The following two properties must hold:

1. $C(u_1, u_2, \ldots, u_m, \Sigma)$ is increasing in each component u_i.
2. $C(1, \ldots, 1, u_i, 1, \ldots, 1, \Sigma) = u_i$ for all $i \in \{1, \ldots, m\}$, $u_i \in [0, 1]$.

91.2.2.2 Sklar's Theorem

Sklar (1959) underlined the applications of the copula. Let $F(\bullet)$ be an m-dimensional joint distribution function with marginal distribution $F_1, F_2, \ldots, F_m$. There exist a copula C: $[0, 1]^m \rightarrow [0, 1]$ such that

$$F(x_1, x_2, \ldots, x_m) = C(F_1(x_1), F_2(x_2), \ldots, F_m(x_m)) \tag{91.2}$$

If the margins are continuous, then C is unique.

For any $x_1, \ldots, x_m$ in $\Re = [-\infty, \infty]$ and X has joint distribution function F, then

$$F(x_1, x_2, \ldots, x_m) = \Pr[F_1(X_1) \le F_1(x_1), F_2(X_2) \le F_2(x_2), \ldots, F_m(X_m) \le F_m(x_m)] \tag{91.3}$$

According to Eq. 91.2, the distribution function of $(F_1(X_1), F_2(X_2), \ldots, F_m(X_m))$ is a copula. Let $x_i = F_i^{-1}(u_i)$, then

$$C(u_1, u_2, \ldots, u_m) = F\left(F_1^{-1}(u_1), F_2^{-1}(u_2), \ldots, F_m^{-1}(u_m)\right) \tag{91.4}$$

This gives an explicit representation of C in terms of F and its margins.

91.2.2.3 Copula of F

Li (2000) used the copula function conversely. The copula function links univariate marginals to their full multivariate distribution. For m uniform random variables, U_1, U_2, ..., U_m, the joint distribution function C is defined as

$$C(u_1, u_2, \ldots, u_m, \Sigma) = \Pr[U_1 \leq u_1, U_2 \leq u_2, \ldots, U_m \leq u_m] \tag{91.5}$$

where Σ is correlation matrix of U_1, U_2, ..., U_m.

For given univariate marginal distribution functions $F_1(x_1)$, $F_2(x_2)$, ..., $F_\mathrm{m}(x_\mathrm{m})$. $x_i = F_i^{-1}(u_i)$, the joint distribution function F can be described as follows:

$$F(x_1, x_2, \ldots, x_m) = C(F_1(x_1), F_2(x_2), \ldots, F_m(x_m), \Sigma) \tag{91.6}$$

The joint distribution function F is defined by using a copula.

The property can be easily shown as follows:

$$\begin{aligned}
\mathrm{C}(F_1(x_1), F_2(x_2), \ldots, F_\mathrm{m}(x_\mathrm{m}), \Sigma) &= \Pr[U_1 \leq F_1(x_1), U_2 \leq F_2(x_2), \ldots, U_m \leq F_m(x_m)] \\
&= \Pr\left[F_1^{-1}(U_1) \leq x_1, F_2^{-1}(U_2) \leq x_2, \ldots, F_m^{-1}(U_m) \leq x_m\right] \\
&= \Pr[X_1 \leq x_1, X_2 \leq x_2, \ldots, X_m \leq x_m] \\
&= F(x_1, x_2, \ldots, x_m)
\end{aligned}$$

The marginal distribution of $\mathrm{X_i}$ is

$$\begin{aligned}
&C(F_1(+\infty), F_2(+\infty), \ldots, F_i(x_i), \ldots, F_m(+\infty), \Sigma) \\
&= \Pr[X_1 \leq +\infty, X_2 \leq +\infty, \ldots, X_i \leq x_i, \ldots, X_m \leq +\infty] \\
&= \Pr[X_i \leq x_i] \\
&= F_i(x_i)
\end{aligned} \tag{91.7}$$

Li showed that with given marginal functions, we can construct the joint distribution through some copulas accordingly. But what kind of copula should be chosen corresponding to the realistic joint distribution of a portfolio? For example, the CreditMetrics chose Gaussian copula to construct multivariate distribution.

By Eq. 91.6, this Gaussian copula is given by

$$\begin{aligned}
C^{Ga}(u, \Sigma) &= \Pr(\Phi(X_1) \leq u_1, \Phi(X_2) \leq u_2, \ldots, \Phi(X_m) \leq u_m, \Sigma) \\
&= \Phi_\Sigma\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2), \ldots, \Phi^{-1}(u_m)\right)
\end{aligned} \tag{91.8}$$

where Φ denotes the standard univariate normal distribution, Φ^{-1} denotes the inverse of a univariate normal distribution, and Φ_Σ denotes multivariate normal distribution. In order to easily describe the construction process, we only discuss two random variables u_1 and u_2 to demonstrate the Gaussian copula:

$$C^{Ga}(u_1, u_2, \rho) = \int_{-\infty}^{\Phi^{-1}(u_1)} \int_{-\infty}^{\Phi^{-1}(u_2)} \frac{1}{2\pi\sqrt{(1-\rho^2)}} \exp\left\{-\frac{v_1^2 - 2\rho v_1 v_2 + v_2^2}{2(1-\rho^2)}\right\} dv_2 dv_1 \tag{91.9}$$

where ρ denotes the correlation of u_1 and u_2.

Equation 91.9 is also equivalent to the bivariate normal copula which can be written as follows:

$$C(u_1, u_2, \rho) = \Phi_2\left(\Phi^{-1}(u_1), \Phi^{-1}(u_2)\right) \tag{91.10}$$

Thus, with given individual distribution (e.g., credit migration over 1-year horizon) of each credit asset within a portfolio, we can obtain the joint distribution and default correlation of this portfolio through copula function. In our methodology, we do not use copula function directly. In the next section, we bring in the concept of factor copula for further improvement to form the default correlation. Using factor copula has two advantages. One is to avoid constructing a high-dimensional correlation matrix. If there are more and more instruments ($N > 1{,}000$) in our portfolio, we need to store N-by-N correlation matrix; scalability is one problem. The other advantage is to speed up the computation time because of the lower dimension.

91.2.3 Factor Copula Model

In this section, copula models that have a factor structure will be introduced. It is called factor copula because this model describes dependence structure between random variables not from the perspective of a certain copula form, such as Gaussian copula, but from the factors model. Factor copula models have been broadly used to assess price of collateralized debt obligation (CDO) and credit default swap (CDS). The main concept of factor copula model is that under a certain macro environment, credit default events are independent to each other. And the main causes that affect default events come from potential market economic conditions. This model provides another way to avoid dealing with multivariate normal distribution (high-dimensional) simulation problem.

Continuing the above example, a portfolio is consisted of m credits. In the first we consider the simplest example which contains only one factor and define V_i as the asset value of ith credit under single factor copula model. Then this ith credit asset value can be expressed by one factor M (mutual factor) chosen from macroeconomic factors and one error term ε_i:

$$V_i = r_i M + \sqrt{1 - r_i^2}\varepsilon_i \tag{91.11}$$

where r_{i} is weight of M, and the mutual factor M is independent of ε_i.

Let the marginal distribution of $V_1, V_2, \ldots, V_{\mathrm{m}}$ are F_{i}, $\mathrm{i} = 1, 2, \ldots, m$. Then the m-dimensional copula function can be written as

$$\begin{aligned} C(u_1, u_2, \ldots, u_m) &= F\left(F_1^{-1}(u_1), F_2^{-1}(u_2), \ldots, F_m^{-1}(u_m)\right) \\ &= \Pr\left(V_1 \le F_1^{-1}(u_1), V_2 \le F_2^{-1}(u_2), \ldots, V_m \le F_m^{-1}(u_m)\right) \end{aligned} \tag{91.12}$$

where F is the joint cumulative distribution function of $V_1, V_2, \ldots, V_{\mathrm{m}}$.

It has been known that M and ε_i are independent of each other, according to iterated expectation theorem; Eq. 91.12 can be written as

$$\begin{aligned} C(u_1, u_2, \ldots, u_m) &= E\left\{\Pr\left(V_1 \le F_1^{-1}(u_1), V_2 \le F_2^{-1}(u_2), \ldots, V_m \le F_m^{-1}(u_n)\right)|M\right\} \\ &= E\left\{\prod_{i=1}^{m} \Pr\left(r_i M + \sqrt{1 - r_i^2}\varepsilon_i \le F_i^{-1}(u_i)\right)|M\right\} \\ &= E\left\{\prod_{i=1}^{m} F_{\varepsilon, i}\left(\frac{F_i^{-1}(u_i) - r_i M}{\sqrt{1 - r_i^2}}\right)|M\right\} \\ &= \int\left(\prod_{i=1}^{m} F_{\varepsilon, i}\left(\frac{F_i^{-1}(u_i) - r_i M}{\sqrt{1 - r_i^2}}\right)\right) g(M) dM \end{aligned} \tag{91.13}$$

Using the above formula, the m-dimensional copula function can be derived. Moreover, according to Eq. 91.13, the joint cumulative distribution F can also be derived:

$$F(t_1, t_2, \ldots, t_m) = \int\left(\prod_{i=1}^{m} F_{\varepsilon, i}\left(\frac{\left(F_i^{-1}\left(F_{T,i}(t_i)\right) - r_i M\right)}{\sqrt{1 - r_i^2}}\right)\right) g(M) dM \tag{91.14}$$

Let $F_i(t_i) = \Pr(T_i \le t_i)$ represents i credit default probability (default occurs before time $\mathrm{t_i}$), where F_i is the marginal cumulative distribution. We note here that CDX pricing cares about when the default time $\mathrm{T_i}$ occurs. Under the same environment (systematic factor M) (Andersen and Sidenius (2004)), the default probability $\Pr(T_i \le t_i)$ will be equal to $\Pr(V_i \le c_i)$, which represent that the probability asset value V_{i} is below its threshold c_{i}. Then joint default probability of these m credits can be described as follows:

$$F(c_1, c_2, \ldots, c_m) = \Pr(V_1 \le c_1, V_2 \le c_2, \ldots, V_m \le c_m)$$

Now, we bring the concept of principal component analysis (PCA). People use PCA to reduce the high-dimensional or multivariable problems. If someone would like to explain one thing (or some movement of random variables), he/she has to

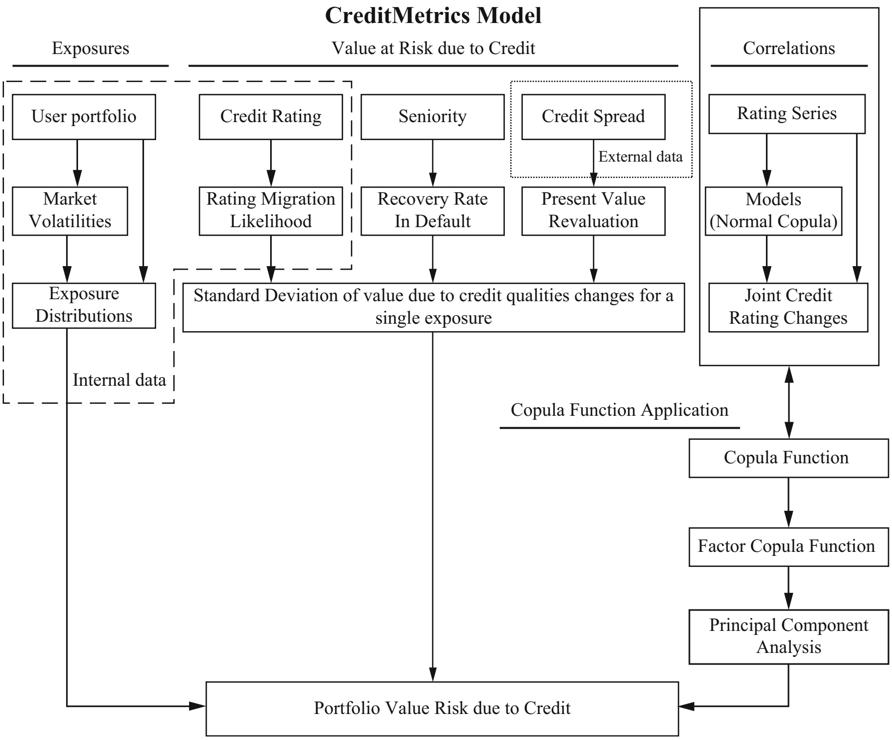


Fig. 91.3 Architecture of proposed model

gather interpreting variables related to those variable movements or their correlation. Once the kinds of interpreting variables are too huge or complicated, it becomes harder to explain those random variables and will produce complex problems. Principal component analysis provides a way to extract approximate interpreting variables to cover maximum variance of variables. Those representative variables may not be "real" variables. Virtual variables are allowed and depend on the explaining meaning. We do not talk about PCA calculation processes; further detail could be referred to Jorion (2000). Based on factor model, the asset value of m credits with covariance matrix Σ can be described as follows:

$$V_i = r_{i1}y_1 + r_{i2}y_2 + \ldots + r_{im}y_m + \varepsilon_i \tag{91.15}$$

where y_i are common factors between these m credits and r_{ij} is the weight (factor loading) of each factor. The factors are independent of each other. The question is: how to determinate those y_i factors and their loading? We use PCA to derive the factor loading. Factor loadings are based on listed price of those companies in the portfolio to calculate their dependence structure. The experimental results will be shown in the next section. We note here that the dependence structure among assets have been absorbed into factor loadings (Fig. 91.3).

91.3 Experimental Results

The purpose of this chapter is to estimate credit risk by using principal component analysis to construct dependence structure without giving any assumptions to specify formulas of copula. In other words, the data were based on itself to describe the dependence structure.

91.3.1 Data

In order to analyze credit VaR empirically through proposed method, this investigation adopts the internal loan account data, loan application data, and customer information data from a commercial bank on current market in Taiwan. For reliability of data authenticities, all the data are in Taiwan stock market instead of virtual one. This also means now the portfolio pool contains only the loans of listed companies and does not contain the loan of unlisted companies. According to the period of these data, we can estimate two future portfolio values. They are values on 2003 and 2004, respectively.

All requirement data are downloaded automatically from database system to workspace for computations. Before going to the detail of the experiments, the relevant data and experimental environment are introduced as follows.

91.3.1.1 Requirements of Data Input

1. Commercial bank internal data: This internal data contains nearly 40,000 entries of customer's data, 50,000 entries of loan data, and 3,000 entries of application data. These data contain maturity dates, outstanding amount, credit ratings, interest rate for lending, market type, etc. up to December 31, 2004.
2. One-year period transition matrix: The data was extracted from Yang (2005), who used the same commercial bank history data to estimate a transition matrix which obeyed Markov chain (Table 91.4).
3. Zero forward rate: Refer to Yang (2005), based on computed transition matrix to estimate the term structure of credit spreads. Furthermore, they added corresponding risk-free interest rate to calculate zero forward rates from discounting zero spot rates (Table 91.5).
4. Listed share prices at exchange market and over-the-counter market: We collected weekly listed share prices of all companies at exchange and over-the-counter markets in Taiwan from January 1, 2000, to December 31, 2003, in total 3 years' data, through Taiwan Economic Journal Data Bank (TEJ).

91.3.2 Simulation

In this section, the simulation procedure of analyzing banking VaR is briefly introduced. There are two main methods of experiments: A and B. A is the method that this chapter proposed which uses factor analysis to explain the dependence

Table 91.4 One-year transition matrix (commercial data)

	Rating at year-end (%)									
Initial rating	1	2	3	4	5	6	7	8	9	D
1	100	0	0	0	0	0	0	0	0	0
2	3.53	70.81	8.90	5.75	6.29	1.39	0.19	2.74	0.06	0.34
3	10.76	0.03	72.24	0.24	10.39	5.78	0.31	0.09	0.06	0.10
4	1.80	1.36	5.85	57.13	18.75	11.31	2.45	0.32	0.70	0.33
5	0.14	0.44	1.58	2.39	75.47	16.97	1.49	0.61	0.49	0.42
6	0.09	0.06	0.94	2.44	13.66	70.58	6.95	1.68	0.76	2.81
7	0.05	0.05	0.27	3.72	3.75	14.49	66.39	8.05	0.12	3.11
8	0.01	0	0.03	0.45	0.21	1.34	2.00	77.10	0.44	18.42
9	0	0	0.02	0.09	1.46	1.80	1.36	3.08	70.06	22.13
D	0	0	0	0	0	0	0	0	0	100

Source: Yang (2005)

Table 91.5 One-year forward zero curves by credit rating category (commercial data)

Yield (%)		1 year	2 year	3 year	4 year	5 year	6 year	7 year	8 year	9 year
Credit rating	1	1.69	2.08	2.15	2.25	2.41	2.53	2.58	2.62	2.7
	2	2.57	2.88	3.19	3.44	3.72	3.94	4.07	4.18	4.33
	3	2.02	2.41	2.63	2.85	3.11	3.32	3.45	3.56	3.71
	4	2.6	2.93	3.28	3.59	3.91	4.17	4.34	4.48	4.65
	5	2.79	3.1	3.48	3.81	4.15	4.42	4.6	4.75	4.93
	6	4.61	5.02	5.16	5.31	5.51	5.67	5.76	5.83	5.93
	7	6.03	6.16	6.56	6.83	7.07	7.23	7.28	7.31	7.36
	8	22.92	23.27	22.54	21.91	21.36	20.78	20.15	19.52	18.94
	9	27.51	27.82	26.4	25.17	24.09	23.03	21.97	20.97	20.08

Source: Yang (2005)

structure and to simulate the distribution of future values. B is the contrast set which are used traditionally and popularly in most applications such as CreditMetrics. We call it the multi-normal (normal/Gaussian copula) simulation method.

Both of these two methods need three input data tables: credit transition matrix, forward zero curves, and share prices of each corporation in the portfolio pool. The detail of normal copula method procedure is not mentioned here; readers can refer to technical documentation of CreditMetrics. Now, the process of factor analysis method is shown as follows:

1. Extract the data entries that do not mature under given date, from database system including credit ratings, outstanding amounts, and interest rates.
2. According to the input transition matrix, we can calculate standardized thresholds for each credit rating.
3. Use the share prices of those corporations in the portfolio pool to calculate equities correlations.

4. Use principal component analysis to obtain each factor loadings for all factors under the assumption that these factors obey some distributions (e.g., standard normal distribution) to simulate their future asset value and future possible credit ratings.
5. According to possible credit ratings, discount the outstanding amounts by their own forward zero curves to evaluate future value distributions.
6. Display the analysis results.

91.3.3 Discussion

91.3.3.1 Tools and Interfaces Preview

For facility and convenience, this chapter uses MATLAB and MySQL to construct an application tool to help tool users analyze future portfolio value more efficiently. Following is this tool's interactive interfaces:

Basic Information of Experimental Data: (Pie Chart)

The non-computing data and pie charts give user the basic view of loan information. These charts present the proportion of each composition of three key elements: loan amount of companies, industry, and credit rating. To assist user, construct an overview of concerned portfolio (Fig. 91.4).

Pie chart of loan amount weight in terms of enterprise, industry, and credit rating.

Information According to Experimental data: (Statistic Numbers)

Besides graphic charts, the second part demonstrates a numerical analysis. The first part is the extraction of the company data which has maturity more than the given months, and the second part is the extraction of the essential data of top weighted companies. Parts I and II extract data without any computation; the only thing has been done is to sort or remove some useless data (Figs. 91.5 and 91.6).

Set Criteria and Derive Fundamental Experimental Result

This portion is the core of proposed tool; it provides several functions of computations. Here are the parameters that users must decide themselves:

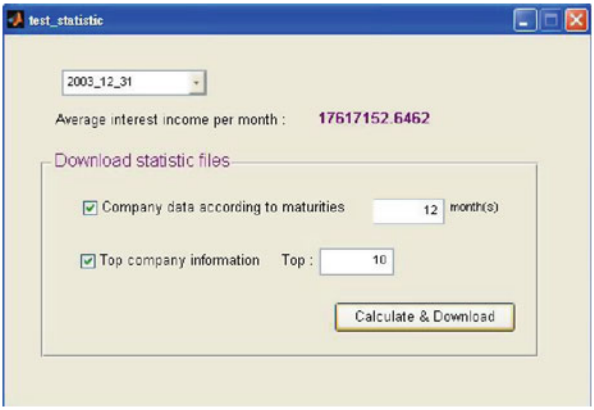
1. Estimated year.
2. Confidence level.
3. Simulation times. Of course, the more simulation time user chooses, the more computational time will need.
4. Percentage of explained factors which is defined for PCA method. Using the eigenvalues of given normalized assets (equities) values, we can determinate the explained percentage.
5. This function gives user the option to estimate all or portion of the companies of portfolio pool. The portion part is sorted according to the loan amount of each enterprise. User can choose multiple companies they are most concerned. The computational result is written to a text file for further analysis.

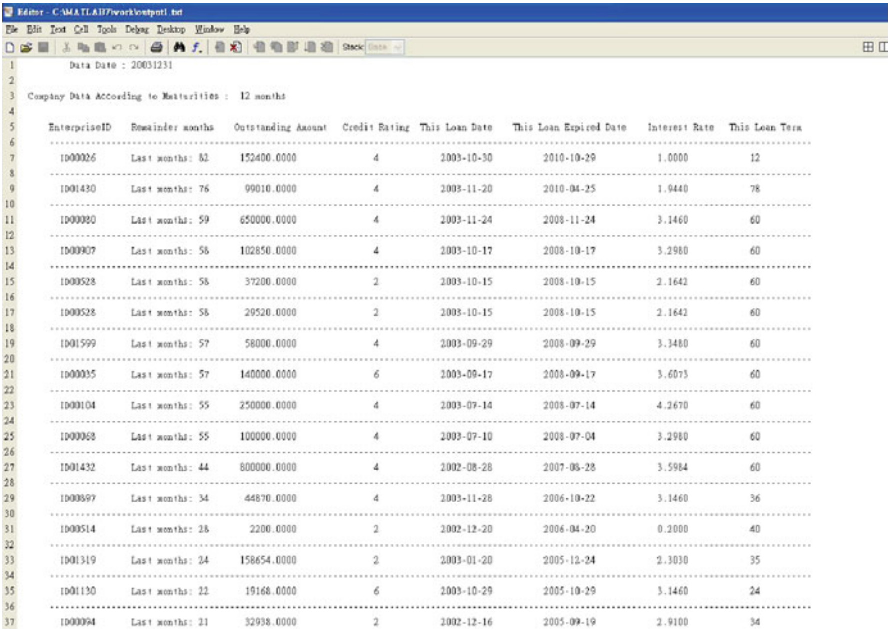
At the end of 2003

At the end of 2004

Fig. 91.4 Interface of part I

Fig. 91.5 Interface of part II




Editor - C:\MATLAB7\work\output1.txt

Data Date : 20031231

Company Data According to Maturities : 12 months

EnterpriseID	Remainder months	Outstanding Amount	Credit Rating	This Loan Date	This Loan Expired Date	Interest Rate	This Loan Term
ID00026	Last months: 82	152400.0000	4	2003-10-30	2010-10-29	1.0000	12
ID01430	Last months: 76	99010.0000	4	2003-11-20	2010-04-25	1.9440	78
ID00080	Last months: 59	650000.0000	4	2003-11-24	2008-11-24	3.1460	60
ID00907	Last months: 58	102850.0000	4	2003-10-17	2008-10-17	3.2980	60
ID00528	Last months: 58	37200.0000	2	2003-10-15	2008-10-15	2.1642	60
ID00528	Last months: 58	29520.0000	2	2003-10-15	2008-10-15	2.1642	60
ID01599	Last months: 57	58000.0000	4	2003-09-29	2008-09-29	3.3480	60
ID00035	Last months: 57	140000.0000	6	2003-09-17	2008-09-17	3.6073	60
ID00104	Last months: 55	250000.0000	4	2003-07-14	2008-07-14	4.2670	60
ID00068	Last months: 55	100000.0000	4	2003-07-10	2008-07-04	3.2980	60
ID01432	Last months: 44	800000.0000	4	2002-08-28	2007-08-28	3.5984	60
ID00897	Last months: 34	44870.0000	4	2003-11-28	2006-10-22	3.1460	36
ID00514	Last months: 28	2200.0000	2	2002-12-20	2006-04-20	0.2000	40
ID01319	Last months: 24	158654.0000	2	2003-01-20	2005-12-24	2.3030	35
ID01130	Last months: 22	19168.0000	6	2003-10-29	2005-10-29	3.1460	24
ID00094	Last months: 21	32938.0000	2	2002-12-16	2005-09-19	2.9100	34

Fig. 91.6 Companies data downloads from part II interface

6. Distribution of factors. This is defined for PCA method, too. There are two distributions that user can choose: standard normal distribution or Student's t-distribution. The default freedom of Student's t-distribution is set as one (Fig. 91.7).

Report of Overall VaR Contributor

User may be more interested in the detail of risk profile at various levels. In this part, industries are discriminated from 19 sections, and credits are discriminated from nine levels. This allow user to see where the risk is concentrated visually (Fig. 91.8).

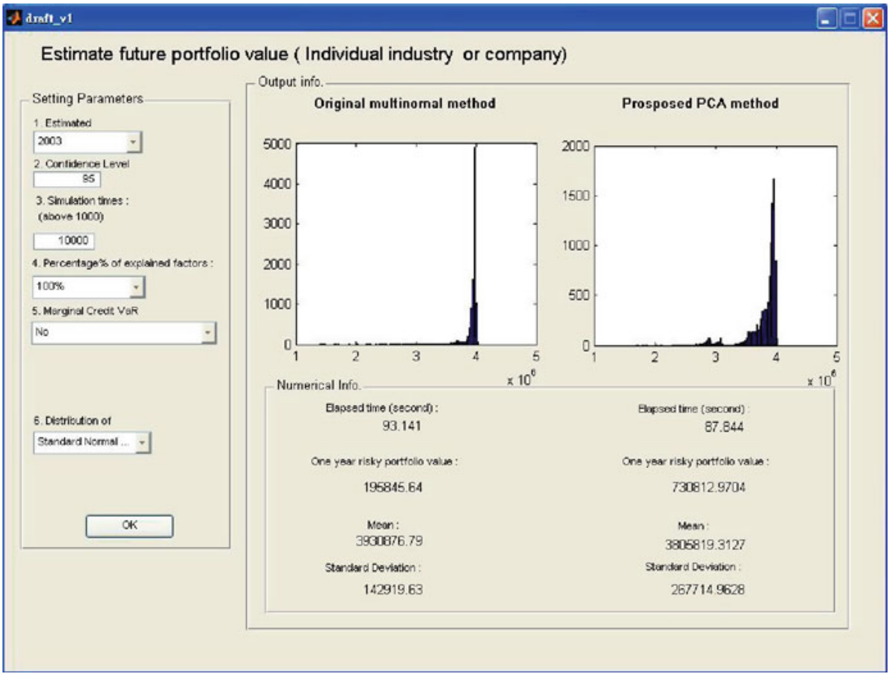


Fig. 91.7 Interface of part III

91.3.3.2 Experimental Result and Discussion

Table 91.6 represents the experimental results. For objectivity, all simulation times are set to 100,000 times which is large enough to obtain stable numerical results.[2] Based on the provided data, the 1-year future portfolio value of listed corporations on 2003 and 2004 can be estimated. In other words, standing on January 1, 2002, we can estimate the portfolio value on December 31, 2003. Or standing on January 1, 2003, we can estimate the portfolio value on December 31, 2004. The following tables listed the experimental results of factor copula methods of different factor distributions and compared with multi-normal method by CreditMetrics. The head of the tables are parameter setting, and the remained fields are experimental results. We note here the formula Eq. 91.15

$$V_i = r_{i1}y_1 + r_{i2}y_2 + \ldots + r_{im}y_m + \varepsilon_i$$

where the distribution of factors $y_1, y_2 \ldots y_m$ listed in the following table is standard normally distributed and Student t-distributed (assumes freedoms are 2, 5, and 10).

[2]We have examine the simulation times; 100,000 times is enough to have a stable computational result.

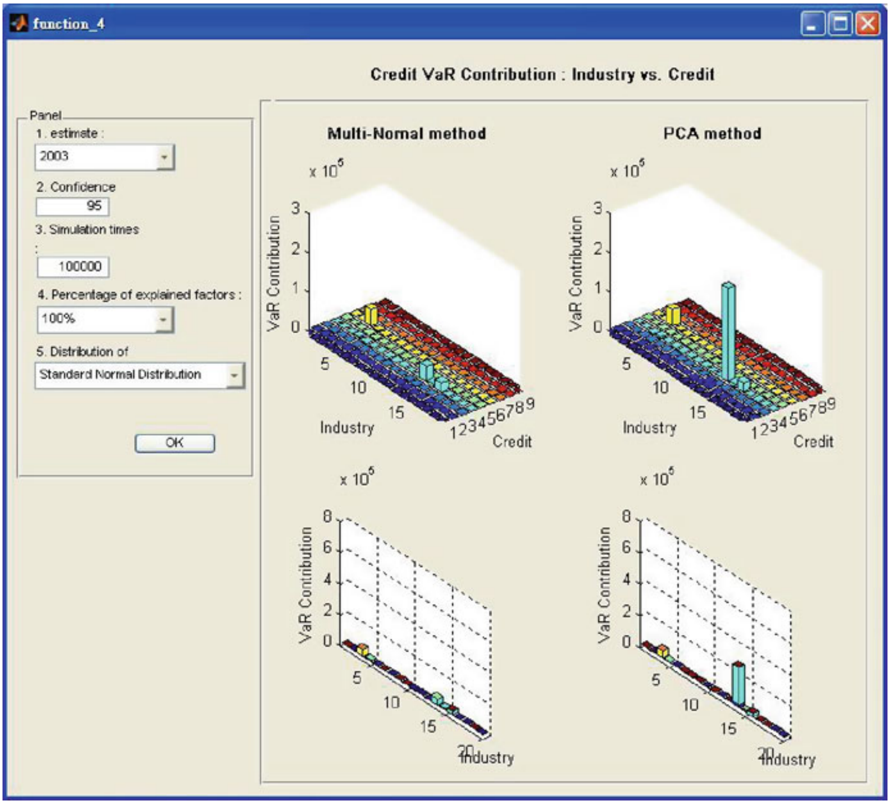


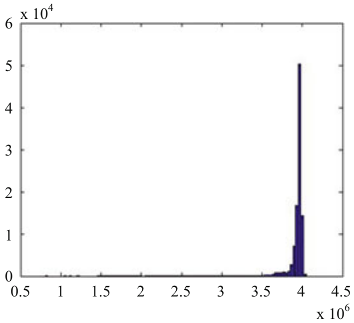
Fig. 91.8 VaR contribution of individual credits and industries

There are some messages that can be derived from the above table. First, obviously, risk of future portfolio value by multi-normal method is less than by proposed method. The risk amount of proposed method is 3–5 times over multi-normal method. This result corresponds to most research that copula function can capture the fat-tail phenomenon which prevails over practical market more adequately. Second, the distribution of future portfolio value by proposed method is more diversified than multi-normal method which concentrated on nearly 400,000 with 50,000 times while proposed method with 17,000 times. Third, it is very clear to see that risks with factors using Student's t-distribution to simulate are more than with normal distribution, and the risk amount tends toward the same while the degree of freedom becomes larger. Fourth, the mean of portfolio of proposed method is smaller than that of multi-normal method, but the standard deviation of proposed method is much more than multi-normal method. It shows that the overall possible portfolio values by proposed method have the trend to become less worth and also fluctuate more rapidly.

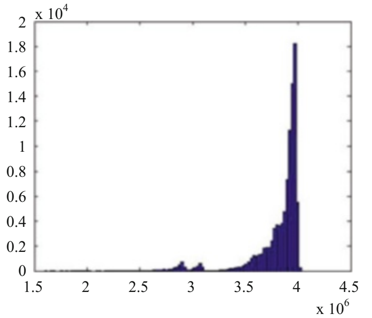
The above discrepancies between two methods give us some inferences. First, the proposed method provides another way to estimate more actual credit risks of

Table 91.6 Experimental result of estimated portfolio value at the end of 2003

Estimate year: 2003				
Parameter setting				
Simulation time: **100,000**			Percentage of explained factors: **100.00 %**	
Involved listed enterprise number: **40**			Loan account entries: **119**	
Result				
Factor distribution assumption : normal distribution $F \sim N(0,1)$				
	Credit VaR 95 %	Credit VaR 99 %	Portfolio mean	Portfolio s.d.
Multi-normal	192,113.4991	641,022.0124	3,931,003.1086	136,821.3770
PCA	726,778.6308	1,029,766.9285	3,812,565.6170	258,628.5713

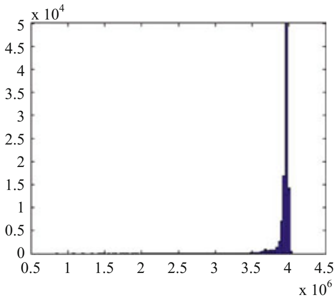


Multi-Normal method

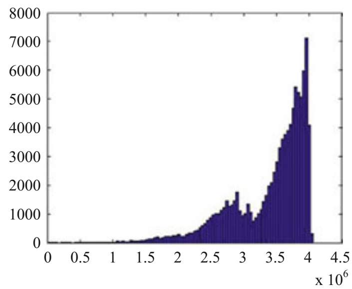


PCA method

Factor distribution assumption : Student's t-distribution, freedom = (2)				
	Credit VaR 95 %	Credit VaR 99 %	Portfolio mean	Portfolio s.d.
Multi-normal	191,838.2019	620,603.6273	3,930,460.5935	136,405.9177
PCA	1,134,175.1655	1,825,884.8901	3,398,906.5097	579,328.2159



Multi-Normal method



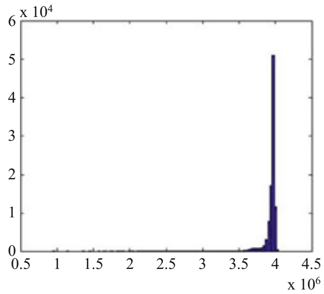
PCA method

Factor distribution assumption : Student's t-distribution, freedom = (5)				
	Credit VaR 95 %	Credit VaR 99 %	Portfolio mean	Portfolio s.d.
Multi-normal	192,758.7482	610,618.5048	3,930,923.6708	135,089.0618
PCA	839,129.6162	1,171,057.2562	3,728,010.5847	337,913.7886

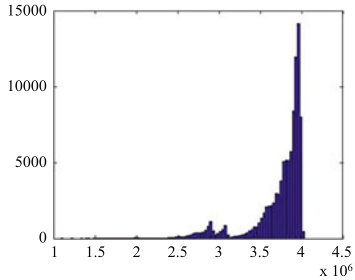
(*continued*)

Table 91.6 (continued)

Factor distribution assumption : Student's t-distribution, freedom = (5)				
	Credit VaR 95 %	Credit VaR 99 %	Portfolio mean	Portfolio s.d.

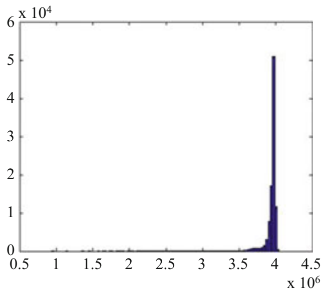


Multi-Normal method

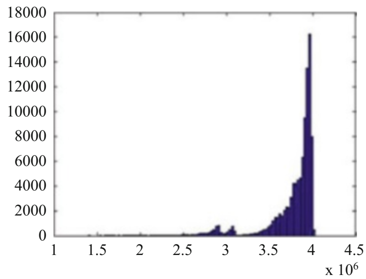


PCA method

Factor distribution assumption : Student's t-distribution, freedom = (10)				
	Credit VaR 95 %	Credit VaR 99 %	Portfolio mean	Portfolio s.d.
Multi-normal	192,899.0228	600,121.1074	3,930,525.7612	137,470.3856
PCA	773,811.8411	1,080,769.3036	3,779,346.2750	291,769.4291



Multi-Normal method



PCA method

Table 91.7 CPU time for factor computation (simulation time: 100,000 year: 2003)

	Explained ratio (s)				
Method	100 %	90~95 %	80~85 %	70~80 %	Below 60 %
Multi-normal	2.5470	2.6090	2.2350	2.2500	2.3444
PCA	1.2030	0.7810	0.7030	0.6720	0.6090

portfolio containing risky credits through market data, and this method captures fat-tail event more notably. Second, the computation time of proposed method is shorter than multi-normal method. In Table 91.7, when using fully explained factors, computation time by proposed method is still faster than by multi-normal method. The computation time decreases as the required explained ratio is set lower.

Table 91.8 Individual credit VaR of top 5 industries

	Credit VaR	
Industry	Multi-normal method	PCA method
(No. 1) Electronics	40,341	252,980
(No. 2) Plastic	42,259	42,049
(No. 3) Transportation	22,752	22,391
(No. 4) Construction	7,011	7,007
(No. 5) Textile	2,884	2,765

Table 91.9 Estimate portfolio value at the end of 2004 with different explained level

95 % confidence level, F ~ (0,1)					
	100 %	90~95 %	80~85 %	70~80 %	60~70 %
Multi-normal	208,329.40	208,684.38	209,079.72	208,686.22	207,710.63
PCA	699,892.33	237,612.60	200,057.74	187,717.73	183,894.91

This means less numbers of factors are used for the expected explained level. Third, Table 91.8 which retrieves individual credit VaR contribution to whole portfolio from 19 industries shows that the main risk comes from electronics industry. Based on the commercial data, we find out that among its loan account entries, the electronics industry customers have the proportion of exceeding half of loan entries (63/119). The credit VaR of electronics industry computed by proposed method is six times more than by multi-normal method. This effect reveals that the multi-normal method lacks the ability to catch concentrative risks. On the contrary, based on factor structure, the mutual factor loadings extracted by the correlation among companies express more actual risks. Fourth, for finite degree of freedom, the t-distribution has fatter tails than Gaussian distribution and is known to generate tail dependence in the joint distribution.

Table 91.9 shows the impact on risk amount by using different factor numbers. According to Table 91.9, the risks decrease as the explained level decreases; this is a trade-off between time-consuming and afforded risk amount. Most research and reports say 80 % explained level is large enough to be accepted.

91.4 Conclusion

Credit risk and default correlation issues have been probed in recent research, and many solutions have been proposed. We take another view to examine credit risks and derivative tasks. On our perspective, the loan credits in target portfolio like the widely different risk characteristics from a portfolio of debt instruments and their properties and behavior are the same in the main.

In this chapter, we propose a new approach which connects the principal component analysis and copula functions to estimate credit risks of bank loan businesses. The advantage of this approach is that we do not need to specify particular copula functions to describe dependence structure among credits. On the contrary, we use a factor structure which covers market factor and idiosyncratic factor, and the computed risks have heavy-tail phenomenon. Another benefit is that it reduces the difficulties to estimate parameters which copula functions will use. This approach provides another way and has better performance than conventional method such as assume the dependence structures are normally distributed.

In order to describe the risk features and other messages that bank policymakers may like to know, we wrote a tool for risk estimation and results display. It contains basic data information preview which just downloads data from database and does some statistic analyses. It also provides different parameter settings and uses Monte Carlo simulation to calculate credit VaR and finally gives an overview of individual credit VaR contributions. The experimental results are consistent with previous studies that the risk will be underestimated compared with real risks if people assume dependence structure are normally distributed. In addition, the aforementioned approach and tool still have some rooms to be improved such as recovery rate estimations, how to chose distributions of factors, and more friendly user interface.

References

Andersen, L., & Sidenius, J. (2004). Extensions to the Gaussian copula: Random recovery and random factor loadings. *Journal of Credit Risk, 1*(1), 29–70.

Bielecki, T. R., Crepey, S., Jeanblanc, M., & Zargari, B. (2012). Valuation and hedging of CDS counterparty exposure in a Markov copula model. *International Journal of Theoretical and Applied Finance, 15*(1), 85–123.

Bouye, E. V., Durrleman, A., Nikeghbali, G. R., & Roncalli, T. (2000). *Copulas for finance – A reading guide and some application*. Groupe de Recherche.

Embrechts, P., McNeil, A., & Straumann, D. (1999). *Correlation and dependence in risk management: Properties and pitfalls*. Zurich: Mimeo. ETHZ Zentrum.

Gollinger, T. L., & Morga, J. B. (1993). Calculation of an efficient frontier for a commercial loan portfolio. *Journal of Portfolio Management, 19*, 39–46.

Gupton, G. M., Finger, C. C., & Bhatia, M. (1997). *CreditMetrics –technical document*. New York: Morgan Guaranty Trust Company.

Hull, J., & White, A. (2004). Valuing of a CDO and an n-th to default CDS without Monte Carlo simulation. *Journal of Derivatives, 12*(2), 8–48.

Jorion, P. (2000). *Value at risk*. New York: McGraw Hill.

Li, D. X. (2000). On default correlation: A copula function approach. *Journal of Fixed Income, 9*, 91–54.

Shen, D. B., Tsai, C. C. & Jing, Y. K. (2003). Research of Taiwan recovery rate with TEJ Data Bank. *Accounting Research Monthly, 215,* 113–125.

Sklar, A. (1959). Functions de repartition an n dimension et leurs marges. *Publications de l'Institut de Statistique de L'Université de Paris*, *8*, 229–231.

Stevenson, B. G., & Fadil, M. W. (1995). Modern portfolio theory: Can it work for commercial loans? *Commercial Lending Review, 10*(2), 4–12.

Yang, T. C. (2005). *The pricing of credit risk of portfolio base on listed corporation in Taiwan market*. Master Thesis, National Tsing Hua University, Taiwan.

Instantaneous Volatility Estimation by Nonparametric Fourier Transform Methods

92

Chuan-Hsiang Han

Contents

Abstract

Malliavin and Mancino (2009) proposed a nonparametric Fourier transform method to estimate the instantaneous volatility under the assumption that the underlying asset price process is a semi-martingale. Based on this theoretical result, this chapter first conducts some simulation tests to justify the effective-ness of the Fourier transform method. Two correction schemes are proposed to improve the accuracy of volatility estimation. By means of these Fourier transform methods, some documented phenomena such as volatility daily

Work supported by NSC 101-2115-M-007-011, Taiwan.

C.-H. Han
Department of Quantitative Finance, National Tsing Hua University, Hsinchu, Taiwan, Republic of China
e-mail: chhan@mx.nthu.edu.tw

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_92,

effect and multiple risk factors of volatility can be observed. Then, a linear hypothesis between the instantaneous volatility and VIX derived from Zhang and Zhu (2006) is investigated. We extend their result and adopt a general linear test for empirical analysis.

Keywords
Information content • Instantaneous volatility • Fourier transform method • Bias reduction • Correction method • Local volatility • Stochastic volatility • VIX • Volatility daily effect • Online estimation

92.1 Introduction

Volatility estimation has been recognized as a core problem along the development of modern financial markets. There are enormous literatures devoted to this subject. The concept of "information content" is introduced to categorize a huge amount of studies. More precisely, backward and forward information contents of volatility are used and distinguished by time.

For any given time point, the backward information refers to the usage of a set of past information. A segment of historical data of an underlying risky asset, such as historical stock prices, is a typical example. In contrast, the forward information refers to the usage of financial data that contain risk exposure in the future. Financial derivatives such as futures and options are typical examples.

The backward information content of volatility is extensively investigated in the fields of financial statistics and econometrics. See Tsai (2005), Engle (2009), and references therein. The forward information content of volatility is almost exclusively studied in the field of mathematical finance and financial engineering. Within all these academic fields, parametric models play the key role to analyze financial data such as stocks and options because certain mathematical structures allow for analytic or computational assessments to relevant estimation procedures.

Relatively few results on nonparametric models can be found to analyze volatility. Dupire formula and VIX Gatheral (2006) are frontiers to compute some kinds of volatility using traded option data. Of course, these volatilities contain forward information. In the context of backward information, the dual concept of VIX, i.e., the integrated variance, is the historical volatility squared, which is defined as the variance of the standardized returns and easy to calculate. The instantaneous volatility, denoted by σ_t, can be viewed as the dual concept of Dupire formula, in which the volatility $\sigma(T, K)$ depends on the maturity T and the strike price K. Unfortunately, estimation of the instantaneous volatility is hard.

One accessible way to estimate the instantaneous volatility is taken directly from the result of a differentiation on the quadratic variation $\langle \cdot, \cdot \rangle_t$ of the underlying price process S. That is, for small $\Delta > 0$,

$$\sigma_t \approx \frac{\langle S, S\rangle_{t+\Delta} - \langle S, S\rangle_t}{\Delta}$$

This approximation is theoretically consistent but not plausible for practical implementation. One main reason is that the differentiation is sensitive to data frequency as seen from the above approximation. See Zhang and Mykland (2005) for improved methods on this direction.

Not until recently, Malliavin and Mancino (2009) proposed a Fourier transform method for estimation of the instantaneous volatility. This alternative approach is integral based, not differentiable based, as the aforementioned difference approximation. The authors claim that this approach is particularly suitable for the analysis of high-frequency time series and for the computation of cross volatilities.

However, Reno (2008) alerts that the Fourier algorithm performs badly near boundaries of estimated volatility time series data, i.e., estimated volatility of the first and last 1 % time series is not accurate enough. The author recommends discarding those volatility estimates near the boundary. Yet, this compromise may constitute a drawback in estimation. One example is that, when exclusively following Reno (2008), dropping the most recent 1 % volatility estimates will distort the prediction of a short-time volatility, say 1-day volatility.

To avoid this "boundary effect" pitfall, price correction schemes by matching the estimated volatility with observed price returns have been proposed in Han et al. (2014) and Han (2014). These schemes only require solving some regression equations derived from the maximum likelihood method so they are easy to implement. Additional advantages include (i) no loss of data observations and (ii) reduction of the volatility bias generated from the Fourier transform method.

Based on those developed Fourier transform methods to estimate the instantaneous volatility, this chapter further conducts two empirical studies as applications. The first empirical study investigates dynamic behaviors of volatility under three different sampling frequencies: high, medium, and low. In particular, the daily effect, a U shape of volatility, and evidence of multiple risk factors of volatility are observed. These observations are consistent with empirical findings in financial literatures. The second empirical study reveals the linear relationship between VIX and the instantaneous volatility. We derive a theoretical result and apply a general linear test to justify the linearity. Two datasets are used for empirical examination. They include TAIEX (January 2001–March, 2011) and S&P 500 index and its VIX (January 1990–January 2011). Data period covers both tranquil and turbulent times.

The organization of this chapter is as follows. Section 92.2 reviews the Fourier transform method and two price correction schemes, including a linear and a nonlinear correction method. Section 92.3 conducts some simulation tests for typical local and stochastic volatility models. Section 92.4 investigates the dynamic behavior of volatility under different sampling frequencies. Section 92.5 conducts a linear test for the instantaneous volatility and its VIX. Section 92.6 concludes this chapter.

92.2 Volatility Estimation: Introduction to Fourier Transform Method

Fourier transform method (Malliavin and Mancino (2009)) is a nonparametric method to estimate multivariate volatility process. Its main idea is to reconstruct volatility as time series in terms of sine and cosine basis under the following continuous semi-martingale assumption. Let u_t be the log-price of an underlying asset S at time t, i.e., $u_t = 1\text{n}\, S_t$, and follow a diffusion process

$$du_t = \mu_t dt + \sigma_t dW_t, \tag{92.1}$$

where μ_t is the instantaneous growth rate and W_t is a one-dimensional standard Brownian motion. One can estimate the time series volatility σ_t with the following steps.

- Step 1: Compute the Fourier coefficients of the underlying u_t as follows:

$$a_0(du) = \frac{1}{2\pi}\int_0^{2\pi} du_t, \tag{92.2}$$

$$a_k(du) = \frac{1}{\pi}\int_0^{2\pi} \cos(kt) du_t, \tag{92.3}$$

$$b_k(du) = \frac{1}{\pi}\int_0^{2\pi} \sin(kt) du_t, \tag{92.4}$$

  for any $k \geq 1$, so that $u(t) = a_0 + \sum_{k=1}^{\infty}\left[-\frac{b_k(du)}{k}\cos(kt) + \frac{a_k(du)}{k}\sin(kt)\right]$. Note that the original time interval [0, T] can always be rescaled to $[0, 2\pi]$ as shown in above integrals.
- Step 2: Compute the Fourier coefficients of variance σ_{t}^2 as follows:

$$a_k(\sigma^2) = \lim_{N\to\infty} \frac{\pi}{2N+1}\sum_{S=-N}^{N-k}\left[a_s^*(du)a_{s+k}^*(du) + b_s^*(du)b_{s+k}^*(du)\right] \tag{92.5}$$

$$b_k(\sigma^2) = \lim_{N\to\infty} \frac{\pi}{2N+1}\sum_{S=-N}^{N-k}\left[a_s^*(du)b_{s+k}^*(du) - b_s^*(du)a_{s+k}^*(du)\right], \tag{92.6}$$

  for $k \geq 0$, in which $a_s^*(du)$ and $b_s^*(du)$ are defined by

$$a_s^*(du) = \begin{cases} a_s(du), & if\, s > 0 \\ 0, & if\, s = 0 \\ a_{-s}(du), & if\, s < 0 \end{cases} \quad and \quad b_s^*(du) = \begin{cases} b_s(du), & if\, s > 0 \\ 0, & if\, s = 0 \\ -b_{-s}(du), & if\, s < 0. \end{cases}$$

- Step 3: Reconstruct the time series of variance σ_t^2 by

$$\sigma_t^2 = \lim_{N\to\infty} \sum_{k=0}^{N} \varphi(\delta k)\left[a_k\left(\sigma^2\right)\cos(kt) + b_k\left(\sigma^2\right)\sin(kt)\right], \tag{92.7}$$

where $\varphi(x) = \frac{\sin^2(x)}{x^2}$ is a smooth function with the initial condition $\varphi(0) = 1$ and δ is a smooth parameter typically specified as $\delta = \frac{1}{50}$ (Reno 2008).

From Eqs. 92.2, 92.3, and 92.4, it is observed that the integration error of Fourier coefficients is adversely proportional to data frequency. That is, when the data frequency gets higher, each integral becomes more accurate.

This Fourier transform method is easy to implement because, as shown in Eqs. 92.5 and 92.6, Fourier coefficients of the variance time series can be approximated by a finite sum of multiplications of a^* and b^*. This integration method can accordingly avoid drawbacks inherited from those traditional methods based on the differentiation of quadratic variation.

92.2.1 Price Correction Schemes: Bias Reduction

It is documented that this Fourier transform method incurs a "boundary effect." Reno (2008) notes that Fourier algorithm provides inaccurate estimate for volatility time series near the boundary of simulated data. He suggests that all the time series of estimated volatility near the first and last 1 % should be discarded for the purpose of better estimation. This compromise is in contrast to the Markov property, which is a key assumption in the stochastic financial theory Shreve (2000). For example, when one is about to compute the value at risk, evaluate option prices, or hedge financial derivatives, he or she may need the most updated volatility for computational tasks. Two correction schemes to remedy this boundary deficit are reviewed as follows.

Recall that u_t defined in Eq. 92.1 is the natural logarithm of asset price. Based on the Euler discretization, the increment of log-price u_t can be approximated by $\sigma_t\sqrt{\Delta_t \varepsilon_t}$. That is,

$$\Delta u_t \approx \sigma_t \sqrt{\Delta_t \varepsilon_t,} \tag{92.8}$$

where Δ_t denotes a small discretized time interval and ε_t denotes a sequence of i.i.d. standard normal random variables. This approximation is derived from neglecting the drift term of small order Δ_t and using the increment distribution of Brownian motion $\Delta W_t = \sqrt{\Delta_t \varepsilon_t,}$. Given a set of discrete observations of log returns, let $\hat{\sigma}_t$ denote the volatility time series estimated from the original Fourier transform method. Two correction schemes including nonlinear and linear correction methods have been proposed in Han et al. (2014) and Han (2014), respectively. These Fourier transform methods are effective to reduce bias of volatility estimation, known as the boundary effect.

1. **Nonlinear Correction Method**: This method consists of a linear transformation on the natural logarithm of estimated variance process $\hat{\sigma}_t^2$ in order to guarantee the positiveness of estimated volatility. That is, we transform $\hat{Y}_t = 2\ln\hat{\sigma}_t$ to $a + b\hat{Y}_t$ so that the corrected volatility $\sigma_t = \exp\left(\left(a + b\hat{Y}_t\right)/2\right) > 0$ satisfies $\Delta u_t \approx \exp\left(\left(a + b\hat{Y}_t\right)/2\right)\sqrt{\Delta_t \varepsilon_t}$, where $\Delta u_t = u_{t+1} - u_t$, and a and b denote the correction variables. This linear transformation on $\hat{Y}_t$ can be understood as the first-order approximation to a possible nonlinear transformation on estimated volatility $\hat{\sigma}_t$. Then, we can use the maximum likelihood method to regress out correction variables via the relationship between logarithm of squared standardized return $\Delta u_t/\sqrt{\Delta_t}$ and the driving volatility process $a + b\hat{Y}_t$:

$$\ln\left(\frac{\Delta u_t}{\sqrt{\Delta_t}}\right)^2 = a + b\hat{Y}_t + \ln\varepsilon_t^2. \tag{92.9}$$

2. **Linear Correction Method**: This method directly applies a linear transformation on the estimated volatility as

$$\sigma_t = a + b\hat{\sigma}_t. \tag{92.10}$$

Substituting this corrected volatility σ_t into Eq. 92.8, $\Delta u_t \approx \left(a + b\hat{\sigma}_t\right)\sqrt{\Delta t \varepsilon_t}$ is obtained. For regression purpose, we take squares on both sides, then a natural logarithm to obtain the following nonlinear equation:

$$\ln\left(\Delta u_t/\sqrt{\Delta_t}\right)^2 \approx \ln\left(a + b\hat{\sigma}_t\right)^2 + \ln{\varepsilon_t}^2.$$

We remark that there is no guarantee that the corrected volatility estimation defined in Eq. 92.10 remains positive. This is a disadvantage compared with the previous nonlinear correction method.

Note that these correction methods do not involve any model parameters, so they retain the spirit of non-parameterization.

92.3 Simulation Tests

In this section, two well-known volatility models including a local volatility model and a stochastic volatility model are considered for simulation studies in order to justify effectiveness of these Fourier transform methods.

92.3.1 Case I: Local Volatility Model

A local volatility model of the following form

$$dS_t = \alpha(m - S_t)dt + \beta S_t^{\gamma} dw_t$$

is employed. Model parameters are taken from Jiang (1998) based on the following estimation result: $\alpha = 0.093$, $m = 0.079$, $\beta = 0.794$, and $\gamma = 1.474$. A simulation procedure is used to generate sample processes of the price S_t and its volatility $\sigma_t = \beta S_t^{\gamma}$. This simulation is done by the Euler discretization with time step size $\Delta_t = 1/250$, and the total sample number is 5,000.

Based on the original Fourier transform method and two proposed price correction schemes, three volatility time series can be estimated. These are compared with the actual volatility series generated from the local volatility model. Criteria for error measurements include mean squared error (MSE) and maximum absolute error (MAE). Comparison results are shown below:

1. MSE: 7.52E-04 (original Fourier method), 1.19E-05 (nonlinear correction method), and 7.61E-06 (linear correction method)
2. MAE: 0.04 (original Fourier method), 0.02 (nonlinear correction method), and 0.01 (linear correction method)

Price correction methods are effective to reduce both error criteria at least by half in this simulated example. Other vast simulation studies also show similar results of bias reduction.

92.3.2 Case II: Stochastic Volatility Model

Stochastic volatility models often possess the mean-reverting property. Among various models, the Ornstein-Uhlenbeck process is often taken as the driving one-factor volatility model, which is also known as the exp-OU model in finance. It is defined as

$$\left\{ \begin{array}{c} dS_t = \mu S_t dt + \exp(Y_t/2) S_t dW_{1t}, \\ dY_t = \alpha(m - Y_t)dt + \beta dW_{2t}, \end{array} \right\} \tag{92.11}$$

where S_t denotes the underlying risky asset price, μ the return rate, and W_{1t} and W_{2t} are two correlated Brownian motions. The volatility process σ_t is defined as $\exp(Y_t/2)$, m denotes the long-run mean, α denotes the rate of mean reversion, β denotes the vol-vol (volatility of volatility), and Y_t denotes the Ornstein-Uhlenbeck process. Model parameters are set as follows: $\mu = 0.01$, $S_0 = 50$, $Y_0 = -2$, $m = -2$, $\alpha = 5$, $\beta = 1$, and $\rho = 0$. To simulate sample processes, the time discretization is set as $\Delta_t = 1/5{,}000$ and the total number of samples is 5,000.

The procedure to conduct our simulation study is as follows. First, time series of volatility $\sigma_t = \exp(Y_t/2)$ and the asset price S_t are simulated. Using the original Fourier transform method and the nonlinear correction method, two volatility time series are estimated to compare with the true volatility time series. Comparison results are as follows:

1. Mean squared error: 0.0324 (original FTM), 0.0025 (corrected FTM)
2. Maximum absolute error: 0.3504 (original FTM), 0.1563 (corrected FTM)

From these simulated results, over a half of errors are reduced by the corrected FTM. Similar reduction can also be found among a vast of simulations. It is worth noting that the employed nonlinear correction method is based on the maximum likelihood method, which is fairly easy to compute numerically.

This section conducted simulation tests for two corrected Fourier transform methods for the instantaneous volatility estimation. It is observed that those methods can effectively reduce bias generated from the truncation errors of the original Fourier transform method. Next, we apply these methods to analyze volatility of indices such as S&P 500 in the USA and TAIEX (Taiwan Stock Exchange Capitalization Weighted Stock Index) in Taiwan.

92.4 Volatility Estimation Under Different Sampling Frequencies

92.4.1 Volatility Daily Effect

It is well documented that volatility demonstrates a daily effect under the high sampling frequency. This effect causes a pattern of U shape on volatility, which is often observed from intraday data. TAIEX, complied by Taiwan Stock Exchange, is updated every 15 s and publically available. Such easy access to the data enables an exploration of the daily effect. The sampled data period is from February 10 to 15 for four consecutive trading days in 2011. In Fig. 92.1, the bottom line, labeled as vol with magnitude on the left Y-axis, shows four U shapes of the instantaneous volatility. Using an average-type deseasoning technique, see Wild and Seber (1999), time series of deseasoned volatility, labeled as Vol_Deseasoned with magnitude on the right Y-axis, is shown in the middle part of that figure. This deseasoned volatility is often used to measure the actual activity of volatility.

92.4.2 Multiple Risk Factors of Volatility

From the statistical point of view, volatility is a latent variable, and its estimation causes a lot of attentions during the last two decades. See review papers of Broto and Ruiz (2004), Molina et al. (2010), Yu (2010), and references therein.

In virtue of the nonparametric Fourier transform method, volatility is represented as a Fourier transformation of the underlying asset prices, which can be calculated but is subject to some truncation errors. Then given such estimation modified by our correction schemes, one can further investigate parameters given a prescribed model. Therefore, a new approach to estimate the exp-OU stochastic volatility model parameters can be given as follows:

Step 1: Use a corrected Fourier transform method to estimate the instantaneous volatility $\hat{\sigma}_t$. By taking a natural logarithm $\hat{Y}_t = 2\ln\hat{\sigma}_t$, the time series of the driving volatility OU becomes available.

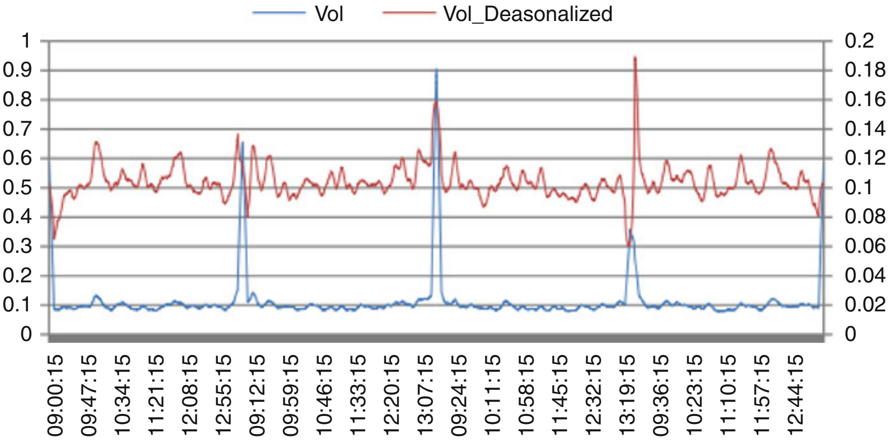


Fig. 92.1 2011/02/10~2011/02/15 TAIEX's estimated Ivol (every 15 s)

Step 2: Discretize the OU process and use the maximum likelihood method to estimate three parameters with the OU process. Detailed description can be found in Han et al. (2014), but estimators for parameters defined in Eq. 92.11 are given below:

$$\hat{\alpha} = \frac{1}{\Delta_t}\left[1 - \frac{\left(\sum_{t=2}^{N}\hat{Y}_t\right)\left(\sum_{t=1}^{N-1}\hat{Y}_t\right) - (N-1)\left(\sum_{t=1}^{N-1}\hat{Y}_t\hat{Y}_{t+1}\right)}{\left(\sum_{t=1}^{N-1}\hat{Y}_t\right)^2 - (N-1)\left(\sum_{t=1}^{N-1}\hat{Y}_t^2\right)}\right],$$

$$\hat{\beta} = \sqrt{\frac{1}{(N-1)\Delta_t}\sum_{t=1}^{N-1}\left[\hat{Y}_{t+1} - \left(\hat{\alpha}\, m\Delta_t + (1-\hat{\alpha}\,\Delta_t)\hat{Y}_t\right)\right]^2},$$

$$\hat{m} = \frac{-1}{\hat{\alpha}\Delta_t}\left[\frac{\left(\sum_{t=2}^{N}\hat{Y}_t\right)\left(\sum_{t=1}^{N-1}\hat{Y}_t^2\right) - \left(\sum_{t=1}^{N-1}\hat{Y}_t\right)\left(\sum_{t=1}^{N-1}\hat{Y}_t\hat{Y}_{t+1}\right)}{\left(\sum_{t=1}^{N-1}\hat{Y}_t\right)^2 - (N-1)\left(\sum_{t=1}^{N-1}\hat{Y}_t^2\right)}\right].$$

These estimators are employed to estimate model parameters under three different sampling frequencies, i.e., high (5 months_1 day, 5-min data for one trading day), medium (1 day_2 years, daily data for 2 years), and low (1 week_10 years, weekly data for 10 years). Parameter m, the long-run mean, is referred as the risk level of volatility. Figure 92.2 illustrates its estimation from January 3 to March 31 of 2011. To be more precise, for any given date, we use historical data of 1 day, 2 years, and 10 years separately, then adopt the M estimator for parameter m under

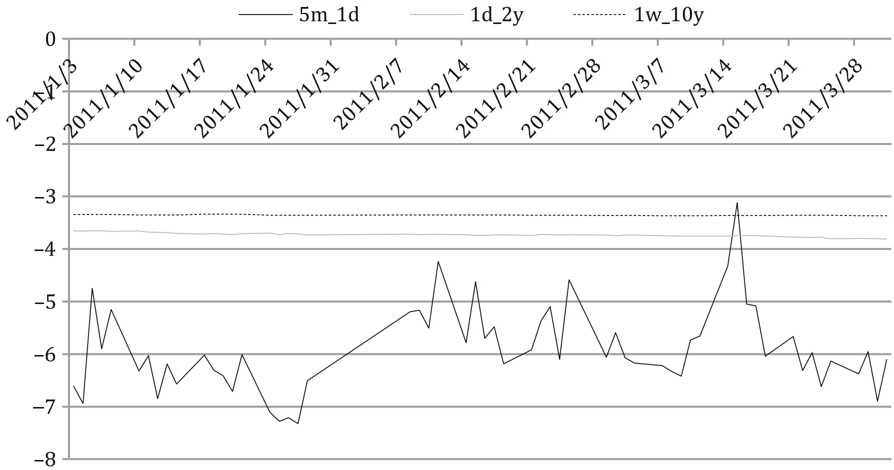


Fig. 92.2 2011/01/03~2011/03/31 long-run mean (m) of the exp-OU stochastic volatility model

these three different frequencies. Our dataset is TAIEX and the data resource is from Taiwan Stock Exchange.

Two other model parameters which include rate of mean reversion α and vol-vol β are related to invariance property of time scales. See Fouque et al. (2000) for detailed discussions.

From Table 92.1, both mean-reverting rate α and vol-vol β are proportional to data sampling frequency. That is, the higher the data frequency, the larger those parameter sizes are. Note that those parameters are well separated, especially the rate of mean reversion. These estimations show a strong evidence of multiple time scales on volatility.

92.5 Hypothesis of Linearity Between the Instantaneous Volatility and VIX

VIX (Volatility Index) is complied by Chicago Board Options Exchange (CBOE). Its formula is mainly based on out-of-money S&P 500 index options with weights depending on the corresponding strike prices. The full definition and formula can be found on Hull (2008) or the VIX white paper published by CBOE.[1] VIX is often used to measure an aggregated risk exposure in the next 30 calendar days. Hence, the information content of VIX is forward, as opposed to its dual backward information content of volatility, i.e., the 30-day historical volatility.

[1]www.cboe.com/micro/vix/vixwhite.pdf

Table 92.1 Model parameter estimations under different data frequencies

	α	(α _ std)	β	(β _ std)
5 months_1 day	11,661.39	(2,681.39)	114.83	(14.70)
1 day _ 2 years	23.81	(2.08)	6.29	(0.27)
1 week _ 10 years	2.62	(0.17)	2.18	(0.07)

In addition, due to a strong negative correlation with the S&P 500 index, VIX is popularly used as a fear gauge.

92.5.1 Theoretical Result

Zhang and Zhu (2006) obtained a linear relationship between the VIX squared and the square of the instantaneous volatility under the Heston model Heston (1993). This result is based on the following derivation. Under a risk-neutral probability measure, the instantaneous variance V_s is assumed to follow a square root process

$$dV_s = [\alpha m - (\alpha + \lambda)V_s]dt + \beta\sqrt{V_s dW_s},$$

where m denotes the long-run mean, α the rate of mean reversion, β the variance of variance, W_s the standard Brownian motion, and λ the volatility risk premium under the probability measure change. Taking mathematical expectations to the previous stochastic differential equation on both sides, we obtain

$$dE_t[V_s] = \alpha m - (\alpha + \lambda)E_t[V_s]dt. \tag{92.12}$$

This is because an exchange of integral and differential is assumed, and the expectation (conditioned at time t) over the Brownian motion term is zero. The notation $E_t[.]$ means a conditional expectation given the filtration at time t under a risk-neutral probability measure.

Equation 92.10 is a linear ordinary equation, and its solution is given as

$$E_t[V_s] = \frac{\alpha m}{\alpha + \lambda} + \left(V_t - \frac{\alpha m}{\alpha + \lambda}\right)e^{-(\alpha+\lambda)(s-t)}.$$

Since VIX^2 can also be defined as a variance swap rate Hull (2008), one can evaluate it by its definition

$$\begin{aligned} VIX_t^2 &= E_t\left[\frac{1}{\tau_0}\int_t^{t+\tau_0} V_s ds\right] \\ &= \frac{1}{\tau_0}\int_t^{t+\tau_0} E_t(V_s)ds, \end{aligned}$$

where τ_0 is defined by $\frac{30}{365}$. In summary, under the Heston model, a linear relationship between VIX^2 and the instantaneous variance V_t is deduced by

$$VIX_t^2 = A + BV_t,$$

where constant coefficients are

$$A = \frac{\alpha\, m}{\alpha+\lambda}\left[1 - \frac{1-e^{-(\alpha+\lambda)\tau_0}}{(\alpha+\lambda)\tau_0}\right], \quad B = \frac{1-e^{-(\alpha+\lambda)\tau_0}}{(\alpha+\lambda)\tau_0}.$$

Next, we generalize this result to a more general class of models, within which the long-run mean m and vol-vol β can be time dependent but deterministic. Even under his larger model class, we still can derive and obtain a linear relationship between VIX square and the instantaneous variance.

Theorem 1 Assume that the instantaneous variance σ_t follows a mean-reverting process

$$d\sigma_t = \alpha(m(t) - \sigma_t)dt + g(\sigma_t)dW_t, \tag{92.13}$$

where functions m and g exist such that the classical assumption of stochastic differential equations (see Oksendal (1998)) is satisfied. A linear relationship between VIX square and the instantaneous variance

$$VIX_t^2 = A + B\sigma_t^2,$$

where $A = -\frac{1}{\tau}\int_0^\tau\int_0^t e^{2\alpha(s-t)}h(s)dsdt$ and $B = \frac{1}{2\alpha\tau}\left(e^{2\alpha\tau} - 1\right)$.

Proof Apply Ito's Lemma to σ_t^2 to obtain

$$d\sigma_t^2 = \left(2\alpha m(t)\sigma_t - 2\alpha\sigma_t^2 + g^2(\sigma_t)\right)dt + 2\sigma_t g(\sigma_t)dW_t.$$

Then, take a mathematical expectation over both sides

$$dE\left[\sigma_t^2\right] = \left(-2\alpha E\left[\sigma_t^2\right] + h(t)\right)dt, \tag{92.14}$$

where $h(t) = E\lfloor 2\alpha m(t)\sigma_t + g^2(\sigma_t)\rfloor$ is a time-dependent deterministic function. This equation is known as the linear ordinary differential equation with an inhomogeneous term. The unique solution is given by

$$E\left[\sigma_t^2\right] = \sigma_0^2 e^{2\alpha t} - \int_0^t e^{2\alpha(s-t)}h(s)ds,$$

and VIX^2 defined as $VIX_0^2 = \frac{1}{\tau}\int_0^\tau E\sigma_t^2 dt = A + B\sigma_0^2$. Coefficients are $A = -\frac{1}{\tau}\int_0^\tau\int_0^t e^{2\alpha(s-t)}h(s)dsdt$ and $B = \frac{1}{2\alpha\tau}\left(e^{2\alpha\tau} - 1\right)$.

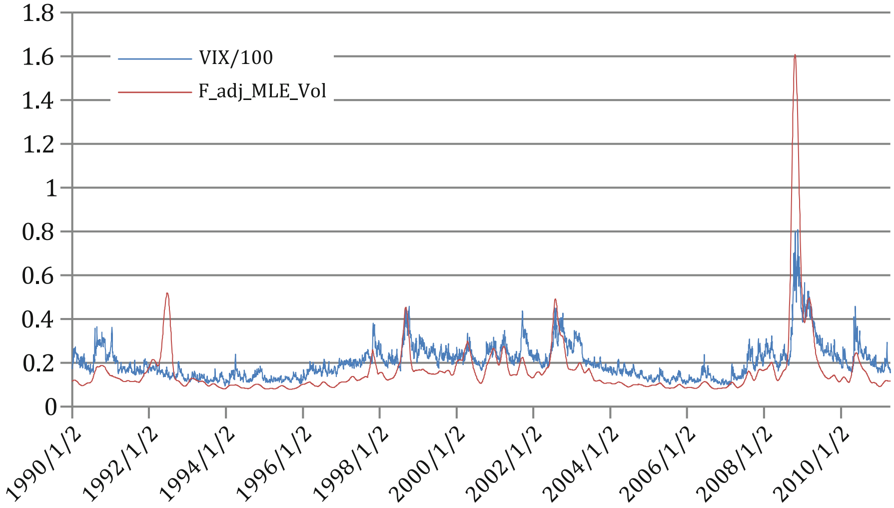


Fig. 92.3 Daily VIX, denoted by VIX/100, and the estimated instantaneous volatility, denoted by F_adj_MLE_Vol, of S&P 500 index within the time period from January 2, 1990, to January 31, 2011

92.5.2 Empirical Results

The previous theorem advocates a linear relationship between the VIX squared and the instantaneous variance. No empirical analysis on this theoretical result is found to our best knowledge. The reason may come from the fact that estimation for the instantaneous variance was difficult. In virtue of the corrected Fourier transform methods, one can estimate time series σ_t^2. Daily VIX and the estimated instantaneous volatility of S&P 500 index within the time period from January 2, 1990, to January 31, 2011, are shown in Fig. 92.3. Our data resource is from Yahoo! Finance.

We adopt the general linear test approach, see Kutner et al. (2005), to examine the linear relationship between VIX^2 and the instantaneous variance σ^2. Assuming that the null hypothesis is

$$\begin{cases} H_0 : \beta_1 = 0, \\ H_1 : \beta_1 \neq 0. \end{cases}$$

and the alternative is $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$, one can use the least squares method to obtain the sum of squared errors, denoted by $SSE(F)$, which is chi-square distributed with the degree of freedom (n-2) and denoted by df_F. Similarly under the null hypothesis, the model is assumed to be $Y_i = \beta_O + \varepsilon_i$, and the sum of squared errors, denoted by $SSE(R)$, which is also chi-square distributed with the degree of freedom (n-1), denoted by df_R. The F statistic is defined by

$$F^* = \frac{\frac{SSE(R)-SSE(F)}{df_R - df_F}}{\frac{SSE(F)}{df_F}} \sim F(1-\alpha, df_R - df_F, df_F),$$

where the confidence level is $(1 - \alpha)$. Our data size is $n = 1{,}091$ and estimated linear equation is

$$\text{VIX\_2} = 0.039342 + 0.180687\text{V}.$$

Given the significance level alpha = 1 %, the value of F statistic is equal to 6192.8 > F(0.99, 1, n − 2) = 6.639621 so that the null is rejected.

92.6 Concluding Remarks and Future Works

This chapter presents methods of nonparametric Fourier transform for estimating the instantaneous volatility in one dimension. Based on these results, we conduct simulation tests for a local model and a stochastic volatility model and two empirical studies. They include (1) volatility behaviors under different sampling frequencies and (2) linear hypothesis between the instantaneous volatility and VIX.

As a matter of fact, this method can be used for high-dimensional case. That allows for estimating dynamic volatility matrices. As a result, this whole approach may be suitable to study subjects of portfolio risk management, systemic risk analysis, etc. We leave these as future works.

References

Broto, C., & Ruiz, E. (2004). Estimation methods for stochastic volatility models: A survey. *Journal of Economic Surveys, 18*(5), 613–649.

Engle, R. (2009). *Anticipating correlations: A new paradigm for risk management*. Princeton, New Jersey: Princeton University Press.

Fouque, J.-P., Papanicolaou, G., & Sircar, R. (2000). *Derivatives in financial markets with stochastic volatility*. Cambridge: Cambridge University Press.

Gatheral, J. (2006). *The volatility surface*. Hoboken, New Jersey: Wiley.

Han, C.-H., Liu, W.-H., & Chen, T.-Y. (2014). VaR/CVaR Estimation under Stochastic Volatility Models. To appear on International Journal of Theoretical & Applied Finance.

C.H. Han, C.-H. Chang, C.-S. Kuo, S.-T. Yu. (2014) Robust hedging performance and volatility risk in option markets: Application to Standard and Poor's 500 and Taiwan index options. To appear on the special issue "Advances in Financial Risk Management and Economic Policy Uncertainty" of International Review of Economics and Finance.

Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies, 6*(2), 327–343.

Hull, J. (2008). *Options, futures, and other derivatives* (7th ed.). Prentice Hall.

Jiang, G. J. (1998). Nonparametric modeling of U.S. interest term structure dynamic and implications on the prices of derivative securities. *The Journal of Financial and Quantitative Analysis, 33*(4), 465–497.

Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). *Applied linear statistical models* (5th ed.). McGraw-Hill/Irwin.

Malliavin, P., & Mancino, M. E. (2009). A Fourier transform method for nonparametric estimation of multivariate volatility. *Annals of Statistics, 37*(4), 1983–2010.

Molina, G., Han, C.-H., & Fouque, J. P. (2010). MCMC estimation of multiscale stochastic volatility models. In *Handbook of quantitative finance and risk management*. Springer.

Oksendal, B. (1998). Stochastic differential equations (5th ed.). Springer.

Reno, R. (2008). Nonparametric estimation of the diffusion coefficient of stochastic volatility models. *Econometric Theory, 24*(5), 1174–1206.

Shreve, S. E. (2000). *Stochastic calculus for finance II: Continuous-time models*. Springer.

Tsai, R. S. (2005). *Analysis of financial time series* (2nd ed.). Wiley.

Wild, C. J. & Seber, G. A. F. (1999). *Chance encounters: A first course in data analysis and inference*. Wiley.

Yu, Jun. (2010). Simulation-based estimation methods for financial time series models. In J.-C. Duan, J. E. Gentle, & W. Hardle (Eds.), *Handbook of computational finance*. Springer.

Zhang, L., & Mykland, P. (2005). A tale of two time scales: Determining integrated volatility with noise high frequency data. *Journal of American Statistics, 100*, 1394–1411.

Zhang, J. E., & Zhu, Y. (2006). VIX Futures. *The Journal of Futures Markets, 26*(6), 521–531.

A Dynamic CAPM with Supply Effect Theory and Empirical Results

93

Cheng-Few Lee, Chiung-Min Tsai, and Alice C. Lee

Contents

C.-F. Lee (✉)
Department of Finance and Economics, Rutgers Business School, Rutgers, The State University of New Jersey, Piscataway, NJ, USA

Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
e-mail: cflee@business.rutgers.edu; lee@business.rutgers.edu

C.-M. Tsai
Central Bank of the Republic of China (Taiwan), Taipei, Taiwan, Republic of China
e-mail: cmtsai@mail.cbc.gov.tw

A.C. Lee
State Street Corp., USA
e-mail: alice.finance@gmail.com

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_93,

Abstract

Breeden (1979) and Grinols (1984) and Cox et al. (1985) have described the importance of supply side for the capital asset pricing. Black (1976) derives a dynamic, multiperiod CAPM, integrating endogenous demand and supply. However, Black's theoretically elegant model has never been empirically tested for its implications in dynamic asset pricing. We first theoretically extend Black's CAPM. Then, we use price, dividend per share, and earnings per share to test the existence of supply effect with US equity data. We find the supply effect is important in US domestic stock markets. This finding holds as we break the companies listed in the S&P 500 into ten portfolios by different level of payout ratio. It also holds consistently if we use individual stock data.

A simultaneous equation system is constructed through a standard structural form of a multiperiod equation to represent the dynamic relationship between supply and demand for capital assets. The equation system is exactly identified under our specification. Then, two hypotheses related to supply effect are tested regarding the parameters in the reduced form system. The equation system is estimated by the seemingly unrelated regression (SUR) method, since SUR allow one to estimate the presented system simultaneously while accounting for the correlated errors.

Keywords

CAPM • Asset • Endogenous supply • Simultaneous equations • Reduced form • Seemingly unrelated regression (SUR) • Exactly identified • Cost of capital • Quadratic cost • Partial adjustment

93.1 Introduction

Breeden (1979) and Grinols (1984) and Cox et al. (1985) have described the importance of supply side for the capital asset pricing. Cox et al. (1985) study a restricted technology to allow them to explicitly solve their model for reduced form. Grinols (1984) focuses on describing market optimality and supply decisions which guide firms in incomplete markets in the absence of investor unanimity. Black (1976) extends the static CAPM by Sharpe (1964), Litner (1965), and Mossin (1966) explicitly allowing for the endogenous supply effect of risky securities to derive the dynamic asset pricing model.[1] Black modifies the static model by explicitly allowing for the existence of the supply effect of risky securities. In addition, the demand side for the risky securities is derived from a negative exponential function for the investor's utility of wealth. Black finds that the static

[1]This dynamic asset pricing model is different from Merton's (1973) intertemporal asset pricing model in two key aspects. First, Black's model is derived in the form of simultaneous equations. Second, Black's model is derived in terms of price change, and Merton's model is derived in terms of rates of return.

CAPM is unnecessarily restrictive in its neglect of the supply side and proposes that his dynamic generalization of the static CAPM can provide the basis for many empirical tests, particularly with regard to the intertemporal aspects and the role of the endogenous supply side. Assuming that there is a quadratic cost structure of retiring or issuing securities and that the demand for securities may deviate from supply due to anticipated and unanticipated random shocks, Black concludes that if the supply of a risky asset is responsive to its price, large price changes will be spread over time as specified by the dynamic capital asset pricing model. One important implication in Black's model is that the efficient market hypothesis holds only if the supply of securities is fixed and independent of current prices. In short, Black's dynamic generalization model of static wealth-based CAPM adopts an endogenous supply side of risky securities by setting equal quantity demanded and supplied of risky securities. Lee and Gweon (1986) extend Black's framework to allow time-varying dividend payments and then test the existence of supply effect in the situation of market equilibrium. Their results reject the null hypothesis of no supply effect in the US domestic stock market. The rejection seems to imply a violation of efficient market hypothesis in the US stock market.

It is worth noting that some recent studies also relate return on portfolio to trading volume (e.g., Campbell et al. 1993; Lo and Wang 2000). Surveying the relationship between aggregate stock market trading volume and the serial correlation of daily stock returns, (Campbell et al. 1993) suggest that a stock price decline on a high-volume day is more likely than a stock price decline on a low-volume day. They propose an explanation that trading volume occurs when random shifts in the stock demand of non-informational traders are accommodated by the risk-averse market makers. Lo and Wang (2000) also examine the CAPM in the intertemporal setting. They derive an intertemporal CAPM (ICAPM) by defining preference for wealth instead of consumption, by introducing three state variables into the exponential types of investor's preference as we do in this paper. This state-dependent utility function allows one to capture the dynamic nature of the investment problem without explicitly solving a dynamic optimization problem. Thus, the marginal utility of wealth depends not only on the dividend of the portfolio but also on future state variables. This dependence introduces dynamic hedging motives in the investors' portfolio choices. That is, this dependence induces investors to care about future market conditions when choosing their portfolio. In equilibrium, this model also implies that an investor's utility depends not only on his wealth but also on the stock payoffs directly. This "market spirit," in their terminology, affects investor's demand for the stocks. In other words, for even the investor who holds no stocks, his utility fluctuates with the payoffs of the stock index.

Black (1976), Lee and Gweon (1986), and Lo and Wang (2000) develop models by using either outstanding shares or trading volumes as variables to connect the decisions in two different periods, unlike consumption-based CAPM which uses consumption or macroeconomic information. Black (1976) and Lee and Gweon (1986) both derive the dynamic generalization models from the wealth-based CAPM by adopting an endogenous supply schedule of risky securities.[2] Thus, the information of quantities demanded and supplied can now play a role in determining the asset price. This proposes a wealth-based model as an alternative method to investigate intertemporal CAPM.

In this chapter, we first theoretically extend the Black's dynamic, simultaneous CAPM to be able to test the existence of the supply effect in the asset pricing determination process. We use two datasets of price per share and dividend per share to test the existence of supply effect with US equity data. The first dataset consists most companies listing in the S&P 500 of the US stock market. The second dataset is the companies listed in the Dow Jones Index. In this study, we find the supply effect is important in the US stock market. This finding holds as we break the companies listed in the S&P 500 into ten portfolios. It also holds if we use individual stock data. For example, the existence of supply effect holds consistently in most portfolios if we test the hypotheses by using individual stock as many as 30 companies in one group. We also find that one cannot reject the existence of supply effect by using the stocks listed in the Dow Jones Index.

This chapter is structured as follows. In Sect. 93.2, a simultaneous equation system of asset pricing is constructed through a standard structural form of a multiperiod equation to represent the dynamic relationship between supply and demand for capital assets. The hypotheses implied by the model are also presented in this section. Section 93.3 describes the two sets of data used in this paper. The empirical finding for the hypotheses and tests constructed in previous section is then presented. Our summary is presented in Sect. 93.4.

93.2 Development of Multiperiod Asset Pricing Model with Supply Effect

Based on the framework of Black (1976), we derive a multiperiod equilibrium asset pricing model in this section. Black modifies the static wealth-based CAPM by explicitly allowing for the endogenous supply effect of risky securities. The demand for securities is based on the well-known model of James Tobin (1958) and Harry Markowitz (1959). However, Black further assumes a quadratic cost function of changing short-term capital structure under long-run optimality condition. He also assumes that the demand for security may deviate from supply due to anticipated and unanticipated random shocks.

Lee and Gweon (1986) modify and extend Black's framework to allow time-varying dividends and then test the existence of supply effect. In Lee and Gweon's model, two major differing assumptions from Black's model are: (1) the model allows for time-varying dividends, unlike Black's assumption constant dividends,

[2]It should be noted that Lo and Wang's model did not explicitly introduce the supply equation in asset pricing determination. Also, one can identify the hedging portfolio using volume data in the Lo and Wang model setting.

and (2) there is only one random, unanticipated shock in the supply side instead of two shocks, anticipated and unanticipated shocks, as in Black's model. We follow the Lee and Gweon set of assumptions. In this section, we develop a simultaneous equation asset pricing model. First, we derive the demand function for capital assets, then we derive the supply function of securities. Next, we derive the multiperiod equilibrium model. Thirdly, the simultaneous equation system is developed for testing the existence of supply effects. Finally, the hypotheses of testing supply effect are developed.

93.2.1 The Demand Function for Capital Assets

The demand equation for the assets is derived under the standard assumptions of the CAPM.[3] An investor's objective is to maximize their expected utility function. A negative exponential function for the investor's utility of wealth is assumed:

$$U = a - h \times e^{\{-bW_{t+1}\}}, \tag{93.1}$$

where the terminal wealth $W_{t+1} = W_t(1 + R_t)$; W_t is initial wealth; and R_t is the rate of return on the portfolio. The parameters a, b, and h are assumed to be constants.

The dollar returns on N marketable risky securities can be represented by

$$X_{j,\ t+1} = P_{j,\ t+1} - P_{j,\ t} + D_{j,\ t+1},\ j = 1, \ldots, N, \tag{93.2}$$

where
$P_{j,\ t+1}$ = (random) price of security j at time t + 1
$P_{j,\ t}$ = price of security j at time t
$D_{j,\ t+1}$ = (random) dividend or coupon on security at time t + 1

These three variables are assumed to be jointly normally distributed. After taking the expected value of Eq. 93.2 at time t, the expected returns for each security, $x_{j,\ t+1}$, can be rewritten as

$$x_{j,t+1} = E_t X_{j,t+1} = E_t P_{j,t+1} - P_{j,t} + E_t D_{j,t+1}, \quad j = 1, \ldots, N, \tag{93.3}$$

where

$$\begin{aligned} E_t P_{j,t+1} &= E\left(P_{j,t+1} | \Omega_t\right) \\ E_t D_{j,t+1} &= E\left(D_{j,t+1} | \Omega_t\right) \\ E_t X_{j,t+1} &= E\left(X_{j,t+1} | \Omega_t\right) \end{aligned}$$

Ω_t is the given information available at time t.

[3]The basic assumptions are as follows: (1) a single period moving horizon for all investors; (2) no transaction costs or taxes on individuals; (3) the existence of a risk-free asset with rate of return, r*; (4) evaluation of the uncertain returns from investments in terms of expected return and variance of end-of-period wealth; and (5) unlimited short sales or borrowing of the risk-free asset.

Then, a typical investor's expected value of end-of-period wealth is

$$w_{t+1} = E_t W_{t+1} = W_t + r^* \left(W_t - q_{t+1}{}' P_t \right) + q_{t+1}{}' x_{t+1}, \tag{93.4}$$

where
$P_t = (P_{1,\,t}, P_{2,\,t}, P_{3,\,t}, \ldots, P_{N,\,t})'$
$x_{t+1} = (x_{1,\,t+1}, x_{2,\,t+1}, x_{3,\,t+1}, \ldots, x_{N,\,t+1})' = E_t P_{t+1} - P_t + E_t D_{t+1}$
$q_{t+1} = (q_{1,\,t+1}, q_{2,\,t+1}, q_{3,\,t+1}, \ldots, q_{N,\,t+1})'$
$q_{j,t+1}$ = number of units of security j after reconstruction of his portfolio
r^* = risk-free rate

In Eq. 93.4, the first term on the right hand side is the initial wealth, the second term is the return on the risk-free investment, and the last term is the return on the portfolio of risky securities. The variance of W_{t+1} can be written as

$$V(W_{t+1}) = E(W_{t+1} - w_{t+1})(W_{t+1} - w_{t+1})' = q_{t+1}{}' S q_{,\,t+1}, \tag{93.5}$$

where $S = E(X_{t+1} - x_{t+1})(X_{t+1} - x_{t+1})'$ = the covariance matrix of returns of risky securities.

Maximization of the expected utility of W_{t+1} is equivalent to:

$$\text{Max } w_{t+1} - \frac{b}{2} V(W_{t+1}), \tag{93.6}$$

By substituting Eqs. 93.4 and 93.5 into Eq. 93.6, Eq. 93.6 can be rewritten as:

$$\text{Max}(1 + r^*) W_t + q_{t+1}{}'(x_{t+1} - r^* P_t) - (b/2) q_{t+1}{}' S q_{t+1.} \tag{93.7}$$

Differentiating Eq. 93.7, one can solve the optimal portfolio as:

$$q_{t+1} = b^{-1} S^{-1} (x_{t+1} - r^* P_t). \tag{93.8}$$

Under the assumption of homogeneous expectation, or by assuming that all the investors have the same probability belief about future return, the aggregate demand for risky securities can be summed as:

$$Q_{t+1} = \sum_{k=1}^{m} q_{t+1}^{k} = cS^{-1}[E_t P_{t+1} - (1 + r^*) P_t + E_t D_{t+1}], \tag{93.9}$$

where $c = \Sigma(b^k)^{-1}$.

In the standard CAPM, the supply of securities is fixed, denoted as Q^*. Then, Eq. 93.9 can be rearranged as $P_t = (1/r^*)(x_{t+1} - c^{-1} S Q^*)$, where c^{-1} is the market price of risk. In fact, this equation is similar to the Lintner's (1965) well-known equation in capital asset pricing.

93.2.2 Supply Function of Securities

An endogenous supply side to the model is derived in this section, and we present our resulting hypotheses, mainly regarding market imperfections. For example, the existence of taxes causes firms to borrow more since the interest expense is tax-deductible. The penalties for changing contractual payment (i.e., direct and indirect bankruptcy costs) are material in magnitude, so the value of the firm would be reduced if firms increase borrowing. Another imperfection is the prohibition of short sales of some securities.[4] The costs generated by market imperfections reduce the value of a firm, and, thus, a firm has incentives to minimize these costs. Three more related assumptions are made here. First, a firm cannot issue a risk-free security; second, these adjustment costs of capital structure are quadratic; and third, the firm is not seeking to raise new funds from the market.

It is assumed that there exists a solution to the optimal capital structure and that the firm has to determine the optimal level of additional investment. The one-period objective of the firm is to achieve the minimum cost of capital vector with adjustment costs involved in changing the quantity vector, $\mathrm{Q}_{\mathrm{i,\ t+1}}$:

$$\begin{aligned}&\text{Min } \mathrm{E_t}D_{\mathrm{i,t+1}}\mathrm{Q}_{\mathrm{i,t+1}} + (1/2)\left(\Delta\mathrm{Q}_{\mathrm{i,t+1}}{}'\mathrm{Ai}\Delta\mathrm{Q}_{\mathrm{i,t+1}}\right),\\ &\text{subject to} P_{\mathrm{i,t}}\Delta\mathrm{Q}_{\mathrm{i,t+1}} = 0,\end{aligned} \tag{93.10}$$

where $\mathrm{A_i}$ is a $\mathrm{n_i} \times \mathrm{n_i}$ positive-definite matrix of coefficients measuring the assumed quadratic costs of adjustment. If the costs are high enough, firms tend to stop seeking raise new funds or retire old securities. The solution to Eq. 93.10 is

$$\Delta\mathrm{Q}_{\mathrm{i,t+1}} = \mathrm{A_i}^{-1}\left(\lambda_{\mathrm{i}}P_{\mathrm{i,t}} - \mathrm{E_t}D_{\mathrm{i,t+1}}\right), \tag{93.11}$$

where λ_i is the scalar Lagrangian multiplier.

Aggregating Eq. 93.11 over N firms, the supply function is given by

$$\Delta\mathrm{Q}_{\mathrm{t+1}} = \mathrm{A}^{-1}(\mathrm{B}P_{\mathrm{t}} - \mathrm{E_t}D_{\mathrm{t+1}}), \tag{93.12}$$

where $A^{-1} = \begin{bmatrix} A_1^{-1} & & & \\ & A_2^{-1} & & \\ & & \ddots & \\ & & & A_N^{-1} \end{bmatrix}$, $B = \begin{bmatrix} \lambda_1 I & & & \\ & \lambda_2 I & & \\ & & \ddots & \\ & & & \lambda_N I \end{bmatrix}$, and

$$Q = \begin{bmatrix} Q_1 \\ Q_2 \\ \vdots \\ Q_N \end{bmatrix}.$$

[4]Theories as to why taxes and penalties affect capital structure are first proposed by Modigliani and Miller (1958) and then Miller (1977). Another market imperfection, prohibition on short sales of securities, can generate "shadow risk premiums" and, thus, provide further incentives for firms to reduce the cost of capital by diversifying their securities.

Equation 93.12 implies that a lower price for a security will increase the amount retired of that security. In other words, the amount of each security newly issued is positively related to its own price and is negatively related to its required return and the prices of other securities.

93.2.3 Multiperiod Equilibrium Model

The aggregate demand for risky securities presented by Eq. 93.9 can be seen as a difference equation. The prices of risky securities are determined in a multiperiod framework. It is also clear that the aggregate supply schedule has similar structure. As a result, the model can be summarized by the following equations for demand and supply, respectively:

$$\mathrm{Q}_{t+1} = \mathrm{cS}^{-1}(\mathrm{E}_t P_{t+1} - (1 + \mathrm{r}^*)P_t + \mathrm{E}_t D_{t+1}), \tag{93.9}$$

$$\Delta\mathrm{Q}_{t+1} = \mathrm{A}^{-1}(\mathrm{B}P_t - \mathrm{E}_t D_{t+1}). \tag{93.12}$$

Differencing Eq. 93.9 for period t and t+1 and equating the result with Eq. 93.12, a new equation relating demand and supply for securities is

$$\mathrm{cS}^{-1}[\mathrm{E}_t P_{t+1} - \mathrm{E}_{t-1}P_t - (1 + \mathrm{r}^*)(P_t - P_{t-1}) + \mathrm{E}_t D_{t+1} - \mathrm{E}_{t-1}D_t] = \mathrm{A}^{-1}(\mathrm{B}P_t - \mathrm{E}_t D_{t+1}) + \mathrm{V}_t, \tag{93.13}$$

where V_t is included to take into account the possible discrepancies in the system. Here, V_t is assumed to be random disturbance with zero expected value and no autocorrelation.

Obviously, Eq. 93.13 is a second-order system of stochastic differential equation in P_t and conditional expectations $\mathrm{E}_{t-1}P_t$ and $\mathrm{E}_{t-1}D_t$. By taking the conditional expectation at time t−1 on Eq. 93.13, and because of the properties of $\mathrm{E}_{t-1}[\mathrm{E}_t P_{t+1}] = \mathrm{E}_{t-1}P_{t+1}$ and $\mathrm{E}_{t-1}\mathrm{E}(\mathrm{V}_t) = 0$, Eq. 93.13 becomes

$$\begin{aligned}&\mathrm{cS}^{-1}[\mathrm{E}_{t-1}P_{t+1} - \mathrm{E}_{t-1}P_t - (1 + \mathrm{r}^*)(\mathrm{E}_{t-1}P_t - P_{t-1}) + \mathrm{E}_{t-1}D_{t+1} - \mathrm{E}_{t-1}D_t]\\ &= \mathrm{A}^{-1}(\mathrm{B}\mathrm{E}_{t-1}P_t - \mathrm{E}_{t-1}D_{t+1}).\end{aligned} \tag{93.13$'$}$$

Subtracting Eq. 93.13′ from Eq. 93.13,

$$\begin{aligned}&\left[(1 + \mathrm{r}^*)\mathrm{cS}^{-1} + \mathrm{A}^{-1}\mathrm{B}\right](P_t - \mathrm{E}_{t-1}P_t) = \mathrm{cS}^{-1}(\mathrm{E}_t P_{t+1} - \mathrm{E}_{t-1}P_{t+1})\\ &+\left(\mathrm{cS}^{-1} + \mathrm{A}^{-1}\right)(\mathrm{E}_t D_{t+1} - \mathrm{E}_{t-1}D_{t+1}) - \mathrm{V}_t.\end{aligned} \tag{93.14}$$

Equation 93.14 shows that prediction errors in prices (the left hand side) due to unexpected disturbance are a function of expectation adjustments in price (first term on the right hand side) and dividends (the second term on the right hand side) two periods ahead. This equation can be seen as a generalized capital asset pricing model.

One important implication of the model is that the supply side effect can be examined by assuming the adjustment costs are large enough to keep the firms from seeking to raise new funds or to retire old securities. In other words, the assumption of high enough adjustment costs would cause the inverse of matrix A in Eq. 93.14 to vanish. The model is, therefore, reduced to the following certain equivalent relationship:

$$P_t - E_{t-1}P_t = (1+r^*)^{-1}(E_tP_{t+1} - E_{t-1}P_{t+1}) + (1+r^*)^{-1}(E_tD_{t+1} - E_{t-1}D_{t+1}) + U_t, \tag{93.15}$$

Where $U_t = -c^{-1}S(1 + r^*)^{-1}V_t$.

Equation 93.15 suggests that current forecast error in price is determined by the sum of the values of the expectation adjustments in its own next-period price and dividend discounted at the rate of 1 + r*.

93.2.4 Derivation of Simultaneous Equation System

From Eq. 93.15, if price series follow a random walk process, then the price series can be represented as $P_t = P_{t-1} + a_t$, where a_t is white noise. It follows that $E_{t-1}P_t = P_{t-1}$, $E_tP_{t+1} = P_t$and $E_{t-1}P_{t+1} = P_{t-1}$. According to the results in Appendix 1, the assumption that price follows a random walk process seems to be reasonable for both datasets. As a result, Eq. 93.14 becomes

$$-\left(r^*cS^{-1} + A^{-1}B\right)(P_t - P_{t-1}) + \left(cS^{-1} + A^{-1}\right)(E_tD_{t+1} - E_{t-1}D_{t+1}) = V_t. \tag{93.16}$$

Equation 93.16 can be rewritten as

$$G\,p_t + H\,d_t = V_t, \tag{93.17}$$

where
$G = -\ (r * cS^{-1} + A^{-1}B)$
$H = (cS^{-1} + A^{-1})$
$d_t = E_tD_{t+1} - E_{t-1}D_{t+1}$
$p_t = P_t - P_{t-1}$.

If Eq. 93.17 is exactly identified and matrix G is assumed to be nonsingular, then as shown in Greene (2004), the reduced form of this model may be written as[5]

$$p_t = \Pi d_t + U_t, \tag{93.18}$$

[5]The identification of the simultaneous equation system can be found in Appendix 2.

where Π is a n-by-n matrix of the reduced form coefficients and U_t is a column vector of n reduced form disturbances. Or

$$\Pi = -G^{-1}H, \text{and } U_t = G^{-1}V_t. \tag{93.19}$$

Equations 93.18 and 93.19 are used to test the existence of supply effect in the next section.

93.2.5 Test of Supply Effect

Since the simultaneous equation system as in Eq. 93.17 is exactly identified, it can be estimated by the reduced form as Eq. 93.18. A proof of identification problem of Eq. 93.17 is shown in Appendix 2. That is, Eq. 93.18, $p_t = \Pi d_t + U_t$, can be used to test the supply effect. For example, in the case of two portfolios, the coefficient matrix G and H in Eq. 93.17 can be written as[6]

$$G = \begin{bmatrix} g_{11} & g_{12} \\ g_{21} & g_{22} \end{bmatrix} = \begin{bmatrix} -(r^* cs_{11} + a_1 b_1) & -r^* cs_{12} \\ -r^* cs_{21} & -(r^* cs_{22} + a_2 b_2) \end{bmatrix},$$

$$H = \begin{bmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{bmatrix} = \begin{bmatrix} cs_{11} + a_1 & cs_{12} \\ cs_{21} & cs_{22} + a_2 \end{bmatrix}. \tag{93.20}$$

Since $\Pi = -G^{-1}H$ in Eq. 93.21, Π can be calculated as

$$\begin{aligned} -G^{-1}H &= \begin{bmatrix} r^* cs_{11} + a_1 b_1 & r^* cs_{12} \\ r^* cs_{21} & r^* cs_{22} + a_2 b_2 \end{bmatrix}^{-1} \begin{bmatrix} cs_{11} + a_1 & cs_{12} \\ cs_{21} & cs_{22} + a_1 \end{bmatrix} \\ &= \frac{1}{|G|} \begin{bmatrix} r^* cs_{22} + a_2 b_2 & -r^* cs_{12} \\ -r^* cs_{21} & r^* cs_{11} + a_1 b_1 \end{bmatrix} \begin{bmatrix} cs_{11} + a_1 & cs_{12} \\ cs_{21} & cs_{22} + a_1 \end{bmatrix} \\ &= \frac{1}{|G|} \begin{bmatrix} (r^* cs_{22} + a_2 b_2)(cs_{11} + a_1) - r^* cs_{12} cs_{21} & (r^* cs_{22} + a_2 b_2) cs_{12} - r^* cs_{12}(cs_{22} + a_1) \\ -r^* cs_{21}(cs_{11} + a_1) + (r^* cs_{11} + a_1 b_1) cs_{21} & -r^* cs_{21} cs_{12} + (r^* cs_{11} + a_1 b_1)(cs_{22} + a_1) \end{bmatrix} \\ &= \begin{bmatrix} \pi_{11} & \pi_{12} \\ \pi_{21} & \pi_{22} \end{bmatrix}. \end{aligned} \tag{93.21}$$

From Eq. 93.21, if there is a high enough quadratic cost of adjustment, or if $a_1 = a_2 = 0$, then with $s_{12} = s_{21}$, the matrix would become a scalar matrix in which diagonal elements are equal to $r^* c^2 (s_{11} s_{22} - s_{12}^2)$, and the off-diagonal elements are all zero. In other words, if there is high enough cost of adjustment, firm tends to stop seeking to raise new funds or to retire old securities. Mathematically, this will be represented in a way that all off-diagonal elements are all zero and all diagonal

[6] s_{ij} is the ith row and jth column of the variance-covariance matrix of return. a_i and b_i are the supply adjustment cost of firm i and overall cost of capital of firm i, respectively.

elements are equal to each other in matrix Π. In general, this can be extended into the case of more portfolios. For example, in the case of N portfolios, Eq. 93.18 becomes

$$\begin{bmatrix} p_{1t} \\ p_{2t} \\ \vdots \\ p_{Nt} \end{bmatrix} = \begin{bmatrix} \pi_{11} & \pi_{12} & \cdots & \pi_{1N} \\ \pi_{21} & \pi_{22} & \cdots & \pi_{2N} \\ \vdots & \vdots & \ddots & \vdots \\ \pi_{N1} & \pi_{N2} & \cdots & \pi_{NN} \end{bmatrix} \begin{bmatrix} d_{1t} \\ d_{2t} \\ \vdots \\ d_{Nt} \end{bmatrix} + \begin{bmatrix} u_{1t} \\ u_{2t} \\ \vdots \\ u_{Nt} \end{bmatrix}. \tag{93.22}$$

Equation 93.22 shows that if an investor expects a change in the prediction of the next dividend due to additional information (e.g., change in earnings) during the current period, then the price of the security changes. Regarding the US equity market, if one believes that how the expectation errors in dividends are built into the current price is the same for all securities, then, the price changes would be only influenced by its own dividend expectation errors. Otherwise, say if the supply of securities is flexible, then the change in price would be influenced by the expectation adjustment in dividends of all other stocks as well as that in its own dividend.

Therefore, two hypotheses related to supply effect to be tested regarding the parameters in the reduced form system shown in Eq. 93.18 are as follows:

Hypothesis 1: All the off-diagonal elements in the coefficient matrix Π are zero if the supply effect does not exist.

Hypothesis 2: All the diagonal elements in the coefficients matrix Π are equal in the magnitude if the supply effect does not exist.

These two hypotheses should be satisfied jointly. That is, if the supply effect does not exist, price changes of a security should be a function of its own dividend expectation adjustments, and the coefficients should all be equal across securities. In the model described in Eq. 93.16, if an investor expects a change in the prediction of the next dividend due to the additional information during the current period, then the price of the security changes.

Under the assumption of the efficiency in the domestic stock market, if the supply of securities is fixed, then the expectation errors in dividends are built in the current price is the same for all securities. This phenomenon implies that the price changes would only be influenced by its own dividend expectation adjustments. If the supply of securities is flexible, then the change in price would be influenced by the expectation adjustment in dividends of all other securities as well as that of its own dividend.

93.3 Data and Empirical Results

In this section, we derive the test by analyzing the US domestic stock market. Most details of the model, the methodologies, and the hypotheses for empirical tests are previously discussed in Sect. 93.2. However, before testing the hypotheses, some other details of the related tests that are needed to support the assumptions used in the model are also briefly discussed in this section.

This section examines the hypotheses derived earlier for the US domestic stock market by using the companies listed in S&P 500 and, then, by using the companies listing in Dow Jones Index. If the supply of risky assets is responsive to its price, then large price changes, which are due to the change in expectation of future dividend, will be spread over time. In other words, there exists supply effect in the US domestic stock markets. This implies that the dynamic instead of static CAPM should be used for testing capital assets pricing in the equity markets of the United States.

93.3.1 Data and Descriptive Statistics

Three hundred companies are selected from the S&P 500 and grouped into ten portfolios with equal numbers of 30 companies by their payout ratios. The data are obtained from the Compustat North America industrial quarterly data. The data starts from the first quarter of 1981 to the last quarter of 2002. The companies selected satisfy the following two criteria. First, the company appears on the S&P 500 at some time period during 1981 through 2002. Second, the company must have complete data available – including price, dividend, earnings per share, and shares outstanding – during the 88 quarters (22 years). Firms are eliminated from the sample list if one of the following two conditions occurs:

(i) Reported earnings are either trivial or negative.
(ii) Reported dividends are trivial.

Three hundred fourteen firms remain after these adjustments. Finally, excluding those seven companies with highest and lowest average payout ratio, the remaining 300 firms are grouped into ten portfolios by the payout ratio. Each portfolio contains 30 companies. Figure 93.1 shows the comparison of S&P 500 index and the value-weighted index of the 300 firms selected (M). Figure 93.1 shows that the trend is similar to each other before the third quarter of 1999. However, there exist some differences after third quarter of 1999.

To group these 300 firms, the payout ratio for each firm in each year is determined by dividing the sum of four quarters' dividends by the sum of four quarters' earnings; then, the yearly ratios are further averaged over the 22-year period. The first 30 firms with highest payout ratio comprise portfolio one, and so on. Then, the value-weighted average of the price, dividend, and earnings of each portfolio is computed. Characteristics and summary statistics of these ten portfolios are presented in Tables 93.1 and 93.2, respectively. Table 93.1 presents information of return, payout ratio, size, and beta for ten portfolios. From the results of this table, there appears to exist an inverse relationship between return and payout ratio, payout ratio and beta. However, the relationship between payout ratio and beta is not so clear. This finding is similar to that of Fama and French (1992).

Table 93.2 shows the first four moments of quarterly returns of the market portfolio and ten portfolios. The coefficients of skewness, kurtosis, and Jarque-Bera statistics show that one cannot reject the hypothesis that log return of most

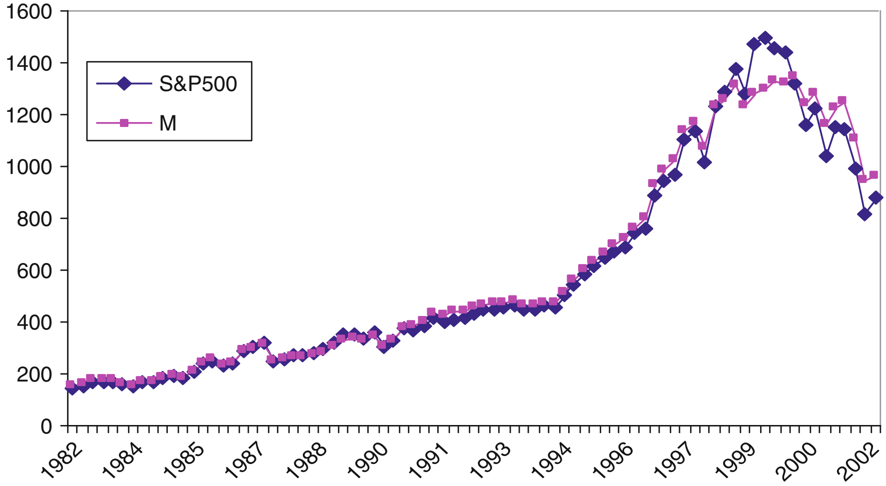


Fig. 93.1 Comparison of S&P 500 and market portfolio

Table 93.1 Characteristics of ten portfolios

Portfolio[a]	Return[b]	Payout[c]	Size (000)	Beta (M)
1	0.0351	0.7831	193,051	0.7028
2	0.0316	0.7372	358,168	0.8878
3	0.0381	0.5700	332,240	0.8776
4	0.0343	0.5522	141,496	1.0541
5	0.0410	0.5025	475,874	1.1481
6	0.0362	0.4578	267,429	1.0545
7	0.0431	0.3944	196,265	1.1850
8	0.0336	0.3593	243,459	1.0092
9	0.0382	0.2907	211,769	0.9487
10	0.0454	0.1381	284,600	1.1007

[a]The first 30 firms with highest payout ratio comprise portfolio one, and so on
[b]The price, dividend, and earnings of each portfolio are computed by value-weighted of the 30 firms included in the same category
[c]The payout ratio for each firm in each year is found by dividing the sum of four quarters' dividends by the sum of four quarters' earnings; then, the yearly ratios are then computed from the quarterly data over the 22-year period

portfolios is normal. The kurtosis statistics for most sample portfolios are close to three, which indicates that heavy tails are not an issue. Additionally, Jarque-Bera coefficients illustrate that the hypotheses of Gaussian distribution for most portfolios are not rejected. It seems to be unnecessary to consider the problem of heteroskedasticity in estimating domestic stock market if the quarterly data are used.

Table 93.2 Summary statistics of portfolio quarterly returns[a]

Country	Mean (quarterly)	Std. dev. (quarterly)	Skewness	Kurtosis	Jarque-Bera
Market portfolio	0.0364	0.0710	−0.4604	3.9742	6.5142*
Portfolio 1	0.0351	0.0683	−0.5612	3.8010	6.8925*
Portfolio 2	0.0316	0.0766	−1.1123	5.5480	41.470**
Portfolio 3	0.0381	0.0768	−0.3302	2.8459	1.6672*
Portfolio 4	0.0343	0.0853	−0.1320	3.3064	0.5928
Portfolio 5	0.0410	0.0876	−0.4370	3.8062	5.1251
Portfolio 6	0.0362	0.0837	−0.2638	3.6861	2.7153
Portfolio 7	0.0431	0.0919	−0.1902	3.3274	0.9132
Portfolio 8	0.0336	0.0906	0.2798	3.3290	1.5276
Portfolio 9	0.0382	0.0791	−0.2949	3.8571	3.9236
Portfolio 10	0.0454	0.0985	−0.0154	2.8371	0.0996

[a]Quarterly returns from 1981:Q1to 2002:Q4 are calculated

* and ** denote statistical significance at the 5 % and 1 % level, respectively

93.3.2 Dynamic CAPM with Supply Side Effect

If one believes that the stock market is efficient (i.e., if one believes the way in which the expectation errors in dividends are built in the current price is the same for all securities), then price changes would be influenced only by its own dividend expectation errors. Otherwise, if the supply of securities is flexible, then the change in price would be influenced by the expectation adjustment in dividends of other portfolios as well as that in its own dividend. Thus, two hypotheses related to supply effect are to be tested and should be satisfied jointly in order to examine whether there exists a supply effect.

Recalling from the previous section, the structural form equations are exactly identified, and the series of expectation adjustments in dividend, d_t, are exogenous variables (d_t can be estimated from earnings per share and dividends per share by using a partial adjustment model as presented in Appendix 3). Now, the reduced form equations can be used to test the supply effect. That is, Eq. 93.22 needs to be examined by the following hypotheses:

Hypothesis 1: All the off-diagonal elements in the coefficient matrix Π are zero if the supply effect does not exist.

Hypothesis 2: All the diagonal elements in the coefficients matrix Π are equal in the magnitude if the supply effect does not exist.

These two hypotheses should be satisfied jointly. That is, if the supply effect does not exist, price changes of each portfolio would be a function of its own dividend expectation adjustments, and the coefficients should be equal across all portfolios.

The estimated coefficients of the simultaneous equation system for ten portfolios are summarized in Table 93.3.[7] Results of Table 93.3 indicate that the estimated

[7]The results are similar when using either the FIML or SUR approach. We report here the estimates of the SUR method.

Table 93.3 Coefficients for matrix Π' (ten portfolios)[a]

	P_P1	P_P2	P_P3	P_P4	P_P5	P_P6	P_P7	P_P8	P_P9	P_P10
P1	15.57183	−12.60844	13.15747	−8.58455	8.62495	−2.486287	10.48123	1.959701	−1.274653	−13.4239
	−23.5688	−24.513	−22.2507	−22.3461	−25.6377	−26.1305	−24.906	−24.181	−16.358	−29.1236
	[0.6607]	[−0.5144]	[0.5913]	[−0.3842]	[0.3364]	[−0.0952]	[0.4208]	[0.0810]	[−0.0779]	[−0.4609]
P2	−16.67868	−18.7728	−24.73303	−12.19542	−18.61126	5.326864	−16.99283	−5.675232	−1.795597	13.98581
	−14.2287	−14.7988	−13.433	−13.4906	−15.4778	−15.7753	−15.036	−14.5984	−9.8755	−17.5823
	[−1.1722]	[−1.2685]	[−1.8412]	[−0.9040]	[−1.2025]	[0.3377]	[−1.1301]	[−0.3888]	[−0.1818]	[0.7955]
P3	140.7762	117.8989	180.973	128.0238	161.9093	44.47442	115.7946	103.2686	74.30349	74.72393
	−73.6813	−76.6333	−69.5607	−69.8588	−80.1493	−81.69	−77.8617	−75.5953	−51.1387	−91.047
	[1.9106]	[1.5385]	[2.6017]	[1.8326]	[2.0201]	[0.5444]	[1.4872]	[1.3661]	[1.4530]	[0.8207]
P4	−79.569	−82.9826	−16.2607	−71.5316	−38.36708	−29.88297	−43.8957	−20.7400	−10.4372	−2.02316
	−64.5317	−67.1171	−60.9228	−61.1839	−70.1966	−71.5459	−68.193	−66.208	−44.7884	−79.741
	[−1.2330]	[−1.2364]	[−0.2669]	[−1.1691]	[−0.5466]	[−0.4177]	[−0.6437]	[−0.3133]	[−0.2330]	[−0.0254]
P5	25.63953	29.0526	54.39686	6.087413	31.12653	7.582502	30.88937	19.3122	17.58315	−0.01716
	−25.521	−26.5435	−24.0937	−24.197	−27.7613	−28.2949	−26.9689	−26.1839	−17.7129	−31.5359
	[1.0047]	[1.0945]	[2.2577]	[0.2516]	[1.1212]	[0.2680]	[1.1454]	[0.7376]	[0.9927]	[−0.0005]
P6	−12.46593	−8.734942	−45.85208	−25.53128	−17.06422	−18.11443	−23.51969	−1.723033	−4.492465	−31.53814
	−12.1881	−12.6764	−11.5065	−11.5558	−13.2581	−13.5129	−12.8796	−12.5047	−8.45921	−15.0607
	[−1.0228]	[−0.6891]	[−3.9849]	[−2.2094]	[−1.2871]	[−1.3405]	[−1.8261]	[−0.1378]	[−0.5311]	[−2.0941]
P7	−84.5262	−35.03964	−114.7987	−19.48548	−97.9274	4.402397	−57.69584	−58.88397	−68.04914	3.566607
	−56.1062	−58.354	−52.9685	−53.1955	−61.0314	−62.2046	−59.2894	−57.5636	−38.9406	−69.3296
	[−1.5065]	[−0.6005]	[−2.1673]	[−0.3663]	[−1.6045]	[0.0708]	[−0.9731]	[−1.0229]	[−1.7475]	[0.0514]
P8	−5.497057	−4.463256	−31.77293	29.38345	−8.488357	0.394223	−21.59846	−45.72339	19.80597	−107.4715
	−62.0465	−64.5323	−58.5765	−58.8276	−67.4932	−68.7905	−65.5667	−63.6582	−43.0635	−76.67
	[−0.0886]	[−0.0692]	[−0.5424]	[0.4995]	[−0.1258]	[0.0057]	[−0.3294]	[−0.7183]	[0.4599]	[−1.4017]

(*continued*)

Table 93.3 (continued)

	P_P1	P_P2	P_P3	P_P4	P_P5	P_P6	P_P7	P_P8	P_P9	P_P10
P9	20.70817	28.77904	15.61156	23.14069	25.93932	35.08121	23.73591	15.46799	18.15523	25.27915
	−15.5463	−16.1691	−14.6768	−14.7398	−16.911	−17.236	−16.4283	−15.9501	−10.7899	−19.2103
	[1.3320]	[1.7799]	[1.0637]	[1.5700]	[1.5339]	[2.0353]	[1.4448]	[0.9698]	[1.6826]	[1.3159]
P10	−14.64016	−51.1797	−49.51991	−64.67943	−23.53575	67.38674	7.053653	−30.23067	−15.54273	36.60222
	−112.584	−117.094	−106.288	−106.743	−122.467	−124.821	−118.971	−115.508	−78.1391	−139.118
	[−0.1300]	[−0.4371]	[−0.4659]	[−0.6059]	[−0.1922]	[0.5399]	[0.0593]	[−0.2617]	[−0.1989]	[0.2631]
R^2	0.083841	0.096546	0.283079	0.134377	0.088212	0.075947	0.091492	0.027763	0.065971	0.138979
F-st	0.772786	0.902404	3.334318	1.310894	0.816966	0.694038	0.850408	0.241141	0.596435	1.363029

[a]Standard errors in ( ) and t-statistics in [ ]

diagonal elements seem to vary across portfolios and most of the off-diagonal elements are significant from zero. However, simply observing the elements in matrix Π directly cannot justify either accept or reject the null hypotheses derived for testing the supply effect. Two tests should be done separately to check whether these two hypotheses are both satisfied.

For the first hypothesis, the test of supply effect on off-diagonal elements, the following regression in accordance with Eq. 93.22 is run for each portfolio:

$$p_{i,t} = \beta_i d_{i,t} + \Sigma_{j \neq i} \beta_j d_{j,t} + \varepsilon_{i,t}, \mathrm{i}, \mathrm{j} = 1, \ldots, 10. \tag{93.23}$$

The null hypothesis then can be written as H_0: $\beta_j = 0$, $j = 1, \ldots, 10$, $j \neq i$. The results are reported in Table 93.4. Two test statistics are reported. The first test uses an F distribution with 9 and 76 degrees of freedom, and the second test uses a chi-squared distribution with 9 degrees of freedom. The null hypothesis is rejected at 5 % significance level in six out of ten portfolios, and only two portfolios cannot be rejected at 10 % significance level. This result indicates that the null hypothesis can be rejected at conventional levels of significance.

For the second hypothesis of supply effect on all diagonal elements of Eq. 93.22, the following null hypothesis needs to be tested:

$$H_0 : \pi_{i,i} = \pi_{j,j} \quad \text{for all } i, j = 1, \ldots, 10.$$

To do this null hypothesis test, we need to estimate Eq. 93.22 simultaneously, and then, we calculate Wald statistics by imposing nine restrictions on this equation system. Under the above nine restrictions, the Wald test statistic has a chi-square distribution with 9 degrees of freedom. The statistic is 18.858, which is greater than 16.92 at 5 % significance level. Since the statistic corresponds to a p-value of 0.0265, one can reject the null hypothesis at 5 %, but it cannot reject H_0 at a 1 % significance level. In other words, the diagonal elements are not similar to each other in magnitude. In conclusion, the above empirical results are sufficient to reject two null hypotheses of nonexistence of supply effect in the US stock market.

In order to check whether the individual stocks can hold up to the same testing, we use individual stock data as many as 30 companies in one group. The results are summarized in Table 93.5. From Table 93.5, we find that the above conclusion seems to be sustainable if we use individual stock data. More specifically, the diagonal elements are not equal to each other at any conventional significant level and the off-diagonal elements are significantly from zero in each group composed of 30 individual stocks.

We also find that one cannot reject the existence of supply effect by using the stocks listed in the Dow Jones Index. Again, to test the supply effect on off-diagonal elements, Eq. 93.23 is run as the following for each company:

$$p_{i,t} = \beta_i d_{i,t} + \Sigma_{j \neq i} \beta_j d_{j,t} + \varepsilon_{i,t}, \quad \mathrm{i}, \mathrm{j} = 1, \ldots, 29. \tag{93.23$'$}$$

Table 93.4 Test of supply effect on off-diagonal elements of matrix $\Pi^{a,b}$

	R^2	F- statistic	*p*-value	Chi-square	*p*-value
Portfolio 1	0.1518	1.7392	0.0872	17.392*	0.0661
Portfolio 2	0.1308	1.4261	0.1852	14.261	0.1614
Portfolio 3	0.4095	5.4896	0.0000	53.896***	0.0000
Portfolio 4	0.1535	1.9240	0.0607	17.316**	0.0440
Portfolio 5	0.1706	1.9511	0.0509	19.511**	0.0342
Portfolio 6	0.2009	1.2094	0.2988	12.094	0.2788
Portfolio 7	0.2021	1.8161	0.0718	18.161*	0.0523
Portfolio 8	0.1849	1.9599	0.0497	19.599**	0.0333
Portfolio 9	0.1561	1.8730	0.0622	18.730**	0.0438
Portfolio 10	0.3041	3.5331	0.0007	35.331***	0.0001

[a] $p_{i,\,t} = \beta_i{}' d_{i,\,t} + \Sigma_{j \neq i}\beta_j{}' d_{j,\,t} + \varepsilon'_{i,\,t}$ i, $j = 1, \ldots, 10$.
Hypothesis: all $\beta_j = 0, j = 1, \ldots, 10, j \neq i$
[b] The first test uses an F distribution with 9 and 76 degrees of freedom, and the second uses a chi-squared distribution with 9 degrees of freedom
*, **, and *** denote statistical significance at the 10 %, 5 %, and 1 % level, respectively

The null hypothesis then can be written as H_0: $\beta_j = 0$, $j = 1, \ldots, 29$, $j \neq i$. The results are summarized in Table 93.6. The null hypothesis is rejected at 1 % significance level in 26 out of 29 companies. For the second hypothesis of supply effect on all diagonal elements, the following null hypothesis is also tested: H_0: $\pi_{i,\,i} = \pi_{j,\,j}$, for all i, j = 1, . . ., 29.

The Wald test statistic has a chi-square distribution with 28 degrees of freedom. The statistic is 86.35. That is, one can reject this null hypothesis at 1 % significance level.

93.4 Summary

We examine an asset pricing model that incorporates a firm's decision concerning the supply of risky securities into the CAPM. This model focuses on a firm's financing decision by explicitly introducing the firm's supply of risky securities into the static CAPM and allows the supply of risky securities to be a function of security price. And thus, the expected returns are endogenously determined by both demand and supply decisions within the model. In other words, the supply effect may be one important factor in capital assets pricing decisions.

Our objective is to investigate the existence of supply effect in the US stock markets. We find that supply effect is important in the US stock market. This finding holds as we break the companies listed in the S&P 500 into ten portfolios. It also holds if we use individual stock data. These test results show that two null hypotheses of the nonexistence of supply effect do not seem to be satisfied jointly. In other words, this evidence seems to be sufficient to support the existence of supply effect and, thus, imply a violation of the assumption in the one-period static CAPM, or to imply a dynamic asset pricing model may be a better choice in the US domestic stock markets.

Table 93.5 Test of supply effect (by individual stock)

	Test of supply effect on the diagonal elements: H_0: $\pi_{ii} = \pi_{jj}$ for all i, $j = 1, 2, \ldots, 30$	Test supply effect on off-diagonal elements: $p_{i,t} = \beta_i{}' d_{i,t} + \Sigma_{j \neq i}\beta_j{}' d_{j,t} + \varepsilon'_{i,t}$, for i, $j = 1, 2, \ldots, 30$; H_0: all $\beta_j = 0$, $j = 1, 2, \ldots, 30$, $j \neq i$ Different significant level 1 %	5 %	10 %
Group 1	$\chi^2 = 113.65$, p-value = 0.0000 ➔ Reject H_0 at 1 %	Reject 23 in 30 equations	Reject 25 in 30 equations	Reject 25 in 30 equations
Group 2	$\chi^2 = 52.08$, p-value = 0.0053 ➔ Reject H_0 at 1 %	Reject 21 in 30 equations	Reject 24 in 30 equations	Reject 25 in 30 equations
Group 3	$\chi^2 = 86.53$, p-value = 0.0000 ➔ Reject H_0 at 1 %	Reject 26 in 30 equations	Reject 27 in 30 equations	Reject 28 in 30 equations
Group 4	$\chi^2 = 88.58$, p-value = 0.0000 ➔ Reject H_0 at 1 %	Reject 21 in 30 equations	Reject 24 in 30 equations	Reject 25 in 30 equations
Group 5	$\chi^2 = 101.14$, p-value = 0.0000 ➔ Reject H_0 at 1 %	Reject 25 in 30 equations	Reject 26 in 30 equations	Reject 28 in 30 equations
Group 6	$\chi^2 = 69.14$, p-value = 0.0000 ➔ Reject H_0 at 1 %	Reject 17 in 30 equations	Reject 21 in 30 equations	Reject 22 in 30 equations
Group 7	$\chi^2 = 181.10$, p-value = 0.0000 ➔ Reject H_0 at 1 %	Reject 29 in 30 equations	Reject 30 in 30 equations	Reject 30 in 30 equations
Group 8	$\chi^2 = 116.97$, p-value = 0.0000 ➔ Reject H_0 at 1 %	Reject 29 in 30 equations	Reject 29 in 30 equations	Reject 29 in 30 equations
Group 9	$\chi^2 = 117.44$, p-value = 0.0000 ➔ Reject H_0 at 1 %	Reject 27 in 30 equations	Reject 28 in 30 equations	Reject 29 in 30 equations
Group 10	$\chi^2 = 109.50$, p-value = 0.0000 ➔ Reject H_0 at 1 %	Reject 25 in 30 equations	Reject 27 in 30 equations	Reject 27 in 30 equations

For the future research, we will first modify the simultaneous equation asset pricing model defined in Eqs. 93.9 and 93.12 to allow for testing the existence of market disequilibrium in dynamic asset pricing. Then, we will use disequilibrium estimation methods developed by Amemiya (1974), Fair and Jaffe(1972), and Quandt (1988) to test whether there is price adjustment in response to an excess demand in equity market.

Appendix 1: Modeling the Price Process

In Sect. 93.2.3, Eq. 93.16 is derived from Eq. 93.15 under the assumption that all countries' index series follow a random walk process. Thus, before further discussion, we should test the order of integration of these price series. From Hamilton (1994), we know that two widely used unit root tests are the Dickey-Fuller (DF) and the augmented Dickey-Fuller (ADF) tests. The former can be represented as $P_t = \mu + \gamma P_{t-1} + \varepsilon_t$, and the latter can be written as $\Delta P_t = \mu + \gamma\, P_{t-1} + \delta_1 \Delta P_{t-1} + \delta_2 \Delta P_{t-2} + \ldots + \delta_p \Delta P_{t-p} + \varepsilon_t$.

Table 93.6 Test of supply effect (companies listed in the Dow Jones Index)

GVKEY	Security i	R^2 of each equation i	H_0: all $\beta_j = 0, j = 1, 2, \ldots, 29, j \neq i$ Chi-square	p-value
1300	Honeywell International Inc	0.7088	137.08	0.0000
1356	Alcoa Inc	0.6716	120.84	0.0000
1447	American Express	0.4799	55.47	0.0015
1581	AT&T Corp	0.5980	56.16	0.0012
2285	Boeing Co	0.5291	66.75	0.0001
2817	Caterpillar Inc	0.5887	83.10	0.0000
2968	JPMorgan Chase & Co	0.5352	68.12	0.0000
3144	Coca-Cola Co	0.5927	87.04	0.0000
3243	Citigroup Inc	0.6082	88.63	0.0000
3980	Disney (Walt) Co	0.6457	104.06	0.0000
4087	Du Pont (E I) De Nemours	0.6231	98.37	0.0000
4194	Eastman Kodak Co	0.3793	36.14	0.1416
4503	Exxon Mobil Corp	0.5653	76.50	0.0000
5047	General Electric Co	0.5425	61.17	0.0003
5073	General Motors Corp	0.5372	66.73	0.0001
5606	Hewlett-Packard Co	0.4755	53.61	0.0025
5680	Home Depot Inc	0.6753	106.00	0.0000
6008	Intel Corp	0.5174	60.05	0.0004
6066	Intl Business Machines Corp	0.5596	75.31	0.0000
6104	Intl Paper Co	0.5512	72.58	0.0000
6266	Johnson & Johnson	0.5211	67.59	0.0000
7154	McDonalds Corp	0.4416	45.53	0.0195
7257	Merck & Co	0.4109	40.82	0.0558
7435	3M CO	0.6344	105.07	0.0000
8543	Altria Group Inc	0.5751	72.25	0.0000
8762	Procter & Gamble Co	0.5816	84.19	0.0000
9899	SBC Communications Inc	0.5486	72.81	0.0000
10983	United Technologies Corp	0.6595	116.19	0.0000
11259	Wal-Mart Stores	0.6488	111.85	0.0000

Test the off-diagonal elements: $p_{i,t} = \beta_i{}' d_{i,t} + \Sigma_{j \neq i} \beta_j{}' d_{j,t} + \varepsilon'_{i,t}$, for i, $j = 1, \ldots, 29$, null hypothesis H_0: all $\beta_j = 0, j = 1, 2, \ldots, 29, j \neq i$

Test of supply effect on the diagonal elements; H_0: $\pi_{ii} = \pi_{jj}$ for all i, $j = 1, 2, \ldots, 29$ Result: $\chi^2 = 86.35$, p-value $= 0.0000 \rightarrow$ Reject H_0 at 1 %

Microsoft Corp. is not included since it had paid dividends twice for the whole sample period

Similarly, in the US stock markets, the Phillips-Perron test is used to check whether the value-weighted price of market portfolio follows a random walk process. The results of the tests for each index are summarized in Table 93.7. It seems that one cannot reject the hypothesis that all indices follow a random walk process since, for example, the null hypothesis of unit root in level cannot be rejected for all indices but are all rejected if one assumes there is a unit root in the first-order difference of the price for each portfolio. This result is consistent with most studies that find that the financial price series follow a random walk process.

Table 93.7 Unit root tests for P_t

	$P_t = \mu + \gamma P_{t-1} + \varepsilon_t$		Phillips-Perron test[a]	
	Estimated c_2 (std. error)	Adj. R^2	Level	1st difference[b]
Market portfolio	1.0060 (0.0159)	0.9788	−0.52	−8.48**
S&P 500	0.9864 (0.0164)	0.9769	−0.90	−959**
Portfolio 1	0.9883 (0.0172)	0.9746	−0.56	−8.67**
Portfolio 2	0.9877 (0.0146)	0.9815	−0.97	−9.42**
Portfolio 3	0.9913 (0.0149)	0.9809	−0.51	−13.90**
Portfolio 4	0.9935 (0.0143)	0.9825	−0.61	−7.66**
Portfolio 5	0.9933 (0.0158)	0.9787	−0.43	−9.34**
Portfolio 6	0.9950 (0.0150)	0.9808	−0.32	−8.66**
Portfolio 7	0.9892 (0.0155)	0.9793	−0.64	−9.08**
Portfolio 8	0.9879 (0.0166)	0.9762	−0.74	−9.37**
Portfolio 9	0.9939 (0.0116)	0.9884	−0.74	−7.04**
Portfolio 10	0.9889 (0.0182)	0.9716	−0.69	−9.07**

*5 % significant level; ** 1 % significant level
[a]The process assumed to be random walk without drift
[b]The null hypothesis of zero intercept terms, μ, cannot be rejected at 5 %, 1 % level for all portfolios

Appendix 2: Identification of the Simultaneous Equation System

Note that given G is nonsingular, $\Pi = -G^{-1}H$ in Eq. 93.19 can be written as

$$\text{where } A = [G \quad H] = \begin{pmatrix} g_{11} & g_{12} & \cdots\cdots & g_{1n} & h_{11} & h_{12} & \cdots\cdots & h_{1n} \\ g_{21} & g_{22} & \cdots\cdots & g_{1n} & h_{21} & h_{22} & \cdots\cdots & h_{2n} \\ . & & & & & & . & \\ . & & & & & & . & \\ g_{n1} & g_{n2} & \cdots\cdots & g_{nn} & h_{n1} & h_{n2} & \cdots\cdots & h_{nn} \end{pmatrix}$$

$$W = [\Pi \quad I_n]' = \begin{pmatrix} \pi_{11} & \pi_{12} & \cdots & \pi_{1n} & 1 & 0 & \cdots\cdots & 0 \\ \pi_{21} & \pi_{22} & \cdots & \pi_{1n} & 0 & 1 & \cdots\cdots & 0 \\ . & & & & & & . & \\ . & & & & & & . & \\ \pi_{n1} & \pi_{n2} & \cdots & \pi_{nn} & 0 & 0 & \cdots\cdots & 1 \end{pmatrix} \tag{93.24}$$

That is, A is the matrix of all structure coefficients in the model with dimension of $(n \times 2n)$, and W is a $(2n \times n)$ matrix. The first equation in Eq. (93.24) can be expressed as

$$A_1 W = 0, \tag{93.25}$$

where A_1 is the first row of A, i.e., $A_1 = [g_{11}\ g_{12} \ldots.\ g_{1n}\ h_{11}\ h_{12} \ldots..\ h_{1n}]$.

Since the elements of Π can be consistently estimated, and I_n is the identity matrix, Eq. 93.25 contains 2n unknowns in terms of πs. Thus, there should be n restrictions on the parameters to solve Eq. 93.25 uniquely. First, one can try to impose normalization rule by setting g_{11} equal to 1 to reduce one restriction. As a result, there are at least n−1 independent restrictions needed in order to solve Eq. 93.25.

It can be illustrated that the system represented by Eq. 93.17 is exactly identified with three endogenous and three exogenous variables. It is entirely similar to those cases of more variables. For example, if $n = 3$, Eq. 93.17 can be expressed in the form

$$\begin{aligned}&-\begin{pmatrix} r*cs_{11}+a_1b_1 & r*cs_{12} & r*cs_{13} \\ r*cs_{21} & r*cs_{22}+a_2b_2 & r*cs_{23} \\ r*cs_{31} & r*cs_{32} & r*cs_{33}+a_3b_3 \end{pmatrix}\begin{pmatrix} p_{1t} \\ p_{2t} \\ p_{3t} \end{pmatrix} \\ &+\begin{pmatrix} cs_{11}+a_1 & cs_{12} & cs_{13} \\ cs_{21} & cs_{22}+a_2 & cs_{23} \\ cs_{31} & cs_{32} & cs_{33}+a_3 \end{pmatrix}\begin{pmatrix} d_{1t} \\ d_{2t} \\ d_{3t} \end{pmatrix} = \begin{pmatrix} v_{1t} \\ v_{2t} \\ v_{3t} \end{pmatrix}\end{aligned} \tag{93.26}$$

where
r^* = scalar of risk-free rate
s_{ij} = elements of variance-covariance matrix of return
a_i = inverse of the supply adjustment cost of firm i
b_i = overall cost of capital of firm i

For example, in the case of $n = 3$, Eq. 93.17 can be written as

$$\begin{pmatrix} g_{11} & g_{12} & g_{13} \\ g_{21} & g_{22} & g_{23} \\ g_{31} & g_{32} & g_{33} \end{pmatrix}\begin{pmatrix} p_{1t} \\ p_{2t} \\ p_{3t} \end{pmatrix} + \begin{pmatrix} h_{11} & h_{12} & h_{13} \\ h_{21} & h_{22} & h_{23} \\ h_{31} & h_{32} & h_{33} \end{pmatrix}\begin{pmatrix} d_{1t} \\ d_{2t} \\ d_{3t} \end{pmatrix} = \begin{pmatrix} v_{1t} \\ v_{2t} \\ v_{3t} \end{pmatrix}. \tag{93.27}$$

Comparing Eq. 93.26 with Eq. 93.27, the prior restrictions on the first equation take the form $g_{12} = -r^*h_{12}$ and $g_{13} = -r^*h_{13}$ and so on.

Thus, the restriction matrix for the first equation is of the form

$$\Phi = \begin{pmatrix} 0 & 1 & 0 & 0 & r^* & 0 \\ 0 & 0 & 1 & 0 & 0 & r^* \end{pmatrix} \tag{93.28}$$

Then, combining Eq. 93.25 and the parameters of the first equation gives

$$[g_{11} \quad g_{12} \quad g_{13} \quad h_{11} \quad h_{12} \quad h_{13}]\begin{pmatrix} \pi_{11} & \pi_{12} & \pi_{13} & 0 & 0 \\ \pi_{21} & \pi_{22} & \pi_{13} & 1 & 0 \\ \pi_{31} & \pi_{32} & \pi_{33} & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & r* & 0 \\ 0 & 0 & 1 & 0 & r* \end{pmatrix}[0 \;\; 0 \;\; 0 \;\; 0 \;\; 0]. \tag{93.29}$$

That is, extending Eq. 93.29, we have

$$\begin{aligned} &g_{11}\pi_{11} + g_{12}\pi_{21} + g_{13}\pi_{31} + h_{11} = 0,\\ &g_{11}\pi_{12} + g_{12}\pi_{22} + g_{13}\pi_{32} + h_{12} = 0,\\ &g_{11}\pi_{13} + g_{12}\pi_{23} + g_{13}\pi_{33} + h_{13} = 0,\\ &g_{12} + r^*h_{12} = 0, \text{and}\\ &g_{13} + r^*h_{13} = 0. \end{aligned} \tag{93.30}$$

The last two ($n-1 = 3-1 = 2$) equations in Eq. 93.30 give the value h_{12} and h_{13}, and the normalization condition, $g_{11} = 1$, allows us to solve Eq. 93.25 in terms of πs uniquely. That is, in the case $n = 3$, the first equation represented by Eq. 93.25, $A_1W = 0$, can be finally rewritten as Eq. 93.30. Since there are three unknowns, g_{12}, g_{13}, and h_{11}, left for the first three equations in Eq. 93.30, the first equation A_1 is exactly identified. Similarly, it can be shown that the second and the third equations are also exactly identified.

Appendix 3: Derivation of the Formula Used to Estimate d_t

To derive the formula for estimating d_t, we first define the partial adjustment model as

$$D_t = a_1 + a_2D_{t-1} + a_3E_t + u_t \tag{93.31}$$

where D_t = dividend per share in period t, D_{t-1} = dividend per share in period t−1, E_t = earnings per share in period t are dividends and earnings, and u_t = error term in period t. Similarly,

$$D_{t+1} = a_1 + a_2D_t + a_3E_{t+1} + u_{t+1}. \tag{93.31$'$}$$

And thus,

$$\mathrm{E}_{t-1}[D_t] = a_1 + a_2D_{t-1} + a_3\mathrm{E}_{t-1}[E_t], \tag{93.32}$$

$$\mathrm{E}_t[D_{t+1}] = a_1 + a_2D_t + a_3\mathrm{E}_t[E_{t+1}], \tag{93.33}$$

$$\mathrm{E}_{t-1}[D_{t+1}] = a_1 + a_2\mathrm{E}_{t-1}[D_t] + a_3\mathrm{E}_{t-1}[E_{t+1}]. \tag{93.34}$$

Substituting Eq. 93.32 to Eq. 93.34, we have

$$E_{t-1}[D_{t+1}] = a_1 + a_1a_2 + a_2^2D_{t-1} + a_2a_3\mathrm{E}_{t-1}[E_t] + a_3\mathrm{E}_{t-1}[E_{t+1}]. \tag{93.34$'$}$$

Subtracting Eq. 93.34′ from Eq. 93.33 on both hand sides, we have

$$\begin{aligned} \mathrm{E}_t[D_{t+1}] - \mathrm{E}_{t-1}[D_{t+1}] = &-a_1a_2 + a_2D_t - a_2^2D_{t-1} - a_2a_3\mathrm{E}_{t-1}[E_t]\\ &+ a_3\mathrm{E}_t[E_{t+1}] - a_3\mathrm{E}_{t-1}[E_{t+1}]. \end{aligned} \tag{93.35}$$

Equation Eq. 93.35 can be investigated depending upon whether E_t is following a random walk.

Case 1 E_t follows an AR(p) process.

If the time series of E_t is assumed to be stationary and follows an AR(p) process, then after taking the seasonal differences, we obtain

$$dE_t = \rho_0 + \rho_1 dE_{t-1} + \rho_2 dE_{t-2} + \rho_3 dE_{t-3} + \rho_4 dE_{t-4} + \varepsilon_t, \tag{93.36}$$

where $dE_t = E_t - E_{t-4}$.

The expectation adjustment in seasonally differenced earnings, or the revision in forecasting future seasonally differenced earnings, can be solved as

$$\mathrm{E}_t[dE_{t+1}] - \mathrm{E}_{t-1}[dE_{t+1}] = \rho_1(-\rho_0 + dE_t - \rho_1 dE_{t-1} - \rho_2 dE_{t-2} - \rho_3 dE_{t-3} - \rho_4 dE_{t-4}). \tag{93.37}$$

Since $\mathrm{E}_t[dE_{t+1}] - \mathrm{E}_{t-1}[dE_{t+1}] = \mathrm{E}_t[E_{t+1}] - \mathrm{E}_{t-1}[E_{t+1}]$, we have

$$\mathrm{E}_t[E_{t+1}] - \mathrm{E}_{t-1}[E_{t+1}] = \rho_1(-\rho_0 + dE_t - \rho_1 dE_{t-1} - \rho_2 dE_{t-2} - \rho_3 dE_{t-3} - \rho_4 dE_{t-4}). \tag{93.38}$$

Furthermore, from Eq. 93.36, we have

$$\mathrm{E}_{t-1}[dE_t] = \rho_0 + \rho_1 dE_{t-1} + \rho_2 dE_{t-2} + \rho_3 dE_{t-3} + \rho_4 dE_{t-4}. \tag{93.39}$$

Similarly, $\mathrm{E}_{t-1}[dE_t] = \mathrm{E}_{t-1}[E_1 - E_{t-4}] = \mathrm{E}_{t-1}[E_t] - E_{t-4}$; thus, $\mathrm{E}_{t-1}[E_t]$ can be found by

$$\mathrm{E}_{t-1}[E_t] = \rho_0 + \rho_1(dE_{t-1}) + \rho_2(dE_{t-2}) + \rho_3(dE_{t-3}) + \rho_4(dE_{t-4}) + E_{t-4}. \tag{93.40}$$

Finally, the expectation adjustment in dividends, d_t, can be found by plugging Eqs. 93.38 and 93.40 into Eq. 93.35:

$$\begin{aligned} d_t \equiv \mathrm{E}_t[D_{t+1}] - \mathrm{E}_{t-1}[D_{t+1}] &= -a_1 a_2 + a_2 D_t - a_2^2 D_{t-1} \\ &\quad - a_2 a_3(\rho_0 + \rho_1 dE_{t-1} + \rho_2 dE_{t-2} + \rho_3 dE_{t-3} + \rho_4 dE_{t-4} + E_{t-4}) \\ &\quad + a_3 \rho_1(-\rho_0 + dE_t - \rho_1 dE_{t-1} - \rho_2 dE_{t-2} - \rho_3 dE_{t-3} - \rho_4 dE_{t-4}) \end{aligned} \tag{93.41}$$

Or

$$\begin{aligned} d_t = C_0 + C_1 D_t + C_2 D_{t-1} + C_3 dE_t + C_4 dE_{t-1} + C_5 dE_{t-2} + C_6 dE_{t-3} \\ + C_7 dE_{t-4} + C_8 E_{t-4} \end{aligned} \tag{93.42}$$

where C_0 to C_8 are functions of a_1 to a_1 and ρ_0 to ρ_4.

That is, the expectation adjustment in dividends, d_t, can be found by the coefficients estimated in Eqs. 93.31 and 93.36, i.e., a_1 to a_3 and ρ_0 to ρ_4, and the observable data from the time series of D_t and E_t.

Case 2 E_t follows a random walk process.

If the series of earnings, E_t, follows a random walk process, i.e., $\mathrm{E}_t[E_{t+1}] = E_t$, $\mathrm{E}_{t-1}[E_t] = E_{t-1}$, and $\mathrm{E}_{t-1}[E_{t+1}] = E_{t-1}$, then Eq. 93.35 can be redefined:

$$d_t \equiv \mathrm{E}_t[D_{t+1}] - \mathrm{E}_{t-1}[D_{t+1}] = C_0 + C_1 D_t + C_2 D_{t-1} + C_3 E_t + C_4 E_{t-1} \quad (93.43)$$

where $C_0 = -a_1 a_2$ $C_1 = a_2$, $C_2 = -a_2^2$, $C_3 = a_3$, and $C_4 = -a_3(1 + a_2)$.

That is, the expectation adjustment in dividends, d_t, can be found by the observable data from the time series of D_t and E_t.

In this study, we assumed that E_t follows a random walk process. Therefore, we used Eq. 93.43 instead of Eq. 93.42 to estimate d_t in Eqs. 93.22 and 93.23 in the text.

References

Amemiya, T. (1974). A note on a Fair and Jaffee model. *Econometrica, 42*, 759–762.
Black, S. W. (1976). Rational response to shocks in a dynamic model of capital asset pricing. *American Economic Review, 66*, 767–779.
Breeden, D. T. (1979). An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics, 7*, 265–196.
Campbell, J. Y., Grossman, S. J., & Wang, J. (1993). Trading volume and serial correlation in stock returns. *Quarterly Journal of Economics, 108*, 905–939.
Cox, J. C., Jr. Ingersoll, J. E., & Ross, S. A. (1985). An intertemporal general equilibrium model of asset prices. *Econometrica, 53*, 363–384.
Fair, R. C., & Jaffee, D. M. (1972). Methods of estimation of market in disequilibrium. *Econometrica, 40*, 497–514.
Fama, E., & French, K. R. (1992). The cross-section of expected stock return. *Journal of Finance, 47*, 427–465.
Greene, W. H. (2004). *Econometric analysis* (4th ed.). Upper Saddle River: Prentice Hall.
Grinols, E. L. (1984). Production and risk leveling in the intertemporal capital asset pricing model. *The Journal of Finance, 39*(5), 1571–1595.
Hamilton, J. D. (1994). *Time series analysis*. Princeton: Princeton University Press.
Lee, C. F., & Gweon, S. C. (1986). *Rational expectation, supply effect and stock price adjustment*. Paper presented at 1986, Econometrica Society Annual Meeting.
Lintner, J. (1965). The valuation of risky assets and the selection of risky investments in stock portfolios and capital budgets. *Review of Economics and Statistics, 47*, 13–37.
Lo, A. W., & Wang, J. (2000). Trading volume: Definition, data analysis, and implications of portfolio theory. *Review of Financial Studies, 13*, 257–300.
Markowitz, H. (1959). *Portfolio selection: Efficient diversification of investments*. New York: Wiley.
Merton, R. C. (1973). An intertemporal capital assets pricing model. *Econometrica, 41*, 867–887.
Miller, M. H. (1977). Debt and taxes. *Journal of Finance, 32*(2), 261–275.
Modigliani, F., & Miller, M. (1958). The cost of capital, corporation finance and the theory of investment. *American Economic Review, 48*, 261–297.
Mossin, J. (1966). Equilibrium in a capital asset market. *Econometrica, 35*, 768–783.
Quandt, R. E. (1988). *The econometrics of disequilibrium*. New York: Basil Blackwell.
Sharpe, W. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance, 19*, 425–442.
Tobin, J. (1958). Liquidity preferences as behavior toward risk. *Review of Economic Studies, 25*, 65–86.

A Generalized Model for Optimum Futures Hedge Ratio

94

Cheng-Few Lee, Jang-Yi Lee, Kehluh Wang, and
Yuan-Chung Sheu

Contents

C.-F. Lee (✉)
Department of Finance and Economics, Rutgers Business School, Rutgers, The State University of New Jersey, Piscataway, NJ, USA

Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
e-mail: cflee@business.rutgers.edu

J.-Y. Lee
Tunghai University, Taichung, Taiwan

K. Wang
Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
e-mail: lkwang@mail.nctu.edu.tw

Y.-C. Sheu
National Chiao-Tung University, Hsinchu, Taiwan
e-mail: sheu@math.nctu.edu.tw

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_94,

Abstract

Under martingale and joint-normality assumptions, various optimal hedge ratios are identical to the minimum variance hedge ratio. As empirical studies usually reject the joint-normality assumption, we propose the generalized hyperbolic distribution as the joint log-return distribution of the spot and futures. Using the parameters in this distribution, we derive several most widely used optimal hedge ratios: minimum variance, maximum Sharpe measure, and minimum generalized semivariance. Under mild assumptions on the parameters, we find that these hedge ratios are identical. Regarding the equivalence of these optimal hedge ratios, our analysis suggests that the martingale property plays a much important role than the joint distribution assumption.

To estimate these optimal hedge ratios, we first write down the log-likelihood functions for symmetric hyperbolic distributions. Then we estimate these parameters by maximizing the log-likelihood functions. Using these MLE parameters for the generalized hyperbolic distributions, we obtain the minimum variance hedge ratio and the optimal Sharpe hedge ratio. Also based on the MLE parameters and the numerical method, we can calculate the minimum generalized semivariance hedge ratio.

Keywords

Optimal hedge ratio • Generalized hyperbolic distribution • Martingale property • Minimum variance hedge ratio • Minimum generalized semivariance • Maximum Sharpe measure • Joint-normality assumption • Hedging effectiveness

94.1 Introduction

Because of their low transaction cost, high liquidity, high leverage, and ease of short position, stock index futures are among the most successful innovations in the financial markets. Besides the speculative trading, they are widely used to hedge against the market risk of the spot position. One of the most important issues for investors and portfolio managers is to calculate the optimal futures hedge ratio, the proportion of the position taken in futures to the size of the spot so that the risk exposure can be minimized.

The optimal hedge ratios typically depend on the objective functions under consideration. In literature on futures hedging, there are two different types of objective functions: the risk function to be minimized and the utility function to be maximized. Johnson (1960) obtains the minimum variance hedge ratio by minimizing the variance of the change in the value of the hedged portfolios. On the other hand, as Adams and Montesi (1995) indicate, corporate managers are more concerned with the downside risk rather than the upside variation. A measure of the downside risk is the generalized semivariance (GSV) where the risk is computed from the expectation of a power function of shortfalls from the target return (Bawa 1975, 1978; Fishburn 1977). De Jong et al. (1997) and Lien and Tse (1998, 2000, 2001) have calculated several GSV-minimizing hedge ratios. Regarding the utility function approach, we consider the Sharpe measure (SM) criteria, i.e., the ratio of the portfolio's excess return to its volatility. Howard and D'Antonio (1984) formulate the optimal hedge ratio by maximizing the Sharpe measure.

Normally, these optimal hedge ratios under different approaches are not the same. However, with the joint-normality and martingale assumptions, they are identical to the minimum variance hedge ratio. Unfortunately, many empirical studies indicate that major markets typically reject the joint-normality assumption (Chen et al. 2001; Lien and Tse 1998). In particular, the fat-tail property of the return distribution affects the hedging effectiveness substantially. It will be useful to find out the nature of the optimal hedge ratios under more realistic assumption. In this paper we introduce the bivariate generalized hyperbolic distributions as alternative joint distributions for returns in the spot and futures markets.

Barndorff-Nielsen (1977, 1978) develops the generalized hyperbolic (GH) distributions as a mixture of the normal distribution and the generalized inverse Gaussian (GIG) distribution first proposed in 1946 by Etienne Halphen. The class of the generalized hyperbolic distributions includes the hyperbolic distributions, the normal inverse Gaussian distributions, and the variance-Gamma distributions, while the normal distribution is a limiting case of the generalized hyperbolic distributions. Uses of the generalized hyperbolic distributions have been increasing in finance literature. To model the log returns of some financial assets, Eberlein and Keller (1995) consider the hyperbolic distribution and Barndorff-Nielsen (1995) proposes the normal inverse Gaussian distribution. For more recent applications of the generalized hyperbolic distributions in finance, see Bibby and Sørensen (2003), Eberlein et al. (1998), Rydberg (1997, 1999), Kücher et al. (1999), and Bingham and Kiesel (2001).

In terms of the parameters for the bivariate hyperbolic distributions, we have developed in this paper the minimum variance hedge ratio, GSV-minimizing hedge ratio, and the SM-maximizing hedge ratio. Moreover, the relationships between these hedge ratios are explored. In particular, under the martingale assumption, we can still obtain the result that these hedge ratios are the same as the minimum variance hedge ratio (see Theorems 2.1, 2.4 and Proposition 2.2). Based on the maximum likelihood estimation of the parameters and the numerical methods, we calculate and compare the different hedge ratios for TAIEX futures and S&P 500 futures.

The chapter is divided into five sections. Section 94.1 first introduces the definitions and some basic properties for GIG and GH distributions. In Sect. 94.2, we study the optimal hedge ratios under different approaches and estimate these ratios in terms of the parameters for GH distributions. In Sect. 94.3, we discuss the kernel density estimators and MLE method for parameter estimation problem. The last section provides the concluding remarks.

94.2 GIG and GH Distributions

94.2.1 The Generalized Hyperbolic Distributions

To introduce the generalized hyperbolic distribution, we first recall some basic properties of generalized inverse Gaussian (GIG) distributions. Note that for any $\delta, \psi > 0$ and $\lambda \in R$, the function

$$d_{GIG(\lambda,\delta,\psi)}(x) = \frac{(\psi/\delta)^{\lambda}}{2K_{\lambda}(\delta\psi)} x^{\lambda-1} e^{-\frac{1}{2}\left(\delta^2 x^{-1} + \psi^2 x\right)}, \qquad x > 0 \tag{94.1}$$

is a probability density function on $(0, \infty)$. Here, the function

$$K_{\lambda}(x) = \frac{1}{2}\int_0^{\infty} u^{\lambda-1} e^{-\frac{1}{2}x\left(u^{-1}+u\right)} du, \quad x > 0 \tag{94.2}$$

is the Bessel functions of the third kind with index λ. The distribution with the density function $d_{GIG(\lambda,\ \delta,\ \psi)}(x)$ on the positive half-line is called a generalized inverse Gaussian (GIG) distribution with parameters λ, δ, ψ and denoted by $GIG(\lambda, \delta, \psi)$. The moment generating function of the generalized inverse Gaussian distribution is given by

$$M_{GIG(\lambda,\ \delta,\ \psi)}(u) = \int_0^{\infty} e^{ux} d_{GIG(\lambda,\ \delta,\ \psi)}(x)dx = \left(\frac{\psi}{\sqrt{\psi^2 - 2u}}\right)^{\lambda} \frac{K_{\lambda}\left(\delta\sqrt{\psi^2 - 2u}\right)}{K_{\lambda}(\delta\psi)} \tag{94.3}$$

with the restriction $2u < \psi^2$. From this, we obtain

$$\mathbb{E}[GIG] = \frac{\delta}{\psi}\frac{K_{\lambda+1}(\delta\psi)}{K_{\lambda}(\delta\psi)}$$

$$\mathrm{Var}[GIG] = \left(\frac{\delta}{\psi}\right)^2 \left[\frac{K_{\lambda+2}(\delta\psi)}{K_{\lambda}(\delta\psi)} - \frac{K_{\lambda+1}^2(\delta\psi)}{K_{\lambda}^2(\delta\psi)}\right].$$

Barndorff-Nielsen (1977) introduced the class of generalized hyperbolic (GH) distributions as mean-variance mixtures of normal distributions. More precisely, one says that a random variable Z has the generalized hyperbolic distribution $GH(\lambda, \alpha, \beta, \delta, \mu)$ if

$$Z|Y = y \sim N(\mu + \beta y, y),$$

where Y is a random variable with distribution $GIG\left(\lambda, \delta, \sqrt{\alpha^2 - \beta^2}\right)$ and $N(\mu + \beta y, y)$ denotes the normal distribution with mean $\mu + \beta y$ and variance y. From this, one can easily verify that the density function for $GH(\lambda, \alpha, \beta, \delta, \mu)$ is given by the formula

$$\begin{aligned} d_{GH(\lambda,\alpha,\beta,\delta,\mu)}(x) &= \int_0^\infty d_{N(\mu+\beta y, y)}(x) d_{GIG\left(\lambda,\delta,\sqrt{\alpha^2-\beta^2}\right)}(y)dy \\ &= \left(\frac{\psi}{\delta}\right)^\lambda \frac{e^{(x-\mu)\beta}}{\sqrt{2\pi}K_\lambda(\delta\psi)} \left[\frac{\delta^2 + (x-\mu)^2}{\alpha^2}\right]^{\frac{\lambda-\frac{1}{2}}{2}} K_{\lambda-\frac{1}{2}}\left(\alpha\sqrt{\delta^2 + (x-\mu)^2}\right) \end{aligned}$$

where $\psi = \sqrt{\alpha^2 - \beta^2}$.

The class of hyperbolic distributions is the subclass of GH distributions obtained when λ is equal to 1. We write $H(\alpha, \beta, \delta, \mu)$ instead of $GH(1, \alpha, \beta, \delta, \mu)$. Using the fact that $K_{1/2}(z) = (\pi/2z)^{1/2}e^{-z}$, one obtains the density for $H(\alpha, \beta, \delta, \mu)$ is

$$d_{H(\alpha,\beta,\delta,\mu)}(x) = \frac{\sqrt{\alpha^2 - \beta^2}}{2\alpha\delta K_1\left(\delta\sqrt{\alpha^2-\beta^2}\right)} e^{-\alpha\sqrt{\delta^2+(x-\mu)^2} + \beta(x-\mu)}. \qquad (94.4)$$

The normal inverse Gaussian (NIG) distributions were introduced to finance in Barndorff-Nielsen (1995). It is a subclass of the generalized hyperbolic distributions obtained for λ equal to $-1/2$. The density of the NIG distribution is given by

$$d_{NIG(\alpha,\beta,\delta,\mu)}(x) = \frac{\delta}{\pi}\left[\frac{\alpha^2}{\delta^2 + (x-\mu)^2}\right]^{\frac{1}{2}} e^{\delta\psi + (x-\mu)\beta} K_1\left(\alpha\sqrt{\delta^2+(x-\mu)^2}\right).$$

94.2.2 Multivariate Modeling

In finance one does not look at a single asset, but at a bunch of assets. Since the assets in the market are typically highly correlated, it is natural to use multivariate distributions. A straightforward way for introducing multivariate generalized hyperbolic (MGH) distributions is via the mixtures of multivariate normal distributions with the generalized inverse Gaussian distributions. In fact the multivariate generalized hyperbolic distributions were introduced and investigated in Barndorff-Nielsen (1978).

Let Δ be a symmetric positive-definite $d \times d$- matrix with determinant $|\Delta| = 1$. Assume that $\lambda \in R$, β, $\mu \in R^d$, $\delta > 0$, and $\alpha^2 > \beta'\Delta\beta$. We say that a d-dimensional random vector Z has the multivariate generalized hyperbolic distribution $MGH(\lambda, \alpha, \beta, \delta, \mu, \Delta)$ with parameters $(\lambda, \alpha, \beta, \delta, \mu, \Delta)$ if

$$\mathbf{Z}|Y = y \sim N_d(\mu + y\Delta\beta, y\Delta),$$

where $Nd(A, B)$ denotes the d-dimensional normal distribution with mean vector A and covariance matrix B, *and Y distribution as* $GIG\left(\lambda, \delta, \sqrt{\alpha^2 - \beta' \Delta\beta}\right)$. Here we notice that the generalized hyperbolic distributions are symmetric if and only if $\beta = (0, \ldots, 0)'$. For $\lambda = (d + 1)/2$ we obtain the multivariate hyperbolic distributions. For $\lambda = -1/2$ we obtain the multivariate normal inverse Gaussian distribution.

The density function of the distribution $MGH(\lambda, \beta, \delta, \mu, \Delta)$ is given by the formula

$$d_{MGH}(x) = c_d \frac{K_{\lambda-d/2}\left(\alpha\sqrt{\delta^2 + (x-\mu)'\Delta^{-1}(x-\mu)}\right)}{\left(\alpha^{-1}\sqrt{\delta^2 + (x-\mu)'\Delta^{-1}(x-\mu)}\right)^{d/2-\lambda}} e\left(\beta'(x-\mu)\right) \tag{94.5}$$

where $c_d = \frac{\left[\left(\alpha^2 - \beta'\Delta\beta\right)/\delta^2\right]^{\lambda/2}}{(2\pi)^{d/2}K_\lambda\left(\delta\sqrt{\alpha^2-\beta'\Delta\beta}\right)}$. The mean and covariance of MGH are given by

$$\mathbb{E}[MGH(\lambda, \alpha, \beta, \delta, \mu, \Delta)] = \mu + \Delta\beta\mathbb{E}[GIG(\lambda, \delta, \psi)], \tag{94.6}$$

$$\begin{aligned} \mathrm{Var}[MGH(\lambda, \alpha, \beta, \delta, \mu, \Delta)] = {} & \Delta\mathbb{E}[GIG(\lambda, \delta, \psi)] \\ & + \Delta\beta\beta'\Delta\mathrm{Var}[GIG(\lambda, \delta, \psi)] \end{aligned} \tag{94.7}$$

where $\psi = \sqrt{\alpha^2 - \beta'\Delta\beta}$. (For details, see, e.g., Blæsid (1981).)

94.3 Futures Hedge Ratios

We consider a decision maker. At the decision date ($t = 0$), the agent engages in the production of $Q(Q > 0)$ commodity units for sale at the terminal date ($t = 1$) at the random cash price $P1$. In addition, at the decision date the agent can sell X commodity units in the futures market at the price F_0 but must repurchase them back at the terminal date at the random futures price F_1. Let the initial wealth be $V_0 = P_0 Q$ and the end-of-period wealth be $V_1 = P_1 Q + (F_0 - F_1)X$. Then we consider the wealth return that is

$$\begin{aligned} \widetilde{r}_\theta &= \frac{V_1 - V_0}{V_0} = \frac{P_1 Q + F_0 X - F_1 X - P_0 Q}{P_0 Q} \\ &= \frac{P_1 - P_0}{P_0} - \frac{F_1 - F_0}{F_0}\left(\frac{F_0}{P_0}\frac{X}{Q}\right) = \widetilde{r}_p - \theta\widetilde{r}_f \end{aligned} \tag{94.8}$$

where $\widetilde{r}_p = (P_1 - P_0)/P_0$ and $\widetilde{r}_f = (F_1 - F_0)/F_0$ are one-period returns on the spot and futures positions, respectively. $h = X/Q$ is the hedge ratio and $\theta = h(F_0/P_0)$. (Note that θ is so-called the adjusted hedge ratio.)

The main objective of hedging is to choose the optimal hedge ratio θ. However, the optimal hedge ratio will depend on a particular objective function to be optimized. We recall some most widely used theoretical approaches to the optimal futures hedge ratios and compute explicitly these optimal ratios in terms of the parameters for MGH distributions. For a comprehensive review of futures hedge ratios, see Chen et al. (2003).

94.3.1 Minimum Variance Hedge Ratio

The most widely used hedge ratio is the minimum variance hedge ratio which is known as the MV hedge ratio. The objective function to be minimized is the variance of $\widetilde{r}_\theta$.

Clearly we have $\mathrm{Var}[\widetilde{r}_\theta] = \sigma_{r_p}^2 + \theta^2\sigma_{r_f}^2 - 2\theta\rho\sigma_{r_p}\sigma_{r_f}$, where σ_{rp} and σ_{rf} are standard deviations of $\widetilde{r}_p$ and $\widetilde{r}_f$, respectively, and ρ is the correlation coefficient between $\widetilde{r}_p$ and $\widetilde{r}_f$. The MV hedge ratio is obtained by minimizing $\mathrm{Var}[\widetilde{r}_\theta]$. Simple calculation shows that the MV hedge ratio is given by

$$\theta_{MV}^* = \rho\frac{\sigma_{r_p}}{\sigma_{r_f}}. \tag{94.9}$$

Theorem 2.1 *Assume* $(\widetilde{r}_f, \widetilde{r}_p)'$ *is distributed as MGH*$(\lambda, \alpha, \beta, \delta, \mu, \Delta)$, *where* $\beta = (\beta_1, \beta_2)'$, $\mu = (\mu_1, \mu_2)'$, *and* $\Delta = \begin{pmatrix} \Delta_{11} & \Delta_{12} \\ \Delta_{21} & \Delta_{22} \end{pmatrix}$ *is symmetry. Then we have*

$$\theta^*{}_{MV} = \frac{\Delta_{12}\mathbb{E}[GIG] + \delta_{fp}}{\Delta_{11}\mathbb{E}[GIG] + \delta_{ff}} \tag{94.10}$$

where $GIG = GIG\left(\lambda, \delta, \sqrt{\alpha^2 - \beta'\Delta\beta}\right)$ *and*

$$\begin{aligned}\delta_{ff} &= \left[\beta_1^2\Delta_{11}^2 + 2\beta_1\beta_2\Delta_{11}\Delta_{12} + \beta_2^2\Delta_{12}^2\right]\mathrm{Var}[GIG]\\ \delta_{fp} &= \left[\beta_1^2\Delta_{11}\Delta_{12} + \beta_1\beta_2\left(\Delta_{11}\Delta_{22} + \Delta_{12}^2\right) + \beta_2^2\Delta_{12}\Delta_{22}\right]\mathrm{Var}[GIG].\end{aligned}$$

In particular, if $\beta = (0, \ldots, 0)'$, *then* $\theta_{MV}^* = \frac{\Delta_{12}}{\Delta_{11}}$.

94.3.2 Sharpe Hedge Ratio

We consider the optimal hedge ratio that incorporates both risk and expected return. Howard and D'Antonio (1984) considered the optimal level of futures contracts by maximizing the ratio of the portfolio's excess return to its volatility, that is,

$$\max_\theta \frac{\mu_{r_p} - \theta\mu_{r_f} - r_L}{\sigma_\theta}, \tag{94.11}$$

where σ_θ is the standard deviation of $\widetilde{r}_\theta, \mu_{r_p}, \mu_{r_f}$ are expected values for $\widetilde{r}_p$ and $\widetilde{r}_f$, respectively, and r_L is the risk-free interest rate.

Consider the function

$$r(\theta) = \frac{\mu_{r_p} - \theta\mu_{r_f} - r_L}{\sigma_\theta}.$$

Then we have

$$r'(\theta) = \frac{\theta\left[-\sigma_{r_f}^2\left(\mu_{r_p} - r_L\right) + \mu_{r_f}\sigma_{r_f r_p}\right] + \left(\mu_{r_p} - r_L\right)\sigma_{r_p r_f} - \sigma_{r_p}^2\mu_{r_f}}{\sigma_\theta^3} \tag{94.12}$$

where $\sigma_{r_f r_p} = \mathrm{Cov}\left(\widetilde{r}_p, \widetilde{r}_f\right)$ and, hence, the critical point for $r(\theta)$ is given by

$$\theta_s^* = \frac{\left(\frac{\sigma_{rp}}{\sigma_{rf}}\right)^2\mu_{rf} - \rho\frac{\sigma_{rp}}{\sigma_{rf}}\left(\mu_{r_p} - r_L\right)}{\rho\frac{\sigma_{rp}}{\sigma_{rf}}\mu_{rf} - \left(\mu_{r_p} - r_L\right)}. \tag{94.13}$$

It follows from Eq. 94.12 that if $\mu_{r_p} - \mu_L > \rho\frac{\sigma_{rp}}{\sigma_{r_f}}\mu_{rf}$, then $r'(\theta) > 0$ for $\theta < \theta_s^*$ and $r'(\theta) < 0$ for $\theta > \theta_s^*$. Hence, θ_s^* is the optimal hedge ratio (Sharpe hedge ratio) for Eq. 94.11. Similarly, if $\mu_{r_p} - \mu_L < \rho\frac{\sigma_{rp}}{\sigma_{r_f}}\mu_{rf}$, then $r(\theta)$ has a minimum at θ_s^*. (Note that if $\mu_{r_p} - \mu_L = \rho\frac{\sigma_{rp}}{\sigma_{r_f}}\mu_{rf}$, then $r(\theta)$ is strictly monotonic in θ_s^*.)

The measure of hedging effectiveness (abbreviated HE) is given in Howard and D'Antonio (1984) by

$$\mathrm{HE} = r\left(\theta_s^*\right)/\left(\frac{\mu_{r_p} - r_L}{\sigma_{r_p}}\right). \tag{94.14}$$

Write

$$\zeta = \frac{\mu_{r_f}/\sigma_{r_f}}{\left(\mu_{r_p} - \mathrm{r_L}\right)/\sigma_{r_p}}. \tag{94.15}$$

(ζ is also called the risk-return relative.) Then we have

$$\theta_s^* = \frac{\sigma_{rp}}{\sigma_{rf}}\left(\frac{\rho - \zeta}{1 - \zeta\rho}\right)$$

and

$$\mathrm{HE} = \sqrt{\frac{(\rho - \zeta)^2}{1 + \rho^2} + 1}.$$

Clearly the last equality implies that

$$HE\begin{cases} > 1 & \text{when} \quad \rho \neq \zeta \\ = 1 & \text{when} \quad \rho = \zeta. \end{cases}$$

Moreover, without any distribution assumption, we have the following relationship between θ_s^* and θ_{MV}^*. In particular, if the expected return on the futures contract is zero and $\mu_{r_p} > r_L$, then the Sharpe hedge ratio reduces to the minimum variance hedge ratio.

Proposition 2.2 *Assume $\mu_{r_p} > r_L$ and $1 > \zeta\rho$. Then we have*

$$\begin{cases} \theta_s^* > \theta_{MV}^* & \text{when} \quad \mu_f < 0 \\ \theta_s^* = \theta_{MV}^* & \text{when} \quad \mu_f = 0 \\ \theta_s^* < \theta_{MV}^* & \text{when} \quad 0 < \mu_f. \end{cases}$$

Recall that $\sigma_{r_f r_p} = \mathrm{Cov}\left(\widetilde{r}_p, \widetilde{r}_f\right)$. Then we have

$$\theta_s^* = \frac{\sigma_{r_p}^2 \mu_{r_f} - \sigma_{r_f r_p}\left(\mu_{r_p} - r_L\right)}{\sigma_{r_f r_p}\mu_{r_f} - \sigma_{r_f}^2\left(\mu_{r_p} - r_L\right)}. \tag{94.16}$$

From this and by Eqs. 94.6 and 94.7, we obtain

Theorem 2.3 *Assume $\left(\widetilde{r}_f, \widetilde{r}_p\right)'$ is distributed as in Theorem 2.1. Assume that $\zeta_{fp}[(\mu_1 + \beta_1\Delta_{11} + \beta_2\Delta_{12})\mathbb{E}[GIG]] < \zeta_{ff}[(\mu_2 + \beta_1\Delta_{21} + \beta_2\Delta_{22})\mathbb{E}[GIG] - \mathrm{r_L}]$.*

Then we have

$$\theta_s^* = \frac{\zeta_{pp}[(\mu_1 + \beta_1\Delta_{11} + \beta_2\Delta_{12})\mathbb{E}[GIG]] - \zeta_{fp}[(\mu_2 + \beta_1\Delta_{21} + \beta_2\Delta_{22})\mathbb{E}[GIG] - \mathrm{r_L}]}{\zeta_{fp}[(\mu_1 + \beta_1\Delta_{11} + \beta_2\Delta_{12})\mathbb{E}[GIG]] - \zeta_{ff}[(\mu_2 + \beta_1\Delta_{21} + \beta_2\Delta_{22})\mathbb{E}[GIG] - \mathrm{r_L}]} \tag{94.17}$$

where δ_{ff}, δ_{fp}, GIG are the same as in Theorem 2.1 and

$$\begin{aligned} \delta_{pp} &= \left[\beta_1^2\Delta_{21}^2 + 2\beta_1\beta_2\Delta_{21}\Delta_{22} + \beta_2^2\Delta_{22}^2\right]\mathrm{Var}[GIG] \\ \zeta_{ff} &= \Delta_{11}\mathbb{E}[GIG] + \delta_{ff} \\ \zeta_{fp} &= \Delta_{12}\mathbb{E}[GIG] + \delta_{fp} \\ \zeta_{pp} &= \Delta_{22}\mathbb{E}[GIG] + \delta_{pp}. \end{aligned}$$

94.3.3 Minimum Generalized Semivariance Hedge Ratio

In this case, the optimal hedge ratio is obtained by minimizing the generalized semivariance (GSV) given below:

$$L_n(c, X) = \int_{-\infty}^{c} (c - x)^n dF(x), \quad n > 0, \tag{94.18}$$

where $F(\cdot 7)$ is the probability distribution function of the return X. The GSV is specified by two parameters: the target return c and the power of the shortfall n. (Note that if the density function of X is symmetric at c, then we obtain $L_2(c, X) = \mathrm{Var}(X)/2$. Hence, in this case, the GSV approach is the same as that of the minimum variance.) The GSV, due to its emphasis on the returns below the target return, is consistent with the risk perceived by managers (see Lien and Tse 2001). For futures hedge, we consider $L_n(c, \theta) = L_n\left(c, \widetilde{r}_p - \theta \widetilde{r}_f\right)$.

Under some conditions on the joint distribution, we obtain that the minimum GSV hedge ratio is the same as the minimum variance hedge ratio.

Theorem 2.4 *Assume* $\left(\widetilde{r}_f, \widetilde{r}_p\right)$ *is the same as in Theorem 2.1. If* $\beta = 0$ *and* $\mu_1 = \mu_{rf} = 0$, *then the minimum GSV hedge ration is the same as the minimum variance hedge ration i.e.,* $\left(\theta^*_{GSV} = \theta^*_{MV}, = \frac{\Delta 12}{\Delta 11}\right)$.

In empirical studies, the true distribution is unknown or complicated. Then θ^*_{GSV} can be estimated from the sample by using the so-called empirical distribution method adapted in, e.g., Price et al. (1982) and Harlow (1991). Suppose we have m observations of $\left(\widetilde{r}_f, \widetilde{r}_p\right)$, say, $(r_f(i), r_p(i))$, $i = 1, 2, \ldots, m$. From this, the GSV can be estimated by the formula

$$L_n^{obs}(c, \theta) = \frac{1}{m} \sum_{i=1}^{m} \left(c - r_{i,\theta}\right)^n I_{ri,\theta \le c}, \tag{94.19}$$

where $r_{i,\theta} = r_p(i) - \theta r_f(i)$. Given c and n, numerical methods can be used to search the hedge ratio that minimizing the sample GSV, $L_n^{obs}(c,\theta)$.

94.4 Estimation and Simulation

94.4.1 Kernel Density Estimators

Assumed that we have n independent observations $x_1, \ldots, x_n$ from the random variable X with the unknown density function f. The kernel density estimator for the estimation of f is given by

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^{n} K\left(\frac{x - x_i}{h}\right), \quad x \in R \tag{94.20}$$

where K is a so-called kernel function and h is the bandwidth. In this chapter we work with the Gaussian kernel: $K(x) = 1/\sqrt{2\pi}\ \exp\left\{-x^2/2\right\}$ and $h = (4/3)^{1/5} \sigma n^{\frac{-1}{5}}$.

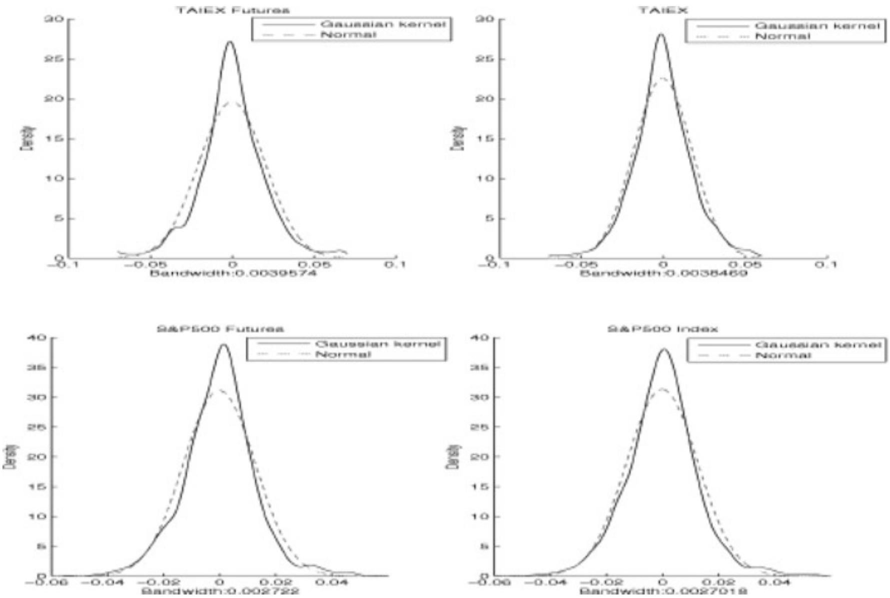


Fig. 94.1 Normal density and Gaussian kernel density estimators

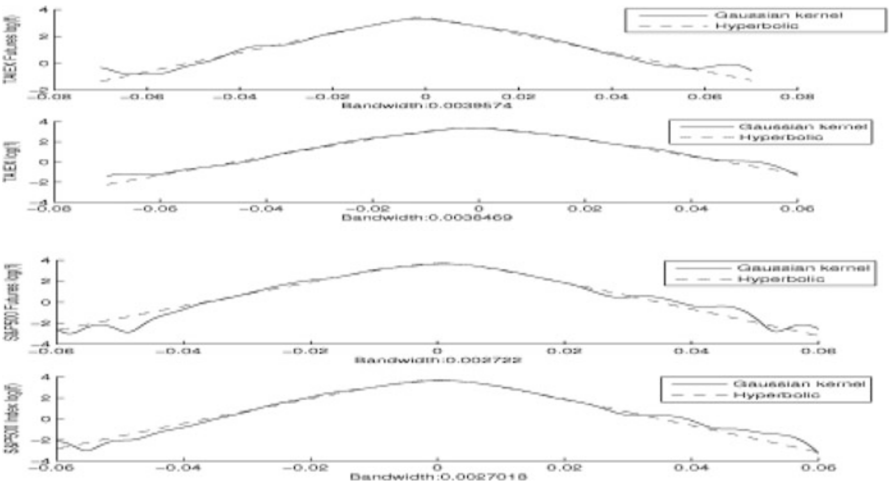


Fig. 94.2 Log-densities of daily log returns of major indices and futures (2000–2004)

(For more details, see Scott (1979).) Meanwhile it is worth noting that Lien and Tse (2000) proposed the kernel density estimation method to estimate the probability distribution of the portfolio return for every θ, and then grid search methods were adapted to find the optimum GSV hedge ratio (Figs. 94.1 and 94.2).

94.4.2 Maximum Likelihood Estimation

We focus on how to estimate the parameters of a density function $f(x; \Theta)$, where Θ is the set of parameters to be estimated. Suppose that we have m independent observations $x1, \ldots, xn$ of a random variable X with the density function $f(x; \Theta)$. The maximum likelihood estimator $\hat{\theta}_{MLE}$ is the parameter set that maximizes the likelihood function

$$L(\Theta) = \prod_{i=1}^{n} f(x_i; \Theta).$$

Clearly this is equivalent to maximizing the logarithm of the likelihood function

$$\log\ L(\Theta) = \sum_{i=1}^{n} \log\ f(x_i; \Theta).$$

The log-likelihood function for hyperbolic distribution $H(\alpha, \beta, \delta, \mu)$ is given by

$$\ell_{H(\alpha,\beta,\delta,\mu)}(\Theta) = n\left(\log\sqrt{\alpha^2 - \beta^2} - \log\ 2 - \log\ \alpha - \log\ \delta - \log\ K_1\left(\delta\sqrt{\alpha^2 - \beta^2}\right)\right) + \sum_{i=1}^{n}\left[-\alpha\sqrt{\delta^2 + (x_i - \mu)^2} + \beta(x_i - \mu)\right].$$

The symmetric MGH density function is given by the formula

$$\frac{(\alpha/\delta)^{\lambda}}{(2\pi)^{d/2} K_{\lambda}(\alpha\delta)} \frac{K_{\lambda-\frac{d}{2}}\left(\alpha\sqrt{\delta^2 + (x-\mu)'\Delta^{-1}(x-\mu)}\right)}{\left(\alpha^{-1}\sqrt{\delta^2 + (x-\mu)'\Delta^{-1}(x-\mu)}\right)^{\frac{d}{2}-\lambda}},$$

and, in particular, the two-dimensional symmetric hyperbolic distributions (i.e., $\beta = 0$ and $\lambda = 3/2$) have the density

$$\mathbf{H}_2 = \frac{(\alpha/\delta)^{3/2}}{2^{3/2}\sqrt{\pi\alpha K_{\frac{3}{2}}(\alpha\delta)}} e^{-\alpha}\sqrt{\delta^2 + (x - \mu')\Delta^{-1}(x-\mu).}$$

From this, we obtain the log-likelihood function for two-dimensional symmetric hyperbolic distributions (Figs. 94.3 and 94.4):

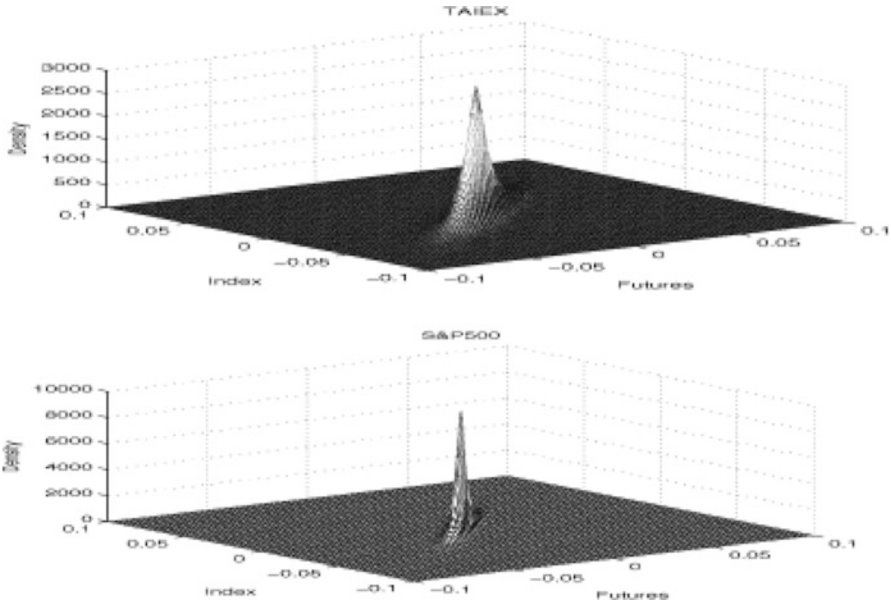


Fig. 94.3 Estimated symmetric H_2 distributions

$$\ell_{\mathbf{H}_2} = n\left[\frac{3}{2}\log\frac{\alpha}{\delta} - \frac{3}{2}\log\ 2 - \frac{1}{2}\log\ \pi - \log\ \alpha - \log\ K_{\frac{3}{2}}(\alpha\delta)\right] - \alpha\sum_{i=1}^{n}\sqrt{\delta^2 + (x_i - \mu)'\Delta^{-1}(x_i - \mu)}.$$

94.4.3 Simulation of Generalized Hyperbolic Random Variables

From the representation of GH distribution as a conditional normal distribution mixed with the generalized inverse Gaussian, a schematic representation of the algorithm reads as follows:

1. Sample Y from GIG(λ, δ, ψ) distribution
2. Sample ε from $N(0, 1)$
3. Return $X = \mu + \beta Y + \sqrt{Y}\varepsilon$

Similarly, for simulating an MGH distributed random vector, we have:

1. Set $\Delta = L^T L$ via Cholesky decomposition
2. Sample Y from GIG(λ, δ, ψ) distribution
3. Sample Z from $N(0, I)$, where I is $d \times d$-identity matrix
4. Return $X = \mu + Y\Delta\beta + \sqrt{Y}L^T Z$

The efficiency of the above algorithms depends on the method of sampling the generalized inverse Gaussian distributions. Atkinson (1982) applied the method of rejection algorithm to sampling GIG. We adopt their method for simulation of estimated hyperbolic random variables.

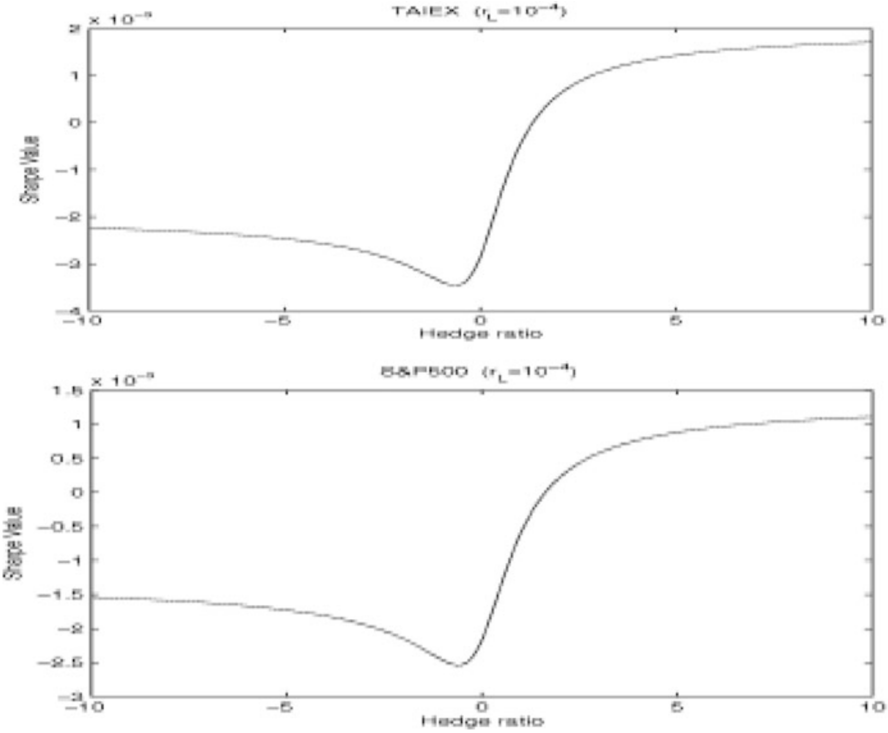


Fig. 94.4 Sharpe value

94.5 Concluding Remarks

Although there are many different theoretical approaches to the optimal futures hedge ratios, under the martingale and joint-normality assumptions, various optimal hedge ratios are identical to the minimum variance hedge ratio. However, empirical studies show that major market data reject the joint-normality assumption. In this paper we propose the generalized hyperbolic distribution as the joint log-return distribution of the spot and futures. In terms of the parameters for generalized hyperbolic distributions, we obtain several most widely used optimal hedge ratios: minimum variance, maximum Sharpe measure, and minimum generalized semivariance. In particular, under mild assumptions on the parameters, we show that these theoretical approaches are equivalent.

To estimate these optimal hedge ratios, we first write down the log-likelihood functions for symmetric hyperbolic distributions. Then we calculate these parameters by maximizing the log-likelihood functions. Using these MLE parameters for the GH distributions, we obtain the MV hedge ratio and the optimal Sharpe hedge ratio by Theorems 2.1 and 2.3, respectively. Also based on the MLE parameters and the numerical method, we calculate the minimum generalized semivariance hedge ratio.

Regarding the equivalence of these three optimal hedge ratios, our results suggest that the martingale property plays a much important role than the joint distribution assumption.

However, conditional heteroskedasticity and stochastic volatility are observed in many spot and futures price series. This implies that the optimal hedge strategy should be time-dependent. To account for this dynamic property, parametric specifications of the joint distribution are required. Based on our work here, it is interesting to extend the results to time-varying hedge ratios.

References

Adams, J., & Montesi, C. J. (1995). *Major issues related to hedge accounting*. Newark: Financial Accounting Standard Board.

Atkinson, A. C. (1982). The simulation of generalized inverse Gaussian and hyperbolic random variables. *SIAM Journal of Scientific and Statistical Computing, 3*, 502–515.

Barndorff-Nielsen, O. E. (1977). Exponentially decreasing distributions for the logarithm of particle size. *Proceedings of the Royal Society London A, 353*, 401–419.

Barndorff-Nielsen, O. E. (1978). Hyperbolic distributions and distributions on hyperbolae. *Scandinavian Journal of Statistics, 5*, 151–157.

Barndorff-Nielsen, O. E. (1995). *Normal inverse Gaussian distributions and the modeling of stock returns*. Research Report no. 300, Department of Theoretical Statistics, Aarhus University.

Bawa, V. S. (1975). Optimal rules for ordering uncertain prospects. *Journal of Financial Economics, 2*, 95–121.

Bawa, V. S. (1978). Safety-first, stochastic dominance, and optimal portfolio choice. *Journal of Financial and Quantitative Analysis, 13*, 255–271.

Bibby, B. M., & Sørensen, M. (2003). Hyperbolic processes in finance. In S. T. Rachev (Ed.), *Handbook of heavy tailed distributions in finance* (pp. 211–248). Amsterdam/The Netherlands: Elsevier Science.

Bingham, N. H., & Kiesel, R. (2001). Modelling asset returns with hyperbolic distribution. In J. Knight & S. Satchell (Eds.), *Return distribution in Finance* (pp. 1–20). Oxford/Great Britain: Butterworth-Heinemann.

Blæsid, P. (1981). The two-dimensional hyperbolic distribution and related distribution with an application to Johannsen's bean data. *Biometrika, 68*, 251–263.

Chen, S. S., Lee, C. F., & Shrestha, K. (2001). On a mean-generalized semivariance approach to determining the hedge ratio. *Journal of Futures Markets, 21*, 581–598.

Chen, S. S., Lee, C. F., & Shrestha, K. (2003). Futures hedge ratio: A review. *The Quarterly Review of Economics and Finance, 43*, 433–465.

De Jong, A., De Roon, F., & Veld, C. (1997). Out-of-sample hedging effectiveness of currency futures for alternative models and hedging strategies. *Journal of Futures Markets, 17*, 817–837.

Eberlein, E., & Keller, U. (1995). Hyperbolic distributions in finance. *Bernoulli, 1*, 281–299.

Eberlein, E., Keller, U., & Prause, K. (1998). New insights into smile, mispricing and value at risk: The hyperbolic model. *Journal of Business, 71*, 371–406.

Fishburn, P. C. (1977). Mean-risk analysis with risk associated with below-target returns. *American Economic Review, 67*, 116–126.

Harlow, W. V. (1991). Asset allocation in a downside-risk framework. *Financial Analysts Journal, 47*, 28–40.

Howard, C. T., & D'Antonio, L. J. (1984). A risk-return measure of hedging effectiveness. *Journal of Financial and Quantitative Analysis, 19*, 101–112.

Johnson, L. L. (1960). The theory of hedging and speculation in commodity futures. *Review of Economic Studies, 27*, 139–151.

Kücher, U., Neumann, K., Sørensen, M., & Streller, A. (1999). Stock returns and hyperbolic distributions. *Mathematical and Computer Modelling, 29*, 1–15.

Lien, D., & Tse, Y. K. (1998). Hedging time-varying downside risk. *Journal of Futures Markets, 18*, 705–722.

Lien, D., & Tse, Y. K. (2000). Hedging downside risk with futures contracts. *Applied Financial Economics, 10*, 163–170.

Lien, D., & Tse, Y. K. (2001). Hedging downside risk: Futures vs. options. *International Review of Economics and Finance, 10*, 159–169.

Price, K., Price, B., & Nantel, T. J. (1982). Variance and lower partial moment measures of systematic risk: Some analytical and empirical results. *Journal of Finance, 37*, 843–855.

Rydberg, T. H. (1997). The normal inverse Gaussian Levy process: Simulation and approximation. *Communications in Statistics: Stochastic models, 13*, 887–910.

Rydberg, T. H. (1999). Generalized hyperbolic diffusion processes with applications in finance. *Mathematical Finance, 9*, 183–201.

Scott, W. (1979). On optimal and data-based histograms. *Biometrika, 33*, 605–610.

Instrumental Variables Approach to Correct for Endogeneity in Finance

95

Chia-Jane Wang

Contents

Abstract

The endogeneity problem has received a mixed treatment in corporate finance research. Although many studies implicitly acknowledge its existence, the literature does not consistently account for endogeneity using formal econometric methods. This chapter reviews the instrumental variables (IV) approach to endogeneity from the point of view of a finance researcher who is implementing instrumental variables methods in empirical studies. This chapter is organized into two parts. Part I discusses the general procedure of the instrumental variables approach, including two-stage least square (2SLS) and generalized method

C.-J. Wang
Manhattan College, Riverdale, NY, USA
e-mail: chia-jane.wang@manhattan.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_95,

of moments (GMM), the related diagnostic statistics for assessing the validity of instruments, which are important but not used very often in finance applications, and some recent advances in econometrics research on weak instruments. Part II surveys corporate finance applications of instrumental variables. We found that the instrumental variables used in finance studies are usually chosen arbitrarily, and very few diagnostic statistics are performed to assess the adequacy of IV estimation. The resulting IV estimates thus are questionable.

Keywords
Endogeneity • OLS • Instrumental variables (IV) estimation • Simultaneous equations • 2SLS • GMM • Overidentifying restrictions • Exogeneity test • Weak instruments • Anderson-Rubin statistic • Empirical corporate finance

95.1 Introduction

Corporate finance often involves a number of decisions that are intertwined and endogenously chosen by managers and/or debtholders and/or shareholders. For example, in order to maximize value, firms must form an effective board of directors and grant their managers an optimal pay-performance compensation contract. Debtholders have to decide how much debt and with what features (junior or senior, convertible, callable, maturity length, etc.) should the debt be structured. Given that these endogenously chosen variables are interrelated and are often partially driven by unobservable firm characteristics, the endogeneity problem can make the standard OLS results hard to interpret (Hermalin and Weisbach 2003).

The usual approach to the endogeneity problem is to implement the instrumental variables estimation method. In particular, one chooses instrumental variables that are correlated to the endogenous regressors, but uncorrelated to the structural equation errors and then employs a two-stage least square (2SLS) procedure.

The endogeneity problem has received a mixed treatment in corporate finance research. Although many studies implicitly acknowledge its existence, the literature does not consistently account for endogeneity using formal econometric methods. There is an increasing emphasis on addressing the endogeneity problem in recent work, and the simultaneous equations model is now being used more commonly. However, the instrumental variables are often chosen arbitrarily and few diagnostic statistics are performed to assess the adequacy of IV estimation.

This chapter reviews the instrumental variables approach to endogeneity from the point of view of a finance researcher who is implementing instrumental variables methods in empirical studies. We do not say much on the distribution theory and the mathematical proofs of estimation methods and test statistics as they are well covered in econometrics textbooks and articles. The chapter proceeds as follows: Sect. 95.2 describes the statistical issue raised by endogeneity. Section 95.3 discusses the commonly used instrumental variables approaches to endogeneity: the two-stage least squares estimation (2SLS) and the generalized method of moments (GMM) estimation. Section 95.4 discusses the conditions of a valid instrument and the related diagnostic statistics, which are critical but not frequently used in finance applications. Section 95.5 considers the weak instrument problem, that is, the instruments are exogenous but have weak explanatory power for explaining the endogenous regressor. Some recent advances in econometrics research on statistical inference with weak instruments are briefly discussed. Section 95.6 surveys corporate finance applications of instrumental variables. Section 95.7 concludes.

95.2 Endogeneity: The Statistical Issue

To illustrate the endogeneity issue, assume that we wish to estimate parameter β of a linear regression model for a population of firms

$$Y = X\beta + u \tag{95.1}$$

where Y is the dependent variable, which is typically an outcome such as return or profitability, X is the regressor that explains the outcome, and u is the unobservable random disturbance or error. If u satisfies the classical regression conditions, the parameters β can be consistently estimated by the standard OLS procedures. This can be shown from the probability limit of OLS estimator

$$\text{Plim } \beta_{OLS} = \beta + \text{Cov}(X, u)/\text{Var}(X) \tag{95.2}$$

When the disturbance and the regressor are not correlated and hence the second term is zero, the OLS estimator will be consistent. But if the disturbance is correlated with the regressor, that is, the explanatory variable in Eq. 95.1 is potentially endogenous,[1] the usual OLS estimation generally results in biased estimator.

In applied finance work, endogeneity is often caused by omitted variables and/or simultaneity. Omitted variables problem arises when the explanatory variable X is hard to measure or depends on some unobservable factors, which are part of the error term u, thus X and u are correlated. The correlation of explanatory variables with unobservable may be due to self-selection: the firm makes the choice of X for the reason that is unobservable. An example is the private information held by a firm in making a debt issue, in which the terms and structure of the debt offering are likely to be correlated with unobserved private information held by the firm. The Heckman two-step procedure (Heckman 1979) is extensively used in corporate

[1]Traditionally, a variable is defined as endogenous if it is determined within the context of a model. However, in applied econometrics the "endogenous" variable is used more broadly to describe the situation where an explanatory variable is correlated with the disturbance and the resulting estimator is biased (Wooldridge 2002).

finance for modeling this omitted variable.[2] Endogeneity may also be caused by simultaneity, in which one or more of the explanatory variables are determined simultaneously along with the dependent variable. For example, if Y is firm value and X is management compensation, the management compensation contract is partly determined by the anticipated firm value Y from choosing X, and the general OLS estimator is biased in this situation.

95.3 Instrumental Variable Approach to Endogeneity

95.3.1 Instrumental Variables and Two-Stage Least Square (2SLS)

The method of instrumental variable (IV) approach provides a general solution to the problem of endogenous explanatory variables. To see how, consider the following setup:

$$Y = X\beta + u \tag{95.3}$$

where Y is $N \times 1$ vector of observations on the dependent variable, X is $N \times K$ matrix of explanatory variables, and u is unobservable mean-zero $N \times 1$ vector of disturbance correlated with some elements of X.

To use the IV approach, we need instruments at least as many as explanatory variables in the model. The instruments should be sufficiently correlated with the endogenous explanatory variables but asymptotically uncorrelated with u, and the explanatory variables that are exogenous can serve as their own instruments as they are uncorrelated with u. Specifically, the valid instruments must satisfy the orthogonality condition and there must be no linear dependencies among the exogenous variables.

In a just-identified setup, using an $N \times K$ matrix Z to instrument for X, the IV estimator can be solved as

$$\beta_{IV} = (Z'X)^{-1}(Z'Y) + (Z'X)^{-1}Z'u \tag{95.4}$$

$$\beta_{IV} = \beta + (Z'X)^{-1}Z'u \tag{95.5}$$

If Z and u are not correlated and $Z'X$ has full rank

$$E(Z'u) = 0 \tag{95.6}$$

$$E(Z'X) = K \tag{95.7}$$

[2]The finance literature using self-selection models has little interest in estimating the endogenous decision itself (the parameter β in Eq. 95.1) but is more interested in using self-selection models to reveal and test for the private information that influences the decision. In contrast, this chapter focuses on how to implement IV approach to estimate the parameter consistently. The readers interested in finance application of self-selection models are referred to Li and Prabhala (2007).

The second term of the right hand side of Eq. 95.5 becomes 0, β_{IV} will be consistent and the unique solution for the true parameters β when the conditions (95.6) and (95.7) hold.

When we have more exogenous variables than needed to identify the parameters, for example, if the instrumental variables Z is an $N \times L$ matrix, where $L > K$, the model is said to be overidentified, and there are $(L - K)$ overidentifying restrictions because $(L - K)$ instruments could be discarded and the parameters can still be identified. If $L < K$, the parameters cannot be identified. Therefore, the order condition $L \geq K$ is necessary for the rank condition, which requires that Z is sufficiently related to X so $Z'K$ has full column rank K. In an overidentifying model, any linear combinations of the instruments can also be used as instruments. Under homoskedasticity, the two-stage least square (2SLS) is the most efficient IV estimator out of all possible linear combinations of the valid instruments since the method of 2SLS chooses the most highly correlated with the endogenous explanatory variable.

The name "two-stage least squares" comes from its two-step procedure:

Step 1: Obtain the fitted value of each endogenous explanatory variable from regressing each endogenous explanatory variable on all instrumental variables. This is called the first-stage regression. Using the matrix notation, the matrix of the fitted values can be expressed as $\hat{X} = Z(Z'Z)^{-1}Z'X$ (if any x_i is exogenous, its fitted value is itself).

Step 2: Run the OLS regression of the dependent variable on the exogenous explanatory variables in the structural Eq. 95.1 and the fitted values obtained from the first-stage regression in place of the observations on the endogenous variables. This is the second-stage regression. The IV/2SLS estimator that uses the instruments $\hat{X}$ can be written as $\hat{\beta} = \left(\hat{X}'X\right)^{-1}\hat{X}'Y$. Substitute $Z(Z'Z)^{-1}Z'X$ for $\hat{X}$; this 2SLS estimator can be written as

$$\hat{\beta} = \left[X'Z(Z'Z)^{-1}Z'X\right]^{-1}X'Z(Z'Z)^{-1}Z'Y \tag{95.8}$$

In practice, it is best to use a software package with 2SLS command rather than carry out the two-step procedure because the reported OLS standard errors of the second-stage regression are not the 2SLS standard errors. The 2SLS residuals and covariance matrix are calculated by the original observations on the explanatory variables instead of the fitted values of the explanatory variables.

In searching valid instrument, both the orthogonality condition and the rank condition are equally important. When an instrument is not fully exogenous and the correlation between the instrument and the explanatory variable is too small, the bias in IV estimator may not be smaller than the bias in OLS estimator. To see this, we compare the bias in the OLS estimator with the bias in the IV estimator for model (95.3):

$$\beta_{OLS} - \beta = Cov(X, u)/Var(X) = \sigma_u \rho_{x,u}/\sigma_x \tag{95.9}$$

$$\beta_{IV} - \beta = \mathrm{Cov}(Z, u)/\mathrm{Cov}(X, Z) = \sigma_u \rho_{z,u}/\sigma_x \rho_{x,z} \tag{95.10}$$

where $\rho_{i,j}$ is the correlation coefficient between variable i and variable j. Thus, the IV estimator has smaller bias than OLS estimator only if $\rho^2_{x,z}$ is larger than $\rho^2_{z,u}/\rho^2_{x,u}$.[3] Although this comparison involves the population correlations with unobservable variable *u* and cannot be estimated directly, it indicates that when the instrument is only weakly correlated to the endogenous regressor, the IV estimator is not likely to improve upon on the OLS (see Bartels 1991 for the discussions of quasi-instrumental variables). Even if the instruments are perfectly exogenous, the low relevance of the instruments can increase asymptotic standard errors and reduce the power of the tests. We can see from Eq. 95.10 that the inconsistency in the IV estimator can get extremely large if the correlation between X and Z gets close to zero and makes the IV estimator undesirable.

The orthogonality of an instrument is difficult to ascertain because we cannot test the correlation between one observable instrument and the unobservable disturbance. But in the case of overidentifying model (i.e., the number of instruments exceeding the number of regressors), the overidentifying restrictions test can be used to evaluate the validity of the additional instruments under the assumption that at least one instrument is valid. We will discuss the tests for validity of the instrument in Sect. 95.4. For checking the rank condition, often we estimate the reduced form for each endogenous explanatory to make sure that at least one of the instruments not in X is significant. If the reduced form regression fits poorly, the model is said to suffer from weak instrument problem and the standard asymptotic theory cannot be employed to make inference. We will discuss the weak instruments problem in Sect. 95.5.

95.3.2 Hypothesis Testing with 2SLS

Testing hypotheses about a single parameter estimate in model (95.3) is straightforward using an asymptotic t-statistic. For testing restrictions on multiple parameters, Wooldridge (2002) provides a method to compute a residual-based F-statistic. Rewrite the Eq. 95.3 into a partitioned model

$$Y = X_1\beta_1 + X_2\beta_2 + u \tag{95.11}$$

where X_1 is $N \times K_1$, X_2 is $N \times K_2$, and $K_1 + K_2 = K$. Let Z denote an $N \times L$ matrix of instruments and assume the rank and the orthogonality conditions hold. Our interest is to test the K_2 restrictions:

$$H_0 : \beta_2 = 0 \text{ against } H_1 : \beta_2 \neq 0 \tag{95.12}$$

In order to calculate the F-statistics for 2SLS, we need to calculate the sum of squared residuals from the restricted second-stage regression, denoted as $\hat{SSR}_r$; the sum of squared residuals from the unrestricted second-stage regression, denoted as $\hat{SSR}_{ur}$; and the sum of squared residuals from unrestricted 2SLS, denoted as SSR_{ur}.

The F-statistics is calculated as

$$F \equiv \frac{\left(S\hat{S}R_r - S\hat{S}R_{ur}\right)}{SSR_{ur}} \cdot \frac{(N-K)}{K_2} \sim F_{K2,\,N-K} \tag{95.13}$$

When the homoskedasticity assumptions cannot be made, we need to calculate the heteroskedasticity-robust standard errors for 2SLS. Some statistical packages compute these standard errors using a simple command. Wooldridge (2002) shows the robust standard errors can be computed in the following steps:

Step 1: Apply 2SLS procedures and obtain the 2SLS residual, denoted $\hat{u}_i$ for each observation i, $i = 1\ldots N$, then use $\hat{u}_i$ to calculate the SSR and $\hat{\sigma}^2$, where $\hat{\sigma}^2 \equiv \frac{SSR}{N-K}$, and $\mathrm{Var}\left(\hat{\beta}_j\right)$, where $\mathrm{Var}\left(\hat{\beta}_j\right) = \hat{\sigma}^2\left(\hat{X}'\hat{X}\right)^{-1}$, and the standard error is denoted as $se\left(\hat{\beta}_j\right)$, $j = 1\ldots K$.

Step 2: Obtain the fitted value of each explanatory variables, denoted as $\hat{x}_{ij}$, $j = 1\ldots K$.

Step 3: Regress each element of $\hat{x}_{ij}$ on all other $\hat{x}_{ik}$ where $k \neq j$, and obtain the residuals from the regressions, denoted as $\hat{r}_{ij}$for each j.

Step 4: Compute the heteroskedasticity-robust standard errors of $\hat{\beta}_j$, denoted as $se_{heter}\left(\hat{\beta}_j\right)$:

$$se_{heter}\left(\hat{\beta}_j\right) = \left[\frac{N}{(N-K)}\right]^{1/2}\left[\frac{se\left(\hat{\beta}_j\right)}{\hat{\sigma}}\right]^2\left[\frac{1}{\hat{m}_j}\right]^{1/2}, \text{where } \hat{m}_j = \sum_{i=1}^{N} \hat{r}_{ij}\hat{u}_i \tag{95.14}$$

95.3.3 Instrumental Variables and Generalized Method of Moments (GMM)

When we have a system of equations, with sufficient instruments we can still apply 2SLS procedure to each single equation in the system to obtain acceptable results. However, in many cases we can obtain more efficient estimators by estimating parameters in the system equations jointly. This system instrumental variables estimation approach is based on the principle of the generalized method of moments (GMM). As discussed in the previous section, the orthogonality conditions require that the valid instruments are uncorrelated to the disturbance, i.e., the population moment condition $E(Z'\mu)$ is equal to zero. Thus, by this principle the optimal parameter estimate is chosen so that the corresponding sample moment is also equal to zero. To show this, reconsider the linear model (95.3) for a random sample from the population, with Z as an $N \times L$ matrix of instruments orthogonal to u for the set of L linear equations in K unknowns and assume the rank condition holds, the sample moment condition must satisfy

$$N^{-1}\sum_{i=1}^{N} Z_i'(Y_i - X_i\hat{\beta}) = 0 \tag{95.15}$$

where the *i* subscript is included for notational clarity. When we have an exactly identified model where $L = K$ and $Z'X$ is invertible, we can solve for the generalized method of moments (GMM) estimator $\hat{\beta}$ in full matrix notation as $\hat{\beta} = (Z'X)^{-1} Z'Y$, which is the same IV estimator obtained in Eq. 95.4. When $L > K$, the model is overidentified; we cannot choose K unknowns to satisfy the L equations and generally the Eq. 95.15 will not have a solution. If we cannot set the sample moment condition exactly equal to zero, we can at least choose the parameter estimate $\hat{\beta}$ so that the vector in Eq. 95.15 is as close to zero as possible. One idea is to minimize the squared Euclidean length of the $L \times 1$ vector in Eq. 95.15 and the optimal GMM estimator $\hat{\beta}$ is the minimizer.

The starting point for GMM estimation is to specify the GMM criterion function $Q\left(\hat{\beta}, Y\right)$, which is the sample moment in Eq. 95.15 in a quadratic form with an $L \times L$ symmetric and positive definite weighting matrix $\hat{W}$:

$$Q\left(\hat{\beta}, Y\right) \equiv \left[\sum_{i=1}^{N} Z_i'\left(Y_i - X_i\hat{\beta}\right)\right]' \hat{W} \left[\sum_{i=1}^{N} Z_i'\left(Y_i - X_i\hat{\beta}\right)\right] \tag{95.16}$$

The GMM estimator $\hat{\beta}$ can be obtained by minimizing the criterion function $Q\left(\hat{\beta}, Y\right)$ over $\hat{\beta}$ and the unique solution in full matrix notation is

$$\hat{\beta} = \left(X'Z\hat{W}Z'X\right)^{-1}\left(X'Z\hat{W}Z'Y\right) \tag{95.17}$$

It can be shown that if the rank condition holds and the chosen weighting matrix $\hat{W}$ is positive definite, the resulting GMM estimator is asymptotically consistent. There is no shortage of such matrix. However, it is important that we choose the weighting matrix that produces the GMM estimator with the smallest asymptotic variance. To do this, first we need to find the asymptotic variance matrix of $\sqrt{N}\left(\hat{\beta} - \beta\right)$. Plug Eq. 95.17 into Eq. 95.3 to get

$$\hat{\beta} = \beta + \left(X'Z\hat{W}Z'X\right)^{-1}\left(X'Z\hat{W}Z'u\right) \tag{95.18}$$

Thus, the asymptotic variance matrix Avar $\sqrt{N}\left(\hat{\beta} - \beta\right)$ can be shown as

$$\begin{aligned} \mathrm{Avar}\sqrt{N}\left(\hat{\beta} - \beta\right) &= \mathrm{Var}\left[(C'WC)^{-1}C'WZ'u\right] \\ &= (C'WC)^{-1}C'WSWC(C'WC)^{-1} \end{aligned} \tag{95.19}$$

where $Z'X \equiv C$, $S \equiv \mathrm{Var}(Z'u)$. It can be shown that if the weighting matrix W is chosen such that $W = S^{-1}$, the GMM estimator has the least asymptotic variance. The details can be found in Hayashi (2000, p. 212). By setting $W \equiv S^{-1}$, the Eq. 95.19 can be simplified as

$$\text{Avar}\,\sqrt{N}\left(\hat{\beta}-\beta\right)=\left(C'S^{-1}C\right)^{-1}=\left(X'ZS^{-1}ZX\right)^{-1} \tag{95.20}$$

Therefore, the efficiency of the GMM estimator depends on if we can consistently estimate S (the variance of the asymptotic distribution of $Z'u$). Since the 2SLS estimator is consistent though not necessarily efficient, it can be used as an initial estimator to obtain the residuals, which in turn can be used as the estimator for S. The two-step efficient GMM procedure is given in Wooldridge (2002) as follows:

Step 1: Use the 2SLS estimator as an initial estimator since it is consistent, denoted as $\widetilde{\beta}$, to obtain the 2SLS residuals by $\widetilde{\mu}_i = Y_i - X_i\widetilde{\beta}$, $i = 1, 2, \ldots, N$. For a system of equations, apply 2SLS equation by equation. This allows the possibility that different instruments are used for different system equations.

Step 2: Estimate S, the variance of the asymptotic distribution of $Z'u$ by

$$\hat{S}=N^{-1}\sum_{i=1}^{N}Z_i'\widetilde{u}_i\widetilde{u}_i'Z_{ii} \tag{95.21}$$

Then use $W=\hat{S}^{-1}$ as the weighting matrix to obtain the efficient GMM estimator. We can plug $\hat{S}^{-1}$ into Eq. 95.17 and obtain the efficient GMM estimator

$$\hat{\beta}=\left(X'Z\hat{S}^{-1}Z'X\right)^{-1}\left(X'Z\hat{S}^{-1}Z'Y\right) \tag{95.22}$$

Following Eq. 95.19, the asymptotic variance matrix $\hat{V}$ of the optimal GMM estimator can be estimated as

$$\hat{V}=\left(X'Z\hat{S}^{-1}ZX\right)^{-1} \tag{95.23}$$

The square roots of diagonal elements of this matrix $\hat{V}$ are the asymptotic standard errors of the optimal GMM estimator.

95.3.4 Hypothesis Testing Using GMM

The t-statistics for the hypothesis testing after GMM estimation can be directly computed by using the asymptotic standard errors obtained from the variance matrix $\hat{V}$. For testing multiple restrictions, the GMM criterion function can be used to calculate the statistic. For example, suppose our interest is to test Q restrictions for the K unknowns in the system; thus, the Wald statistics is a limiting null χ^2_Q. To apply this statistics, we need to assume the optimal weighting matrix $W = S^{-1}$ is chosen to obtain the GMM estimator with and without imposing the Q restrictions. Define the residuals evaluated at the unrestricted GMM estimator $\hat{\beta}_u$ as $\hat{u}_u \equiv Y_i - X_i\hat{\beta}_u$ and the residuals evaluated at the restricted GMM estimator $\hat{\beta}_r$ as $\hat{u}_r \equiv Y_i - X_i\hat{\beta}_r$. By plugging the calculated residuals $\hat{u}_u$ and $\hat{u}_r$ into the criterion function (95.16) for the unrestricted model and the restricted model, respectively, the criterion function statistic is computed as the difference between the two criterion function values, divided by the sample size N. The criterion function statistic has chi-square distribution with Q degrees of freedom.

95.4 Validity of Instrumental Variables

95.4.1 Test for Exogeneity of Instruments

When researchers explore various alternative ways to solve the endogeneity problem and consider the instrumental variables estimation to be the most promising econometric approach, the next step is to find and justify the instruments used. The orthogonality condition (95.6) is the sufficient condition in order for a set of instruments to be valid. When the condition is not satisfied, the IV estimator is inconsistent. In practice, we cannot test whether an observable instrument is uncorrelated with the unobservable disturbance. But when we have more instruments than needed to identify an equation, we can use any subset of the instruments for estimation and the estimates should not be significantly different if the instruments are valid. Following the principle of Hausman (1978), we can test whether the additional instruments are valid by comparing the estimates of an overidentified model with those of a just-identified model.

The model we wish to test is

$$Y = X\beta + u \tag{95.24}$$

where X is a vector of $1 \times K$ explanatory variables with some elements correlated with u, Z is a vector of $1 \times L$ instruments, and $L > K$ so the model has $L - K$ overidentifying restrictions and we can use any $1 \times K$ subset of Z as instruments for X. Let Z_1 be a vector of $1 \times (L - K)$ extra instruments. The overidentifying restrictions require that $E(Z_1{}' u) = 0$. The LM statistic can be computed by regressing the 2SLS residuals from the original model (95.24) on the full set of instruments Z and use the obtained uncentered R^2 times N. The asymptotic distribution of the statistic is χ^2_{L-K}. Another way to test the overidentifying restrictions is to use a test statistic based on the difference between the minimized values of the IV criterion function for the overidentified model and the just-identified model. By the moment condition, the minimized value of the criterion function is equal to zero for the just-identified model. Thus, the test statistic is the minimized value of the criterion function for the overidentified model (95.24), divided by the estimate of the error variance from the same model. This test is called Sargan's test, after Sargan (1958), and numerically the Sargan's test is identical to the LM statistic discussed above.

The usefulness of the overidentifying restrictions test is that if we cannot reject the null, we can have some confidence in the overall set of instruments used. This suggests that it is preferable to have an overidentified model for empirical application, because the overidentifying restrictions can and should always be tested. However, we also need to note that the overidentifying restrictions test is performed under the assumption that at least one instrument is valid. If this assumption does not hold, we will have a situation that all instruments have similar bias and the test will not reject the null. It could also be that the test has low power [4] for detecting the endogeneity of some of the instruments.

The heteroskedasticity-robust test can be computed as follows. Let H be any $1 \times (L - K)$ subset of Z. Regress each element of H onto the fitted value of X and collect the residuals, denoted as $\hat{v}$. The asymptotic χ^2_{L-K} test statistic is obtained as N minus the sum of squared residuals from regressing 1 on $\hat{u}'\hat{v}$.

For the system equations using GMM estimation, the test of overidentifying restrictions can be computed by using a similar criterion function-based procedure. The overidentification test statistic is the minimized criterion function (95.16) evaluated at the optimal (efficient) GMM estimator, and there is no need to be divided by an estimator of error variance because the GMM criterion function takes account of the covariance matrix of the error terms:

$$\left[N^{-1/2}\sum_{i=1}^{N} Z_i'\left(Y_i - X_i\hat{\beta}\right)\right]' \hat{W}\left[N^{-1/2}\sum_{i=1}^{N} Z_i'\left(Y_i - X_i\hat{\beta}\right)\right] \sim \chi^2_{L-K} \tag{95.25}$$

The GMM estimator using the optimal weighting matrix is called the minimum chi-square estimator because the GMM estimator $\hat{\beta}$ is chosen to make the criterion function minimum. If the chosen weighting matrix is not optimal, the expression (95.25) fails to hold. This test statistic is often called Hansen's J-statistic after Hansen (1982) or Hansen-Sargan statistic for its close relationship with the Sargan's test in the IV/2SLS estimation. The Hansen's J-statistic is distributed as chi-square in the number of overidentifying restrictions, $(L - K)$, since K degrees of freedom are lost for having estimated K parameters,[5] and it is consistent in the presence of heteroskedasticity.

It is strongly suggested that an overidentifying restrictions test on instruments should always be performed before formally using the IV estimation, but we also need to be cautious in interpreting the test results. There are several situations in which the null will be rejected. One possibility is that the model is correctly specified, and some of the instruments are indeed correlated with the disturbance and thus are invalid. The other possibility is that the model is not correctly specified, for instance, some variables are omitted from the regression function. In either case, the overidentifying test statistic leads us to reject the null hypothesis but we cannot be certain about which is the case. Another problem is that in small samples, the actual size of the Hansen's J-test far exceeds the nominal size and the test usually rejects too often.

---

[4] If the instruments are only weakly related to the endogenous explanatory variables, the power of the test can be low.

[5] A potential problem is that the test is not consistent against some failures of the orthogonality condition due to the loss of degrees of freedom from K to K-L.

95.4.2 Whether IV Estimator Is Really Needed

In many cases, we may suspect that certain explanatory variables are endogenous but do not know whether the IV estimation is reasonably preferred to the OLS estimation. However, it is important to know that using instrumental variable to solve for the endogeneity problem in empirical research is like a trade-off between efficiency and consistency. As discussed in Bartels (1991), the asymptotic mean square error of the IV estimator can be partitioned into three components:

$$\mathrm{AMSE}\left(\hat{\beta}_{\mathrm{IV}}\right) = \left[\sigma^2{}_{\mathrm{u}}/\mathrm{N}\sigma^2{}_{\mathrm{X}}\right]\left[1 + \left(1 - \sigma^2{}_{\mathrm{XZ}}\right)/\sigma^2{}_{\mathrm{XZ}} + \mathrm{N}\sigma^2{}_{\mathrm{Zu}}/\sigma^2{}_{\mathrm{XZ}}\right] \quad (95.26)$$

where u is the structural error, X is the explanatory variable, Z is the instrumental variable, and N is the sample size. The first term in the second bracket of (95.26) corresponds to the asymptotic variance of the OLS estimator, the second term corresponds to the additional asymptotic variance produced by using IV estimator rather than the OLS estimator, and the third term captures the inconsistency if the instruments are not truly exogenous. Thus, even when we find perfectly exogenous instruments so that $\sigma^2{}_{\mathrm{Zu}} = 0$, the standard error for the IV estimator will exceed the standard error for the OLS estimator by $1/\sigma^2{}_{\mathrm{XZ}}$. As the correlation between X and Z gets close to zero, the standard error for the IV estimator can get extremely large.

Because the IV estimator is less efficient than the OLS estimator, unless we have strong evidence that the explanatory variable is endogenous, the IV estimation is not preferred to the OLS estimation. Therefore, it is always useful to run a test for endogeneity before using the IV approach. This endogeneity test dates back to Durbin (1954), subsequently extended by Wu (1973) and Hausman (1978). Thus, we refer to the test of this type as Durbin-Wu-Hausman tests. To illustrate the procedure, write a population model as

$$\mathrm{Y}_1 = \beta\mathrm{Y}_2 + \gamma\mathrm{Z}_1 + \mathrm{u} \quad (95.27)$$

where Y_1 is the dependent variable, Y_2 is $1 \times \mathrm{K}_2$ vector of the possible endogenous explanatory variables, Z_1 is $1 \times \mathrm{K}_1$ vector of exogenous explanatory variables, and $\mathrm{K}_1 + \mathrm{K}_2 = \mathrm{K}$; also Z_1 is a subset of the $1 \times \mathrm{L}$ exogenous variables Z, assuming Z satisfies the orthogonality condition and $\mathrm{E}(\mathrm{Z}'\mathrm{Y}_2)$ has full rank.

Our interest is to test the null hypothesis that Y_2 is actually exogenous:

$$\begin{aligned} \mathrm{H}_0 &: \mathrm{Y}_1 = \beta\mathrm{Y}_2 + \gamma\mathrm{Z}_1 + \mathrm{u}, \mathrm{E}(\mathrm{Y}_2{}'\mathrm{u}) = 0 \text{ against} \\ \mathrm{H}_1 &: \mathrm{Y}_1 = \beta\mathrm{Y}_2 + \gamma\mathrm{Z}_1 + \mathrm{u}, \mathrm{E}(\mathrm{Y}_2{}'\mathrm{u}) \neq 0 \end{aligned}$$

Under H_0 both IV and the OLS estimators are consistent, and they should not differ significantly. Under H_1 only the IV estimator is consistent. Thus, the original idea of the endogeneity test is to check whether the IV estimator is significantly different from the OLS estimator. The test can also be made by just comparing the estimated coefficients of the parameters of interest, which is $\hat{\beta}$ in our case. The suitable Hausman statistic thus can be calculated as $\left(\hat{\beta}_{IV}-\hat{\beta}_{OLS}\right)'\left[\mathrm{Var}\left(\hat{\beta}_{IV}\right)-\mathrm{Var}\left(\hat{\beta}_{OLS}\right)\right]^{-1}\left(\hat{\beta}_{IV}-\hat{\beta}_{OLS}\right)$.[6]

We often use Hausman (1978) regression-based form of the test, which is easier to compute and asymptotically equivalent to the original endogeneity test. Following the procedure described in Wooldridge (2002), the first step is to regress each element of the K_2 possibly endogenous variables against all exogenous variables Z. The $1 \times K_2$ vector of the obtained reduced form error is denoted as ν. Since Z is uncorrelated to the structural error u and the reduced form error ν, Y_2 is endogenous if and only if u and ν are correlated. To test this, we can project *u* onto ν as

$$u = \lambda\nu + e \tag{95.28}$$

where ν is uncorrelated to *e* and *e* is uncorrelated to Z; thus, the test of whether Y_2 is exogenous is equivalent to test whether the joint test of λ (the K_2 restrictions) is significantly different from 0.

By plugging Eqs. 95.9 into 95.8, we have the equation

$$Y_1 = \beta Y_2 + \gamma Z_1 + \lambda\nu + e \tag{95.29}$$

Since e is uncorrelated to ν, Z_1, and Y_2 by construction, the test of the null hypothesis can be done by using a standard joint F-test on λ in an OLS regression (for the single endogenous variable, a t-statistic can be used in the same procedure), and the F-statistic has the $F(K_2, N\text{-}K\text{-}K_2)$ distribution. If we reject the null H_0: $\lambda = 0$, there is evidence that at least some elements of Y_2 are indeed endogenous. So the use of IV approach is justified assuming the instruments are valid. If the heteroskedasticity is suspected under H_0, the test can be made robust to heteroskedasticity in μ (since $\mu = e$ under H_0) by computing the heteroskedasticity-robust standard errors.[7] Another use of this regression-based form of the test is that the OLS estimates of β and λ for Eq. 95.29 should be identical to their 2SLS estimates for the Eq. 95.27 using Z as the instruments (see Davidson and MacKinnon 1993 for the details), and that allows us to examine whether the differences in the OLS and 2SLS point estimates are practically significant.

[6]If the assumption of homoskedasticity cannot be made, this standard error is invalid because the asymptotic variance of the difference is no longer the difference in asymptotic variances.

[7]Since the robust (Hubert-White) standard errors are asymptotically valid to the presence of heteroskedasticity of unknown form including homoskedasticity, these standard errors are often reported in empirical research especially when the sample size is large. Several statistical packages such as Stata now report these standard errors with a simple command, so it is easy to obtain the heteroskedasticity-robust standard errors.

95.5 Identification and Inferences with Weak Instruments

95.5.1 Problems with Weak Instruments and Diagnosis

As stated in the very beginning of this chapter, a valid instrument should be asymptotically uncorrelated with the structural error (orthogonality) but sufficiently correlated with the endogenous explanatory variable for which it is supposed to serve as instrument (relevance). The relevance condition is critical for the structural model to be identified. If the instruments have little relevance, the instruments will not enter the first-stage regression, then the sampling distribution of IV statistics are generally not normal and the IV estimates and standard tests are unreliable.

A simple way to detect the presence of weak instruments is to look at the first-stage F-statistic for the null hypothesis that the instruments are jointly equal to zero. Staiger and Stock (1997) suggest that the instruments are considered to be weak if the first-stage F-statistic is less than 10. Some empirical studies use R^2 in the first-stage regression as a measure of instrument relevance. However, Shea (1997) argues that for the model with multiple endogenous regressors, when instruments are highly collinear, IV may work poorly even if R^2 is high for each first-stage regression. For instance, suppose both the vector of endogenous regressors X and the vector of instruments Z have rank of 2, while only one element of Z is highly correlated to X. In this situation the regression of each element of X onto Z will have high R^2 even though β in the structural model may not be identified. Instead, Shea (1997) proposes a partial R^2, which measures the instrument relevance by taking intercorrelations among the instruments into account. The idea is that the instruments should work best when the part of instruments important to one endogenous regressor is linearly independent of the part important to the other endogenous regressor. Taking the example above, Shea's partial R^2 is the squared correlation between the component of one endogenous regressor X_1 orthogonal to the other endogenous regressor X_2 (i.e., the residuals from regressing X_1 onto X_2) and the component of the fitted values of X_1 orthogonal to the fitted values of X_2 (i.e., the residuals from regressing the fitted values of X_1 onto the fitted values of X_2). Shea's partial R^2 can be corrected for the degrees of freedom by

$$\overline{R}_p^2 = 1 - [(N-1)/(N-K)] \times \left(1 - R_p^2\right) \tag{95.30}$$

where $\overline{R}_p^2$ is the corrected partial R^2, R_p^2 is uncorrected partial R^2, N is the sample size, and K is the number of exogenous variables in the system.

Another type of the tests for instrument relevance is testing whether the equation is identified. Rothenberg et al. (1984) shows that at a formal level the strength of instruments can be characterized in terms of the so-called concentration parameter associated with the first-stage regression. He considers a simple linear model with a single endogenous regressor X and without included exogenous regressors:

$$Y = X\beta + u \tag{95.31}$$

$$X = Z\prod + v \tag{95.32}$$

where Y and X are N × 1 vectors of observations on endogenous variables, Z is an N × K matrix of instruments, and u and v are N × 1 error vectors. The errors are assumed to be i.i.d. as N(0, $\boldsymbol{\Sigma}$) where σ^2_u, σ^2_{uv}, and σ^2_v are elements of $\boldsymbol{\Sigma}$, and the correlation between the error terms $\rho = \sigma^2_{uv}/(\sigma_u\sigma_v)$. In matrix notation the 2SLS estimator can be written as $\hat{\beta}_{2SLS} = (X'P_ZX)^{-1}(X'P_ZY)$, where the idempotent matrix $P_Z = Z'(Z'Z)^{-1}Z'$. The concentration parameter μ^2 associated with the first-stage reduced form is defined as

$$\mu^2 = \prod{}'Z'Z\prod/\sigma^2_v \tag{95.33}$$

The concentration parameter μ^2 can be interpreted in terms of F, the first-stage F-statistic for testing the hypothesis $\Pi = 0$ (i.e., the instruments do not enter the first-stage regression). Let $\widetilde{F}$ be the infeasible counterpart of F using the true value of σ^2_v. Then $K\widetilde{F}$ has a chi-squared distribution of K degrees of freedom and noncentrality parameter μ^2. When the sample size is large, the $E\left(K\widetilde{F}\right) \cong \mu^2/K + 1$; thus, for large values of μ^2/K, (F − 1) can be used as an estimator for μ^2/K. Rothenberg et al. (1984) shows that the concentration parameter μ^2 plays an important role in the approximation to the distributions of 2SLS estimators and test statistics. He emphasizes that for the normal approximation to the distribution of the 2SLS estimator to be precise, the concentration parameter μ^2 must be large. Thus, a small F-statistic (i.e., a smaller value of μ^2/K) can indicate the presence of weak instruments. There are several tests developed using the concentration matrix to test for weak identification. Cragg and Donald (1993) proposed a statistic using the minimum eigenvalue of the concentration parameter to test the null hypothesis of underidentification, which occurs when the concentration matrix is singular.

Stock and Yogo (2002) argue that when the concentration matrix is nonsingular but its eigenvalues are sufficiently small, the inferences based on conventional normal approximation distributions are misleading even though the parameters might be identified. Thus, the minimum eigenvalues of the concentration matrix μ^2/K can be used to detect the presence of weak instruments. By this principle Stock and Yogo (2002) develop two alternative quantitative definition of the weak instrument. A set of instruments is considered to be weak if μ^2/K is small enough, so the bias in IV estimator to the bias in OLS estimator exceeds a certain threshold, depending on the researcher's tolerance, for example, 10 %. Alternatively, a set of instruments is considered to be weak if μ^2/K is small enough, so the conventional α-level Wald test-based IV statistic has an actual size exceeding a certain threshold, again depending on the researcher's tolerance. Stock and Yogo (2002) propose using the first-stage F-statistic for making inferences about μ^2/K and develop the critical values of F-statistic corresponding to the weak instrument threshold μ^2/K. For example, if a researcher requires the 2SLS relative bias no more than 10 %, for a model with three instruments, the computed first-stage F-statistic has to be larger than 9.08 for the threshold value μ^2/K larger than 3.71, so the null hypothesis that the 2SLS relative bias is less than or equal to 10 % (i.e., instruments are weak) will be rejected. The readers are referred to Stock and Yogo (2002) for a tabulation of the critical values for weak instrument tests with multiple endogenous regressors.

95.5.2 Possible Cures and Inferences with Weak Instruments

Many of the key issues of weak instruments have been studied for decades, however, most of the research on the estimation and inferences robust to weak instruments is quite recent and their applications in finance still remain to be seen. Therefore, this section just simply touches upon this topic and refers the reader to the original articles for details.

In their survey of weak instrument and identification, Stock et al. (2002) considered the Anderson-Rubin statistic as "fully robust" to weak instruments, in the sense that this procedure has the correct size regardless of the value of concentration parameter. For testing the null hypothesis for $\beta = \beta_0$, the Anderson-Rubin statistic (Anderson and Rubin 1949) is computed as

$$\mathrm{AR}(\beta_0) = \frac{(\mathrm{Y} - \mathrm{X}\beta_0)'\mathrm{P_Z}(\mathrm{Y} - \mathrm{X}\beta_0)/\mathrm{K}}{(\mathrm{Y} - \mathrm{X}\beta_0)'\mathrm{M_Z}(\mathrm{Y} - \mathrm{X}\beta_0)/(\mathrm{N} - \mathrm{K})} \sim \mathrm{F_{K,\,N-K}} \tag{95.34}$$

where $\mathrm{P_Z} = \mathrm{Z}(\mathrm{Z'Z})^{-1}\mathrm{Z'}$ and $\mathrm{M_Z} = \mathrm{I} - \mathrm{P_Z}$. Testing the null hypothesis that the coefficients of the endogenous regressors in the structural equation are jointly equal to zero is numerically equivalent to estimating the reduced form of the equation (with the full set of instruments as regressors) and testing that the coefficients of the excluded instruments are jointly equal to zero. Therefore, the AR statistic is often used for testing overidentifying restrictions. The Anderson-Rubin procedure provides valid tests and confidence set under weak instrument asymptotics,[8] but it has low power when too many instruments are added (see Dufour 2003). Berkowitz et al. (2012) argue that when there is a mild violation of the orthogonality condition, the Anderson and Rubin (1949) test may be oversized. In order to correct this problem, the authors fractionally resampled Anderson-Rubin test by modifying Wu's (1990) resampling technique and obtain valid but more conservative critical values.

Other fully robust tests discussed in Stock et al. (2002) include Moreira's conditional test by Moreira (2002), which fixes the size distortion of the test in the presence of weak instruments and can be used to make reliable inference about the coefficients of endogenous variables in the structural equation, and Kleibergen's

[8] Stock et al. (2002) defines weak instrument asymptotics as the alternative asymptotics methods that can be used to analyze IV statistics in the presence of weak instruments. Weak instrument asymptotics involves a sequence of models chosen to keep concentration parameters constant as sample size $\mathrm{N} \to \infty$ and the number of instruments held fixed.

statistic by Kleibergen (2002), which is robust under both conventional and weak instrument asymptotics.

Stock et al. (2002) considered several k-class estimators that are "partially robust" to weak instruments, in the sense that these k-class estimators are more reliable than 2SLS. The k-class is a set of estimators defined by the following estimating equation with an arbitrary scalar k:

$$\hat{\beta}(k) = [X'(I - kM_Z)X]^{-1}[X'(I - kM_Z)Y] \tag{95.35}$$

This class includes 2SLS, LIML, and Fuller-k. 2SLS is a k-class estimator with k equal to 1; LIML is a k-class estimator with k equal to the LIML eigenvalue; Fuller-k or called Fuller's modified LIML, proposed by Fuller (1977), sets $k = k_{LIML} - \alpha/(N - K)$, where K is the total number of instruments and α is the Fuller constant, and the Fuller estimator with $\alpha = 1$ yields unbiased results to second order with fixed instruments and normal errors; Jackknife 2SLS estimator, proposed by Angrist et al. (1999), is asymptotically equivalent to k-class estimator with $k = 1 + K/(N - K)$ (Chao and Swanson 2005). For a discussion of LIML and k-class estimators, see Davidson and MacKinnon (1993).

Stock et al. (2002) found that LIML, Fuller-k, and Jackknife estimators have lower critical value for weak instrument test than 2SLS (so the null will not be rejected too often) and thus are more reliable when the instruments are weak. Anderson et al. (2010, 2011) show that LIML estimator has good performance in terms of the bounded loss functions and probabilities in the presence of many weak instruments. However, Hahn et al. (2004) argued that due to the lack of finite moment, LIML sometimes performs well but sometimes poorly in the weak instrument situation. They found that the interquartile range and the root MSE of the LIML often far exceed those of the 2SLS and hence suggested extreme caution in using the LIML in the presence of weak instrument. Instead, Hahn et al. (2004) recommend the Jackknife 2SLS estimator and Fuller-k estimator because the two estimators do not have the "no moment" problem that LIML has. Theoretical calculations and simulations show that Jackknife 2SLS estimator improves on 2SLS when many instruments are used and thus the weak instrument problem usually occurs (see Chao and Swanson (2005) and Angrist et al. (1999)). Hahn et al. (2004) also find that the bias and mean square error using Fuller-k estimator are smaller than those using 2SLS and LIML.

95.6 Empirical Applications in Corporate Finance

In order to provide some insight into the use of IV estimation by finance researchers, we follow the methodology used by Larcker and Rusticus (2010) to conduct a search using the key words "2SLS," "simultaneous equations," "instrumental variables," and "endogeneity" for papers published in *Journal of Finance*, *Journal of Financial Economics*, *Review of Financial Studies*, and *Journal of Financial and Quantitative Analysis* during the period from 1997 to 2012.

Table 95.1 Finance research that uses instrumental variable methods

Capital structure/leverage ratio	Debt covenants
Faulkender and Petersen (RFS 2006)	Dennis et al. (JFQA 2000)
Yan (JFQA 2006)	Chen et al. (JF 2007)
Molina (JF 2005)	**Macro/product market**
Johnson (RFS 2003)	Thorsten et al. (JFE 2000)
Desai et al. (JF 2004)	Campello (JFE 2006)
Harvey et al. (JFE 2004)	Garmaise (RFS 2008)
Agency/ownership structure/governance	**Financial institutions**
Bitler et al. (JF 2005)	Ljungqvist et al. (JF 2006)
Ortiz-Molina (JFQA 2006)	Berger et al. (JFE 2005)
Palia (RFS 2001)	**Microstructure**
Cho (JFE 1998)	Brown et al. (JF 2008)
Wei et al. (JFQA 2005)	Conrad et al. (JFE 2003)
Daines (JFE 2001)	Kavajecz and Odders-White (RFS 2001)
Coles et al. (JFE 2012)	**Diversification/acquisition**
Wintoki et al. (JFE forthcoming)	Campa and Kedia (JF 2002)
Pricing of public offering	Hsieh and Walkling (JFE 2005)
Cliff and Denis (JF 2004)	**Venture capital/private equity**
Lee and Wahal (JFE 2004)	Gompers and Lerner (JFE 2000)
Lowry and Shu (JFE 2002)	

The sample is based on an electronic search for the term "2SLS," "instrumental variables," "simultaneous equations," and "endogeneity" for papers published in *Journal of Finance*, *Journal of Financial Economics*, *Review of Financial Studies*, and *Journal of Financial and Quantitative Analysis* during the period from 1997 to 2008

As shown in Table 95.1, our search produced 30 published articles that use instrumental variables approach to solve the endogeneity bias in the study of capital structure, agency/ownership structure, pricing of public offering, debt covenants, financial institutions, microstructure, diversification/acquisition, venture capital/ private equity, product market, and macroeconomics. Compared with the survey of Larcker and Rusticus (2010) for the IV applications in accounting research (which found 42 such articles in the recent decade), our list shows that the IV estimation is less commonly used in finance research[9] and often employed in corporate finance-related studies.

Similar to the finding of Larcker and Rusticus (2010) on IV applications in accounting research, there is little attempt in empirical finance research to develop a formal structural equation to identify endogenous and exogenous variables in the first place. Most take the endogeneity in the variables of interest as granted and only

[9]Another reason we found a smaller number of finance papers using IV may be that we limit our keywords to appearing in the abstract. Thus, our data may be more representative of the general situation of the finance research using IV as their main tests.

Table 95.2 Descriptive statistics for finance research that uses instrumental variables methods

A. Types of IV applications
Standard two-stage least squares
Capital structure/leverage ratio (4)
Agency/ownership structure/governance (7)
Pricing of public offering (1)
Debt covenants (2)
Financial institutions (1)
Microstructure (2)
Diversification/acquisition (1)
Total: 18
Two-stage Heckman
Pricing of public offering/debt covenants (2)
Financial institutions (1)
Microstructure (1)
Diversification/acquisition (1)
Venture capital/private equity (1)
Total: 6
GMM/3SLS
Capital structure/leverage ratio/governance (3)
Macro/product market (3)
Total: 6

B. Features of IV application	Two-stage	Heckman	GMM/3SLS	Total
Specification/justification				
Discussion of model/instruments	9 (50 %)	3 (50 %)	4 (66 %)	16 (53 %)
Endogeneity test	3 (17 %)	1 (17 %)	1 (17 %)	5 (17 %)
Reported statistics				
IV relevance				
First-stage regression	7 (39 %)	2 (33 %)	3 (50 %)	12 (40 %)
F-statistic for strength of IV	3 (17 %)	2 (33 %)	3 (50 %)	8 (27 %)
Partial R^2 for IV	0	1 (17 %)	0	
IV orthogonality				
Overidentifying restrictions test	2 (11 %)	0	4 (66 %)	6 (20 %)
Concerns on WI	1 (6 %)	1 (17 %)	2 (33 %)	4 (13 %)
Total	18	6	6	

17 % conducted the endogeneity test (Hausman test). Almost all decided using IV approach to solve the endogeneity problem directly without considering the alternatives. As we discussed in Sects. 95.3.1 and 95.5.1, without valid instruments, the IV estimator can produce even more biased results than the OLS estimator and finding good instruments can be very challenging. Therefore, the researcher should investigate the nature of the endogeneity and explores alternative methods before selecting the IV approach. We found that only Sorensen (2007) and Mitton (2006)

evaluated the situation and chose alternative methods and fixed effects model instead to solve the endogeneity problem for the lack of good instruments.

Table 95.2 shows that the standard two-stage least square regression method is the most commonly used procedure (18 out of 30), and the GMM/3SLS procedure is more concentrated on the study of financial economics. Half of the research did not discuss formally why a specific variable is selected as an instrumental variable, and even fewer explicitly justify theoretically or statistically the validity of selected instruments by examining the orthogonality and relevance of the instruments. All studies have overidentified models but only 20 % performed the overidentifying restrictions test to examine the orthogonality of the instruments. One study even cited the low R^2 in the first stage as evidence that the instruments do not appear to be related to the error terms. It raises a great concern about the validity of the instruments used and the possible bias in the resulted IV estimates.

Forty percent reported the first-stage results along with R^2; however, it should be noted that the first-stage R^2 does not represent the relevance of the excluded instruments but the overall explanatory power of all exogenous variables. Thus, the strength of instruments can be overstated if judged by the first-stage R^2. Only 27 % formally performed the first-stage F-statistic for the instrument relevance, one study also used Shea's partial R^2 (Shea 1997), and none used Cragg and Donald's underidentification test (Cragg and Donald 1993) or Stock and Yogo's weak instrument test Stock and Yogo (2002). Given the absence of the weak instrument test in most finance studies, many of the estimation results using IV approach are questionable. It is also not surprising that only four studies among all addressed the concerns with the weak instrument problem. Ljungqvist et al. (2006) used five bank pressure proxies as instruments for analyst behavior in their two-step Heckman MLE model for competing underwriting mandates but detected the weak instrument problem by the low F-statistic. They interpreted the insignificant estimates in the second step as possibly a result of weak instrument. Beck et al. (2000) discussed the possible weak instrument problem (though not judged by any test) associated with a difference estimator using the lagged value of the dependent variable, and then decided to use the alternative method by estimating the regression in difference jointly with the regression in level to reduce the potential finite sample bias. It appears that the recently developed weak-instrument-robust estimators and inferences have not yet applied in finance research.

95.7 Conclusion

The purpose of this chapter is to present a practical procedure for using the instrumental variables approach to solve the endogeneity problem in empirical finance studies. The endogeneity problem has received a mixed treatment in finance research. The literature does not consistently account for endogeneity using formal econometric methods. When the IV approach is used, the instrumental variables are often chosen arbitrarily and few diagnostic statistics are performed to assess the adequacy of IV estimation.

Given the challenge of finding good instruments, it is important that the researcher analyzes the nature of the endogeneity and the direction of the bias if possible, and then explores alternative empirical approaches so the problem can be solved more appropriately. For example, in the presence of omitted variables, if the unobservable effects, which are part of the error term, can be treated as random variables rather than the parameters to be estimated, panel data models can be used to obtain consistent estimates. When the IV approach is considered to be the most appropriate estimation method, the researcher need to find and justify the instruments theoretically and statistically. One way to describe an instrumental variable is that a valid instrument Z for the potential endogenous variable X should be redundant in the structural equation when X is already included, which means that Z will not affect the dependent variable Y in any way other than through X. To examine the orthogonality statistically, an overidentified model is preferred and hence the overidentifying restrictions can be used. If the OID test cannot be rejected, we then can have some confidence in the orthogonality of the overall instruments. To examine the instrument relevance, the partial R^2 and the first-stage F-statistic on the joint significance of the instruments should be performed at the minimum. The newly developed weak instrument tests (Cragg and Donald 1993; Stock and Yogo 2002) can also be used as robust check, especially for the finite samples. The researcher should keep in mind that the IV estimation method provides a general solution to the endogeneity problem, however; without strong and exogenous instruments, the IV estimator is more biased and inconsistent than the simple OLS estimator.

References

Anderson, T. W., & Rubin, H. (1949). Estimation of the parameters of a single equation in a complete system of stochastic equations. *Annuals of Mathematical Statistics, 20*, 46–63.

Anderson, T. W., Kunitomo, N., & Matsushita, Y. (2010). On the asymptotic optimality of the LIML estimator with possibly many instruments. *Journal of Econometrics, 157*, 191–204.

Anderson, T. W., Kunitomo, N., & Matsushita, Y. (2011). On finite sample properties of alternative estimators of coefficients in a structural equation with many instruments. *Journal of Econometrics, 165*, 58–69.

Angrist, J. D., Imbens, G. W., & Krueger, A. B. (1999). Jackknife instrumental variables estimation. *Journal of Applied Econometrics, 14*, 57–67.

Bartels, L. M. (1991). Instrumental and 'quasi-instrumental' variables. *American Journal of Political Science, 35*, 777–800.

Beck, T., Levine, R., & Loayza, N. (2000). Finance and the sources of growth. *Journal of Financial Economics, 58*, 261–300.

Berger, A. N., Miller, N. H., Petersen, M. A., Rajan, R. G., & Stein, J. C. (2005). Does function follow organizational form? Evidence from the lending practices of large and small banks. *Journal of Financial Economics, 76*, 237–269.

Berkowitz, D., Caner, M., & Ying, F. (2012). The validity of instruments revisited. *Journal of Econometrics, 166*, 255–266.

Bitler, M. P., Moskowitz, T. J., & Vissing-Jorgensen, A. (2005). Testing agency theory with entrepreneur effort and wealth. *Journal of Finance, 60*, 539–576.

Brown, J. R., Ivkovic, Z., Smith, P. A., & Weisbenner, S. (2008). Neighbors matter: Causal community effects and stock market participation. *Journal of Finance, 63*, 1509–1531.

Campa, J. M., & Kedia, S. (2002). Explaining the diversification discount. *Journal of Finance, 57*, 1731–1762.

Campello, M. (2006). Debt financing: Does it boost or hurt firm performance in product markets? *Journal of Financial Economics, 82*, 135–172.

Chao, J. C., & Swanson, N. R. (2005). Consistent estimation with a large number of weak instruments. *Econometrica, 73*, 1673–1692.

Chen, L. D., Lesmond, A., & Wei, J. (2007). Corporate yield spreads and bond liquidity. *Journal of Finance, 62*, 119–149.

Cho, M. (1998). Ownership structure, investment and the corporate value: An empirical analysis. *Journal of Financial Economics, 47*, 103–121.

Cliff, M. T., & Denis, D. J. (2004). Do initial public offering firms purchase analyst coverage with underpricing? *Journal of Finance, 59*, 2871–2901.

Coles, J. L., Lemmon, M. L., & Meschke, F. J. (2012). Structural models and endogeneity in corporate finance: The link between managerial ownership and corporate performance. *Journal of Financial Economics, 103*, 149–168.

Conrad, J., Johnson, K. M., & Wahal, S. (2003). Institutional trading and alternative trading systems. *Journal of Financial Economics, 70*, 99–134.

Cragg, J. G., & Donald, S. G. (1993). Testing identifiability and specification in instrumental variable models. *Econometric Theory, 9*, 222–240.

Daines, R. (2001). Does delaware law improve firm value? *Journal of Financial Economics, 62*, 525–558.

Davidson, R., & MacKinnon, J. G. (1993). *Estimation and inference in econometrics*. New York: Oxford University Press.

Dennis, S., Nandy, D., & Sharpe, I. G. (2000). The determinants of contract terms in bank revolving credit agreements. *Journal of Financial and Quantitative Analysis, 35*, 87–110.

Desai, M. A., Foley, C. F., & Hines, J. R. (2004). A multinational perspective on capital structure choice and internal capital markets. *Journal of Finance, 59*, 2451–2487.

Dufour, J. (2003). Identification, weak instruments and statistical inference in econometrics. *Canadian Journal of Economics, 36*, 767–808.

Durbin, J. (1954). Errors in variables. *Review of the International Statistical Institute, 22*, 23–32.

Faulkender, M., & Petersen, M. A. (2006). Does the source of capital affect capital structure? *Review of Financial Studies, 19*, 45–79.

Fuller, W. A. (1977). Some properties of a modification of the limited information estimator. *Econometrica, 45*, 939–953.

Garmaise, M. J. (2008). Production in entrepreneurial firms: The effects of financial constraints on labor and capital. *Review of Financial Studies, 21*, 543–577.

Gompers, P., & Lerner, J. (2000). Money chasing deals? The impact of fund inflows on private equity valuations. *Journal of Financial Economics, 55*, 281–325.

Hahn, J., Hausman, J., & Kuersteiner, G. (2004). Estimation with weak instruments: Accuracy of higher-order bias and MSE approximation. *Econometrics Journal, 7*, 272–306.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica, 50*, 1029–1055.

Harvey, C. R., Lins, K. V., & Roper, A. H. (2004). The effect of capital structure when expected agency costs are extreme. *Journal of Financial Economics, 74*, 3–30.

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica, 46*, 1251–1271.

Hayashi, F. (2000). *Econometrics*. Princeton: Princeton University Press.

Heckman, J. (1979). Sample selection as a specification error. *Econometrica, 47*, 1531–161.

Hermalin, B. E., & Weisbach, M. S. (2003). Boards of directors as an endogenously determined institution: A survey of the economic literature. *Federal Reserve Bank of New York Economic Policy Review, 9*, 7–26.

Hsieh, J., & Walkling, R. A. (2005). Determinants and implications of arbitrage holdings in acquisitions. *Journal of Financial Economics, 77*, 605–648.

Johnson, S. A. (2003). Debt maturity and the effects of growth opportunities and liquidity risk on leverage. *Review of Financial Studies, 16*, 209–236.

Kavajecz, K. A., & Odders-White, E. R. (2001). An examination of changes in specialists' posted price schedules. *Review of Financial Studies, 14*, 681–704.

Kleibergen, F. (2002). Pivotal statistics for testing structural parameters in instrumental variables regression. *Econometrica, 70*, 1781–1803.

Larcker, D. F., & Rusticus, T. O. (2010). On the use of instrumental variables in accounting research. *Journal of Accounting and Economics, 49*, 186–205.

Li, K., & Prabhala, N.R. (2007). Self-Selection models in corporate finance. In: *Handbook of Corporate finance: Empirical corporate finance*, Handbooks in finance series (Chapter 2). Amsterdam: Elsevier/North Holland.

Lee, P. M., & Wahal, S. (2004). Grandstanding, certification and the underpricing of venture capital backed IPOs. *Journal of Financial Economics, 73*, 375–407.

Ljungqvist, A., Marston, F., & Wilhelm, W. J. (2006). Competing for securities underwriting mandates: Banking relationships and analyst recommendations. *Journal of Finance, 61*, 301–340.

Lowry, M., & Shu, S. (2002). Litigation risk and IPO underpricing. *Journal of Financial Economics, 65*, 309–335.

Mitton, T. (2006). Stock market liberalization and operating performance at the firm level. *Journal of Financial Economics, 81*, 625–647.

Molina, C. A. (2005). Are firms underleveraged? An examination of the effect of leverage on default probabilities. *Journal of Finance, 60*, 1427–1459.

Moreira, M. J. (2002). *Tests with correct size in the simultaneous equations model*. Berkeley: University of California.

Ortiz-Molina, H. (2006). Top management incentives and the pricing of corporate public debt. *Journal of Financial and Quantitative Analysis, 41*, 317–340.

Palia, D. (2001). The endogeneity of managerial compensation in firm valuation: A solution. *Review of Financial Studies, 14*, 735–764.

Rothenberg, T. J., Griliches, Z., & Intriligator, M. D. (1984). Approximating the distributions of econometric estimators and test statistics In: *Handbook of econometrics: Vol. II. Handbooks in economics series*. Amsterdam: Elsevier/North Holland.

Sargan, J. D. (1958). The estimation of economic relationships using instrumental variables. *Econometrica, 26*, 393–415.

Shea, J. (1997). Instrument relevance in multivariate linear models: A simple measure. *Review of Economics and Statistics, 79*, 348–352.

Sorensen, M. (2007). How smart is smart money? A two-sided matching model of venture capital. *The Journal of Finance, 62*, 2725–2762.

Staiger, D., & Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica, 65*, 557–586.

Stock, J. H., Wright, J. H., & Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business and Economic Statistics, 20*, 518–529.

Stock, J. H., & Yogo, M. (2002). *Testing for weak instruments in linear IV regression* (National Bureau of Economic Research (NBER) Technical Working Paper 284). Cambridge, MA: National Bureau of Economic Research.

Thorsten, B., Ross, L., & Loayza, N. (2000). Finance and the sources of growth. *Journal of Financial Economics, 58*, 261–300.

Wei, Z., Xie, F., & Zhang, S. (2005). Ownership structure and firm value in China's privatized firms: 1991–2001. *Journal of Financial and Quantitative Analysis, 40*, 87–108.

Wintoki, M. B., Linck, J. S., & Netter, J. M. (2012). Endogeneity and the dynamics of internal corporate governance. *Journal of Financial Economics, 105*, 581–606.

Wooldridge, J. M. (2002). *Econometric analysis of cross section and panel data*. Cambridge, MA: MIT Press.

Wu, D. (1973). Alternative tests of independence between stochastic regressors and disturbances. *Econometrica, 41*, 733–750.

Wu, C. (1990). On the asymptotic properties of the jackknife histogram. *Annals of Statistics, 18*, 1438–1452.

Yan, A. (2006). Leasing and debt financing: Substitutes or complements? *Journal of Financial and Quantitative Analysis, 41*, 709–731.

Application of Poisson Mixtures in the Estimation of Probability of Informed Trading

96

Emily Lin and Cheng-Few Lee

Contents

Abstract

This research first discusses the evolution of probability for informed trading in finance literature. Motivated by asymmetric effects, e.g., return and trading volume in up and down markets, this study modifies a mixture of the Poisson

E. Lin (✉)
St. John's University, New Taipei City, Taiwan
e-mail: mlin@mail.sju.edu.tw

C.-F. Lee
Department of Finance and Economics, Rutgers Business School, Rutgers, The State University of New Jersey, Piscataway, NJ, USA

Graduate Institute of Finance, National Chiao Tung University, Hsinchu, Taiwan
e-mail: lee@business.rutgers.edu; cflee@business.rutgers.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_96,

distribution model by different arrival rates of informe d buys and sells to measure the probability of informed trading proposed by Easley et al. (Journal of Finance 51:1405–1436, 1996).

By applying the expectation–maximization (EM) algorithm to estimate the parameters of the model, we derive a set of equations for maximum likelihood estimation, and these equations are encoded in a SAS Macro utilizing SAS/IML for implementation of the methodology.

Keywords

Probability of informed trading (*PIN*) • Expectation–maximization (EM) algorithm • A mixture of Poisson distribution • Asset-pricing returns • Order imbalance • Information asymmetry • Bid–ask spreads • Market microstructure • Trade direction • Errors in variables • GARCH

96.1 Introduction

This study investigates the probability of informed trading *PIN* which is widely used in existing literature and is introduced by Easley et al. (1996). Easley et al. (1996, 2002, 2008) have proposed the same arrival rate of informed orders μ for both bad and good events, and the likelihood is given by

$$L(\theta|B,S) = (1-\alpha)e^{-\varepsilon_b}\frac{\varepsilon_b^B}{B!}e^{-\varepsilon_s}\frac{\varepsilon_s^S}{S!} + \alpha\delta e^{-\varepsilon_b}\frac{\varepsilon_b^B}{B!}e^{-(\mu+\varepsilon_s)}\frac{(\mu+\varepsilon_s)^S}{S!} + \alpha(1-\delta)e^{-(\mu+\varepsilon_b)}\frac{(\mu+\varepsilon_b)^B}{B!}e^{-\varepsilon_s}\frac{\varepsilon_s^S}{S!} \tag{96.1}$$

where α is the probability of new information, δ is the probability that new information is bad news, μ is the arrival rate of informed buy orders and also that of informed sell orders, and ε_b and ε_s are the arrival rates of uninformed buyers and sellers. Until Easley, Engle, O'Hara and Wu (2008) models a time-varying arrival rate of informed and uninformed traders, the model has been a static approach.

We allow the arrival rate of informed buyers to be different from that of informed sellers in order to match the empirical environment. Furthermore, we examine on intraday data and allow more than one informational event per day. The modified model is given by

$$L(\theta|B,S) = (1-\alpha)e^{-\varepsilon_b}\frac{\varepsilon_b^B}{B!}e^{-\varepsilon_s}\frac{\varepsilon_s^S}{S!} + \alpha\delta e^{-\varepsilon_b}\frac{\varepsilon_b^B}{B!}e^{-(\mu_s+\varepsilon_s)}\frac{(\mu_s+\varepsilon_s)^S}{S!} + \alpha(1-\delta)e^{-(\mu_b+\varepsilon_b)}\frac{(\mu_b+\varepsilon_b)^B}{B!}e^{-\varepsilon_s}\frac{\varepsilon_s^S}{S!} \tag{96.2}$$

where μ_b is the arrival rate of informed buyers and μ_s is the arrival rate of informed sellers. The function provides the structure necessary to exact information on the parameters $\theta = (\alpha, \delta, \mu, \varepsilon_b, \varepsilon_s)$ from the observable variables, buys and sells, to measure $PIN = \frac{\alpha\mu}{\alpha\mu + \varepsilon_s + \varepsilon_b}$. For the parameters in our model, $\theta^* = (\alpha, \delta, \mu_b, \mu_s, \varepsilon_b, \varepsilon_s)$ and probability of informed trading $PIN^* = \frac{\alpha(\delta|\mu_s| + (1-\delta)|\mu_b|)}{\alpha(\delta|\mu_s| + (1-\delta)|\mu_b|) + \varepsilon_s + \varepsilon_b}$. The buys and sells follow one of three Poisson processes on each day. The likelihood of observing any sequence of orders that contains B buys and S sells on a no-event day is given by

$$e^{-\varepsilon_b} \frac{\varepsilon_b^B}{B!} e^{-\varepsilon_s} \frac{\varepsilon_s^S}{S!} \tag{96.3}$$

Similarly, on a bad-event day, the likelihood of observing any sequence of orders that contains B buys and S sells is

$$e^{-\varepsilon_b} \frac{\varepsilon_b^B}{B!} e^{-(\mu + \varepsilon_s)} \frac{(\mu + \varepsilon_s)^S}{S!} \tag{96.4}$$

Finally, on a good-event day, this likelihood is

$$e^{-(\mu + \varepsilon_b)} \frac{(\mu + \varepsilon_b)^B}{B!} e^{-\varepsilon_s} \frac{\varepsilon_s^S}{S!} \tag{96.5}$$

To estimate the order arrival rates of the buy and sell processes, we need only consider the total number of buys, B, and the total number of sells, S, on any day. The likelihood of observing B buys and S sells on a day of unknown type is a mixture of the Poisson distribution, the weighted average of Eqs. 96.3, 96.4, and 96.5 using the probabilities of each type of day occurring to obtain Eq. 96.1.

Church and Gale (1995) claim that a mixture of the Poisson distribution fits the data better than the standard Poisson, producing more accurate estimates of the variance. Johnson and Kotz (1969, pp. 135–136) survey a number of applications of the negative binomial in a variety of fields and conclude that "the negative binomial is frequently used as a substitute for the Poisson, when the strict requirements of the Poisson is doubtful." This is due to the negative binomial, which can be viewed as a continuous mixture of infinitely many Poissons, as suggested by Bookstein and Swanson (1974, p. 317).

Because days are independent, the likelihood of observing the data $M = (B_i, S_i)_{i=1}^I$ over I days is just the product of the daily likelihoods. Therefore,

$$L(\theta|M) = \prod_i^I L(\theta|B_i, S_i), \tag{96.6}$$

To estimate the parameter vector θ from any data set M, we maximize the likelihood defined in Eq. 96.6.

96.2 The Problem with *PIN*

The derivation of *PIN* requires the classification of trades into buyer- or seller-initiated trades, and therefore errors can occur by possible misclassification. Ellis et al. (2000) present that in the case of NASDAQ trades, the Lee and Ready (1991) trade classification algorithm correctly classifies 81.05 % of the trades, with the lowest rate of success among trades that take place inside the spread. Specifically, the authors admit that "the success rate for classifying trades inside the quotes . . . is substantially lower, falling to approximately 60 % for midpoint trades and to only 55 % for trades that are inside the quotes but not at the midpoint." In the case of NYSE trades, Odders-White (2000) reports a success rate of 85 % for the entire sample. Boehmer et al. (2007) and Lei and Wu (2005) also point out that misappropriation of trades may bias *PIN* estimates and arrival rates may not be symmetric and time varying.

To avoid this criticism or error, Popescu and Kumar (2008) use observed bid and ask quotes, assume different depth at the bid and the ask, and include order processing costs for estimating the probability of informed trading by extending the model developed by Copeland and Galai (1983),[1] which is the first model to examine intraday informed trading under an option framework. The measure of Popescu and Kumar (2008) can be computed at any point in time and thus can be used to estimate changes in the level of information asymmetry over a short interval. Although this methodology does not require trades to be distinguished between buyer initiated and seller initiated, the estimation of order processing costs based on three bid–ask spread structure models may introduce bias on the estimated *PIN*. These three models include Glosten and Harris (1988) and Madhavan and Smidt (1991) which model revision in trade price and a proposed model which combines Hasbrouck (1991), Foster and Viswanathan (1993), and Brennan and Subrahmanyam (1996) to model transaction size and price revision.

Copeland and Galai (1983) extend Bagehot (1971) which employs bid–ask spread to derive an adverse selection cost, use an option idea to describe the quote spread of market maker to contain an option value, and consider the trade value of straddle strategy as an adverse selection cost. Bollen et al. (2004) also follow the same strategy by Copeland and Galai (1983) to measure an adverse selection cost. Because the bid–ask spread carries unrealized price information and hides future value, it seems more reasonable to observe the bid–ask spread from the idea of an option.

[1]Copeland and Galai (1983) model informed trading as $(1 - P_I)[P_{BL}(A - S_0) + P_{SL}(B - S_0) + P_{NL}.0]$, where B is the bid price, A is the ask price, S_0 is the dealer's estimate of the "true" value of a security ($B < S_0 < A$), and P_I is the probability that the next trade originates from an informed trader, while P_{BL}, P_{SL}, and P_{NL} are the conditional probabilities that the next liquidity trader will buy, sell, or not trade when he/she faces the market maker. Popescu and Kumar (2008) revise it into $(1 - P_I)[P_{BL}.\, D_A.\, (A - S_0) + P_{SL}.\, D_B.\, (B - S_0) + P_{NL}.0]$, where D_A and D_B denote the depth at the ask and at the bid.

Based on the setting of Easley et al. (1996, 2002, 2008), the informed order imbalance is unavailable. We attain informed order imbalance in November 2008 and our work is the first to measure informed trading based on an order imbalance signal. Although not fully explored here, this measure allows one to measure informed order imbalance by $(\mu_b - \mu_s)/(\mu_b + \mu_s)$. The measure is a proxy for informed trading and is discussed in Lin et al. (2013), while the relationship between *PIN* and arbitrage opportunity is determined in Chang and Lin (2014). Like *PIN*, this measure is an estimate variable, and so it is potentially subject to errors-in-variables bias. To correct the errors-in-variables problem in *PIN*, Easley et al. (2002) suggest to create an instrument variable to use in place of the variable in question. Lee and Chen (2012) include five other methods[2] which can correct this bias. Duarte and Young (2009) also allow the arrival rate of informed buyers μ_b to be different from that of informed sellers μ_s and overcome the estimation dilemma of standard *PIN*. In addition, Duarte and Young (2009) apply a time-varying technique to examine whether *PIN* is priced and includes symmetric order-flow shocks to capture the positive correlation they find between buys and sells. Chang and Lin (2014) ignore to do so as they observe no significant correlation between buys and sells in the data.

Easley et al. (2012) take a similar approach with a time-varying technique, easing the estimation of *PIN* in high-volume markets and referring it to *VPIN*. Easley et al. (2012) claim *VPIN* is updated in volume time and does not require the intermediate estimation of nonobservable parameters or the application of numerical methods.

Nevertheless, Easley et al. (2002) show that *PIN* as a proxy for the risk of informed trading is priced. The results of Easley et al. (2002) provide evidence that information plays a deeper role beyond what is captured in spreads. Duarte and Young (2009) propose a model that decomposes *PIN* into two components, one related to asymmetric information and one related to illiquidity. On the contrary, they find the *PIN* component related to asymmetric information is not priced, while the *PIN* component related to illiquidity is priced. This contrary finding makes itself an open question for finance researchers.

96.3 The Estimation Methodology

To solve the likelihood function in Eqs. 96.1 and 96.2, we apply an expectation–maximization (EM) algorithm. In statistics, an expectation–maximization (EM) algorithm is an iterative method for finding a maximum likelihood or maximum posteriori estimates of parameters in statistical models, where the model depends on unobserved latent variables. The EM iteration alternates between

[2]The other five methods are (i) classical estimation method: either unconstrained or constrained type, (ii) grouping method, (iii) mathematical programming method, (iv) maximum likelihood method, and (v) LISREL and MIMIC methods.

performing an expectation (E) step, which creates a function for the expectation of the log-likelihood evaluated using the current estimate for the parameters, and a maximization (M) step, which computes parameters maximizing the expected log-likelihood found in step E. These parameter estimates are then used to determine the distribution of the latent variables in the next E step. We can derive a set of equations for maximum likelihood estimation when the observed data consists of complete pairs. These equations are encoded in a SAS Macro utilizing SAS/IML for implementation of the methodology. McLachlan and Krishnan (2008) have discussed this algorithm in detail. In addition, Appendix 2 has presented the estimation procedure of *PIN*.

96.4 Empirical Results

It is *PIN* that affects asset-pricing return consistent with economic analysis motivates the empirical work in Chang and Lin (2014) to explore how *PIN* under various cross-section and time-series sample splits is related to cash-futures basis, defined as futures price minus stock price. *PIN* is calculated as per Eq. 96.1 or 96.2 to test this relationship. The resulting higher significant regression coefficient of *PIN* derived from different arrival rates of informed trades at buy side and sell side confirms a conjecture that the revised *PIN* appears to capture the asymmetric information in cash-futures basis spread better than does standard *PIN*. To deeply understand the difference of these two *PIN* measures, this study furthermore compares the distribution of the parameters in each model. The comparison estimates the parameters of the two models using 5-min intraday data sourced from Taiwan index futures and Taiwan stock index markets. To appropriate a trade direction, two general approaches are used to infer the direction of a trade: (1) compare the trade price to the bid/ask prices of the prevailing quote or (2) compare the trade price to adjacent trades (the techniques commonly known as "tick tests"). In this study, the algorithm of Ready and Lee (1991) is used for classifying index futures data, while tick test[3] is used for stock index data because of lack of quote price at market level.

96.4.1 Preliminary Results

Both Tables 96.1 and 96.2 contain time-series averages from March 24, 1999, through September 22, 2005, of means, medians, standard deviations, and the

[3]Tick test classifies each trade into four categories: an uptick, a downtick, a zero-uptick, and a zero-downtick. A trade is an uptick (downtick) if the price is higher (lower) than the price of the previous trade. When the price is the same as the previous trade (a zero tick) and if the last price was an uptick, then the trade is a zero-uptick. Similarly, if the last price change was a downtick, then the trade is a zero-downtick.

Table 96.1 Parameter summary statistics in Taiwan index futures market

Parameter	Mean	Median	Standard deviation	Median standard error
α	0.416	0.408	0.126	0.003
	(0.601)	(0.598)	(0.178)	(0.004)
δ	0.514	0.516	0.197	0.005
	(0.546)	(0.561)	(0.204)	(0.005)
μ	173	160.352	109.167	2.692
μ_b	(147.437)	(146.797)	(132.094)	(3.257)
μ_s	(27.419)	(66.203)	(171.272)	(4.223)
ε_b	75.983	69.405	45.941	1.133
	(77.256)	(70.914)	(42.375)	(1.045)
ε_s	81.688	76.936	47.363	1.168
	(129.918)	(84.524)	(119.959)	(2.958)
PIN	0.300	0.296	0.067	0.002
	(0.305)	(0.299)	(0.064)	(0.002)

Table 96.2 Parameter summary statistics in Taiwan stock index market

Parameter	Mean	Median	Standard deviation	Median standard error
α	0.651	0.655	0.085	0.002
	(0.658)	(0.661)	(0.074)	(0.002)
δ	0.517	0.520	0.115	0.003
	(0.517)	(0.517)	(0.108)	(0.003)
μ	903,591	843,264	352,935	8,689
μ_b	(892,352)	(833,270)	(356,513)	(8,776)
μ_s	(907,195)	(838,015)	(373,210)	(9,187)
ε_b	579,489	536,094	219,850	5,412
	(577,746)	(535,663)	(221,193)	(5,445)
ε_s	577,863	536,838	216,781	5,337
	(574,816)	(536,306)	(215,968)	(5,316)
PIN	0.334	0.332	0.047	0.001
	(0.337)	(0.334)	(0.042)	(0.001)

median of parameter standard errors from the likelihood estimation. Parameters as per Eqs. 96.1 and 96.2 are estimated for the index futures and stock index in the Taiwan market. We achieve the parameters for the Taiwan index futures market $\hat{\theta}_{fut} = \left(\hat{\alpha}, \hat{\delta}, \hat{\mu}, \hat{\varepsilon}_b, \hat{\varepsilon}_s\right) = (0.42, 0.51, 173, 76, 82)$ and $\widehat{PIN}_{fut} = 0.2997$ for the case of the same arrival rate of informed trades, μ_{fut}, at the buy side and sell side. Meanwhile, we achieve $\hat{\theta}^*_{fut} = \left(\hat{\alpha}, \hat{\delta}, \hat{\mu}_b, \hat{\mu}_s, \hat{\varepsilon}_b, \hat{\varepsilon}_s\right) = (0.60, 0.55, 147, 27, 77, 129)$ and $\widehat{PIN}^*_{fut} = 0.3046$ for the case of different arrival rates of informed trades, μ_b^{fut} and μ_s^{fut}, at buy side and sell side. There are 1,005 out of 1,645 days (61 %) that the value of μ^{fut} is between μ_b^{fut} and μ_s^{fut}.

Similarly, we achieve the parameters for the Taiwan stock index $\hat{\theta}_{ind} = \left(\hat{\alpha}, \hat{\delta}, \hat{\mu}, \hat{\varepsilon}_b, \hat{\varepsilon}_s\right) = (0.65, 0.52, 903591, 579489, 577863)$ and $\widehat{PIN}_{ind} = 0.3343$ for the case of the same μ^{ind}, and meanwhile we obtain $\hat{\theta}^*_{ind} = \left(\hat{\alpha}, \hat{\delta}, \hat{\mu}_b, \hat{\mu}_s, \hat{\varepsilon}_b, \hat{\varepsilon}_s\right) = (0.66, 0.52, 892352, 907195, 577746, 574816)$ and $\widehat{PIN}^*_{ind} = 0.3366$ for the case of different arrival rates, μ_b^{ind} and μ_s^{ind}. There are 1,637 out of 1,650 days (99 %) that the value of μ^{ind} is between μ_b^{ind} and μ_s^{ind}.

In Table 96.1, each parameter in the two models has a similar distribution, except μ and the arrival rate of uninformed sell orders, ε_s, while in Table 96.2, there is no such exception for the distribution of each parameter in the two models. The *PIN* value estimated from each model is similar for both index futures and stock index markets. The differential results of these two models may result from volatile data or order imbalance between buys and sells. We observe that the Taiwan index futures behave much more volatile than the Taiwan stock index. The model of Easley et al. assumes a sole informed arrival rate for buys on a good-event day and for sells on a bad-event day. The assumption might fit an individual stock level better than a market level because information tends to be just good or bad for an individual stock than for a market. Meanwhile, the model of Easley et al. appears more appropriate for a stable or a less order imbalance market.

96.4.2 Application of PIN

Maturity effect may also play an information role in the market by helping the market incorporate certain types of information into prices. In Table 96.3, *PIN*, as derived from differential arrival rate of informed trades for index futures denoted by pin(f) and stock index denoted by pin(s), is calculated to explore how they

Table 96.3 Asymmetry of maturity effect

Panel A. Volume effect			
Dependent variables	A/F 2000/10/24	B/F 2000/10/24	P-value
Rhat	1.5039	1.2485	0.0004
pin(f)	0.2960	0.2856	<0.0001
pin(s)	0.3290	0.3540	<0.0001
Panel B. Option introduction			
Dependent variables	A/F 2001/12/24	B/F 2001/12/24	P-value
Rhat	1.3697	1.2499	0.0364
pin(f)	0.2869	0.3075	<0.0001
pin(s)	0.3318	0.3359	<0.0001

This table compares sample split means using 5-min data. Rhat is the ratio of standard deviation for index futures return over that for the spot index return calculated by $\sigma_{fut}/\sigma_{ind}$. Panel A investigates the day futures volume reached 10,000 contracts for the first time and Panel B examines the day options were introduced to the market, respectively. Results reflect a testing period from March 1999 to September 2005. P-values reflect the significance of the difference in means derived from t-tests evolve before and after the day index futures volume reached 10,000 contracts for the first time and before and after the introduction of index futures options. *Rhat* is the ratio of standard deviation for index futures return over standard deviation for the spot index return. MacKinlay and Ramaswamy (1988) find that no arbitrage implies equal volatility in the spot and futures markets. Therefore, the standard deviation ratio *Rhat* should be one if arbitrage is tightly enforced. As shown in Table 96.3, *Rhat* is generally higher than one for both markets indicating that futures are more volatile than spot. In addition, *Rhat* increases after volume effect and option introduction which signals market maturity. Increased pin(f) and smaller pin(s) suggest volume effect facilitates price discovery in the futures market, and this improvement seems to be switched from the spot market. Similarly, reduced pin(f) and pin(s), in particular pin(f), indicate option introduction induces much of the price discovery moved from the futures market to the option market. The occurrence of price discovery in the option market may result from the higher leverage trait of option trading.

96.5 Conclusion

The study focuses on comparing the estimation of probability of informed trading for the model of Easley et al. and of our extension. The study introduces mixtures of Poisson distribution, reports the evolution of *PIN* development, summarizes the problems with *PIN*, and estimates *PIN* by EM algorithm. By using Taiwan stock index and Taiwan index futures, this study examines the distribution of the two model parameters and the price discovery effect before and after futures volume effect and option introduction. The study finds *PIN* calculated by Easley et al.'s model is similar to the *PIN* calculated by different arrival rate of informed trading. This evidences Easley et al.'s simplified assumption has no loss of generalization for obtaining *PIN*, whereas the assumption of ours better fits volatile data and allows us to capture order imbalance of informed trading.

Appendix 1: Poisson Mixtures

Generalized Poisson Mixtures

Poisson mixtures can be thought of as a generalization of two Poisson models where the mixing parameter, a, is replaced with an arbitrary density function, f. The density function f is intended to capture dependencies on hidden variables. The general form of a Poisson mixture is

$$\Pr(x) = \int_0^\infty \phi(\omega)\pi(\omega, x)d\omega \quad \text{for } x = 0, 1, \ldots$$

where π is a Poisson

$$\pi(\omega, k) = \frac{e^{-\theta}\omega^k}{k!} \quad \text{for } k = 0, 1, \ldots$$

and ϕ is an arbitrary density function. A density function should integrate to 1. That is, $\int_0^\infty \phi(\omega)d\omega = 1.$

Binomial Poisson Mixtures

We give an application of Poisson mixtures, e.g., binomial Poisson mixture as follows:

Generate $X_1, X_2, \ldots, X_n$, i.i.d. from

$$\Pr(X_i = k) = 1 - P + P\frac{e^{-\lambda}\lambda^k}{k!}, k = 0, 1, \ldots$$

That is, the random variable X_i has a probability P of being from a Poisson distribution with $\lambda > 0$ and has a probability $1 - P$ of being zero, $0 < P < 1$. When $k = 0$, it degenerates to a Poisson distribution. When $k > 0$, it is a binomial 1, otherwise 0.

Appendix 2: Estimation by EM Algorithm

To write a likelihood function by a framework of mixtures of Poisson distribution, we define, for each $i \in I$, a vector Z_i by $Z_i = (Z_{i1}, Z_{i2}, Z_{i3}) = (1,0,0)$, $(0,1,0)$, or $(0,0,1)$ depending if i has either no event, a good event, or a bad event, respectively.

$$\sum_{j=1}^{3} Z_{ij} = 1, \text{for each } i.$$

$$M^* = (B_i, S_i, Z_i)_{i=1}^{I} \sim \prod_{j=1}^{3} P_j^{Z_{ij}} f_j^{Z_{ij}}(B_i, S_i), \text{then}$$

$$L(\theta|B_i, S_i, Z_i) \sim \prod_{j=1}^{3} P_j^{Z_{ij}} f_j^{Z_{ij}}(B_i, S_i),$$

where $P_1 = 1 - \alpha$, $P_2 = \alpha\delta$, $P_3 = \alpha(1 - \delta)$, P_2 and P_3 are independent, $P_1 = 1 - P_2 - P_3, Z_{ij}, i$ is for date i, state j, and $f_j^{Z_{ij}}$ equals 1 when $Z_{ij} = 0$, while $f_j^{Z_{ij}}$ is one of Eqs. 96.3, 96.4, and 96.5 as $Z_{ij} = 1$.

Let $V = L(\theta|M^*) = \prod_{i=1}^{I} L(\theta|B_i, S_i, Z_i)$, then $V(B,S,Z) \sim \prod_{\mathrm{i}=1}^{I}\prod_{j=1}^{3} P_j^{Z_{ij}} f_j^{Z_{ij}}(B_i,S_i)$.

By taking logarithm and letting $\ell = \ln(V|B,S) = \sum_{i=1}^{I}\sum_{j=1}^{3}\left\{Z_{ij}\ \ln P_j + Z_{ij}\ \ln f_j(B_i,S_i)\right\}$, we start to perform E-step of EM algorithm.

Assume

$$Q = E(\ln\, V|B,S) = \sum_{i=1}^{I}\sum_{j=1}^{3}\left\{E\left(Z_{ij}|B_i,S_i\right)\ln P_j + E\left(Z_{ij}|B_i,S_i\right)\ln\, f_j(B_i,S_i)\right\}, \tag{96.7}$$

where

$$\begin{aligned} E\left(Z_{ij}|B_i,S_i\right) = P\left(Z_{ij}=1|B_i,S_i\right) &= \frac{P\left(Z_{ij}=1, B_i=b, S_i=s\right)}{P(B_i=b,S_i=s)} \\ &= \frac{P_j f_j(b,s)}{P_1 f_1(b,s) + P_2 f_2(b,s) + P_3 f_3(b,s)}. \end{aligned}$$

Then r-step estimator is given by

$$\begin{aligned} Q = &\sum_{i=1}^{I}\sum_{j=1}^{3}\left\{\frac{P_j f_j(b,s)}{P_1 f_1(b,s) + P_2 f_2(b,s) + P_3 f_3(b,s)}\ln P_j \right. \\ &\left. + \frac{P_j f_j(b,s)}{P_1 f_1(b,s) + P_2 f_2(b,s) + P_3 f_3(b,s)}\ln\, f_j(b,s)\right\} \\ = &\sum_{i=1}^{I}\sum_{j=1}^{3} a_{ij}^{(r)} \ln P_j^{(r)} + C, \end{aligned}$$

where $a_{ij}^{(r)} = \frac{P_j^{(r)} f_j^{(r)}(b,s)}{P_1^{(r)} f_1^{(r)}(b,s) + P_2^{(r)} f_2^{(r)}(b,s) + P_3^{(r)} f_3^{(r)}(b,s)}$, and C is a constant.

The M-step is to maximize Q. We find the result of the first-order condition for P_j is independent of j and the result is shown as follows:

$$\begin{aligned} &\frac{\partial Q}{\partial P_j} = \sum_{i=1}^{I}\left(\frac{a_{ij}^{(r)}}{P_j} - \frac{a_{i1}^{(r)}}{P_1}\right) = 0, \quad \text{then} \\ &\sum_{i=1}^{I}\frac{a_{ij}^{(r)}}{P_j} = \sum_{i=1}^{I}\frac{a_{i1}^{(r)}}{P_1} = d, \quad \text{where } d \text{ is a constant} \end{aligned}$$

$$\sum_{i=1}^{I} a_{ij}^{(r)} = dP_j, \forall j = 1,2,3.$$

Since $\sum_{j=1}^{3} P_j = 1$, and $\sum_{j=1}^{3} a_{ij}^{(r)} = 1$, we obtain $\sum_{i=1}^{I} a_{ij}^{(r)} = I$ and $dP_j = d$ and solve

$$P_j = \frac{1}{I}\sum_{\mathrm{i}=1}^{\mathrm{I}} a_{ij}^{(r)}, \forall j = 1,2,3. \tag{96.8}$$

Applying EM algorithm in the estimation of *PIN*, we can simplify the estimation procedure as the following steps:

1. A naïve guess (or estimation) for initials of parameters, e.g., $P_j^{(0)}$, $\varepsilon_b^{(0)}$, $\varepsilon_s^{(0)}$, $\mu^{(0)}$
2. Apply r-step $P_j^{(r)}$, $\varepsilon_b^{(r)}$, $\varepsilon_s^{(r)}$, $\mu^{(r)}$ into

$$E\left(Z_{ij}^{(r)}|B,S\right) = \frac{P_j^{(r)} f_j^{(r)}(B_i,S_i)}{P_1^{(r)} f_1^{(r)}(B_i,S_i) + P_2^{(r)} f_2^{(r)}(B_i,S_i) + P_3^{(r)} f_3^{(r)}(B_i,S_i)} = a_{ij}^{(r)}.$$

3. After step (2), the function becomes complete-data MLE. We maximize likelihood function Q to estimate five parameters, $P_1^{(r)}$, $P_2^{(r)}$, $\varepsilon_b^{(r)}$, $\varepsilon_s^{(r)}$, and $\mu^{(r)}$, and derive $P_j^{(r+1)}$ and $Z_{ij}^{(r+1)}$. We obtain

$$P_j^{(r+1)} = \frac{1}{I}\sum_{i=1}^{\mathrm{I}} a_{ij}^{(r)}$$

and

$$Z_{ij}^{(r+1)} = \frac{P_j^{(r+1)} f_J^{(r)}(B_i,S_i)}{P_1^{(r+1)} f_1^{(r)}(B_i,S_i) + P_2^{(r+1)} f_2^{(r)}(B_i,S_i) + P_3^{(r+1)} f_3^{(r)}(B_i,S_i)}. \tag{96.9}$$

Numerical Example

Below are the results by using real data. An initial naïve guess for $(P_1^{(0)}, P_2^{(0)}, P_3^{(0)}, \varepsilon_b^{(0)}, \varepsilon_s^{(0)}, \mu^{(0)}) = (0.3, 0.4, 0.3, 1, 1, 1)$ when buys, $B_i = 193$ and sells, $S_i = 179$ during a 5-min interval. By Eqs. 96.3, 96.4, and 96.5, we calculate the following:

$$f_1^{(0)}(B_i,S_i) = e^{-\varepsilon_b}\frac{\varepsilon_b^B}{B_i!}e^{-\varepsilon_s}\frac{\varepsilon_s^S}{S_i!} = \frac{e^{-1}\times 1\times e^{-1}\times 1}{193!\times 179!},$$

$$f_2^{(0)}(B_i,S_i) = e^{-\varepsilon_b}\frac{\varepsilon_b^B}{B_i!}e^{-(\mu+\varepsilon_s)}\frac{(\mu+\varepsilon_s)^S}{S_i!} = \frac{e^{-1}\times 1\times e^{-2}\times 2}{193!\times 179!}$$

$$f_3^{(0)}(B_i,S_i) = e^{-(\mu+\varepsilon_b)}\frac{(\mu+\varepsilon_b)^B}{B_i!}e^{-\varepsilon_s}\frac{\varepsilon_s^S}{S_i!} = \frac{e^{-2}\times 2\times e^{-1}\times 1}{193!\times 179!}$$

and

$$a_{i1}^{(0)} = Z_{i1}^{(0)} = \frac{0.3 \times f_1^{(0)}(B_i, S_i)}{0.3 \times f_1^{(0)} + 0.4 \times f_2^{(0)} + 0.3 \times f_3^{(0)}}$$
$$a_{i2}^{(0)} = Z_{i2}^{(0)} = \frac{0.4 \times f_2^{(0)}(B_i, S_i)}{0.3 \times f_1^{(0)} + 0.4 \times f_2^{(0)} + 0.3 \times f_3^{(0)}}$$
$$a_{i3}^{(0)} = Z_{i3}^{(0)} = \frac{0.3 \times f_3^{(0)}(B_i, S_i)}{0.3 \times f_1^{(0)} + 0.4 \times f_2^{(0)} + 0.3 \times f_3^{(0)}}.$$

Hence,

$$P_1^{(1)} = \frac{1}{I}\sum_{\mathrm{i=1}}^{\mathrm{I}} a_{i1}^{(0)}, P_2^{(1)} = \frac{1}{I}\sum_{\mathrm{i=1}}^{\mathrm{I}} a_{i2}^{(0)}, P_3^{(1)} = \frac{1}{I}\sum_{\mathrm{i=1}}^{\mathrm{I}} a_{i3}^{(0)}.$$

By Eq. 96.9, we obtain

$$Z_{i1}^{(1)} = \frac{P_1^{(1)} f_1^{(0)}(B_i, S_i)}{P_1^{(1)} f_1^{(0)}(B_i, S_i) + P_2^{(1)} f_2^{(0)}(B_i, S_i) + P_3^{(1)} f_3^{(0)}(B_i, S_i)}$$
$$Z_{i2}^{(1)} = \frac{P_2^{(1)} f_2^{(0)}(B_i, S_i)}{P_1^{(1)} f_1^{(0)}(B_i, S_i) + P_2^{(1)} f_2^{(0)}(B_i, S_i) + P_3^{(1)} f_3^{(0)}(B_i, S_i)}$$
$$Z_{i3}^{(1)} = \frac{P_3^{(1)} f_3^{(0)}(B_i, S_i)}{P_1^{(1)} f_1^{(0)}(B_i, S_i) + P_2^{(1)} f_2^{(0)}(B_i, S_i) + P_3^{(1)} f_3^{(0)}(B_i, S_i)}$$

The complete-data MLE becomes

$$G = \ln V = \sum_{i=1}^{I}\sum_{j=1}^{3}\left\{Z_{ij}\ln\ P_j + Z_{ij}\ln\ f_j(B_i, S_i)\right\}.$$

When the observed data consists of complete pairs, we derive a set of equations for maximum likelihood estimation. The first-order condition is

$$\frac{\partial G}{\partial P_j} = \sum_{i=1}^{I}\frac{Z_{ij}}{P_j} = 0, \forall \mathrm{j} = 1, 2, 3.$$

If $j = 3$, then $\frac{\partial G}{\partial P_3} = 0$, and $P_3 = \frac{1}{I}\sum_{i=1}^{I} Z_{i3}$

If $j = 2$, then $\frac{\partial G}{\partial P_2} = 0$, and $P_2 = \frac{1}{I}\sum_{i=1}^{I} Z_{i2}$

If $j = 1$, then $P_1 = 1 - \frac{1}{I}\sum_{i=1}^{I} Z_{i3} - \frac{1}{I}\sum_{i=1}^{I} Z_{i3}$.

To solve ε_b, the first-order condition is

$$\frac{\partial Q}{\partial \varepsilon_b} = \sum_{i=1}^{I}\left(Z_{i1}\frac{\partial \ln f_1}{\partial \varepsilon_b} + Z_{i2}\frac{\partial \ln f_2}{\partial \varepsilon_b} + Z_{i3}\frac{\partial \ln f_3}{\partial \varepsilon_b}\right) = 0 \tag{96.10}$$

where

$$\begin{aligned}
\ln f_1 &= -\varepsilon_b + B_i \ln \varepsilon_b - \ln(B_i!) - \varepsilon_s + S_i \ln \varepsilon_S - \ln(S_i!),\\
\ln f_2 &= -\varepsilon_b + B_i \ln \varepsilon_b - \ln(B_i!) - (\mu + \varepsilon_s) + S_i \ln(\mu + \varepsilon_S) - \ln(S_i!),\\
\ln f_3 &= -(\mu + \varepsilon_b) + B_i \ln(\mu + \varepsilon_b) - \ln(B_i!) - \varepsilon_s + S_i \ln \varepsilon_S - \ln(S_i!).
\end{aligned}$$

We then rewrite Eq. 96.10 as

$$\sum_{\mathrm{i}=1}^{\mathrm{I}}\left(Z_{i1}\left(-1 + \frac{B_i}{\varepsilon_b}\right) + Z_{i2}\left(-1 + \frac{B_i}{\varepsilon_b}\right) + Z_{i3}\left(-1 + \frac{B_i}{\mu + \varepsilon_b}\right)\right) = 0$$

and then reorganize it to

$$\sum_{i=1}^{I}(Z_{i1} + Z_{i2} + Z_{i3}) = \sum_{i=1}^{I}\left(\frac{Z_{i1}B_i + Z_{i2}B_i}{\varepsilon_b} + \frac{Z_{i3}B_i}{\mu + \varepsilon_b}\right).$$

Since $\sum_{i=1}^{I}(Z_{i1} + Z_{i2} + Z_{i3}) = I$ and $\sum_{i=1}^{I}\left(\frac{Z_{i1}B_i + Z_{i2}B_i}{\varepsilon_b} + \frac{Z_{i3}B_i}{\mu + \varepsilon_b}\right) = a_b x_b + b_b y_b$, we simplify the equation to

$$\mathrm{I} = a_b x_b + b_b y_b, \ \ldots \text{(i)}$$

$$\text{where} \quad x_b = \frac{1}{\varepsilon_b}, y_b = \frac{1}{\mu + \varepsilon_b}.$$

In a similar way, we can attain

$$\begin{aligned}
&\frac{\partial Q}{\partial \varepsilon_s} = \sum_{i=1}^{I}\left(Z_{i1}\frac{\partial \ln f_1}{\partial \varepsilon_s} + Z_{i2}\frac{\partial \ln f_2}{\partial \varepsilon_s} + Z_{i3}\frac{\partial \ln f_3}{\partial \varepsilon_s}\right) = 0.\\
&I = a_s x_s + b_s y_s, \ \ldots \text{(ii)}\\
&\text{where} \quad x_s = \frac{1}{\varepsilon_s}, y_s = \frac{1}{\mu + \varepsilon_s}.
\end{aligned}$$

$$\begin{aligned}
&\frac{\partial Q}{\partial \mu} = \sum_{i=1}^{I}\left(Z_{i1}\frac{\partial \ln f_1}{\partial \mu} + Z_{i2}\frac{\partial \ln f_2}{\partial \mu} + Z_{i3}\frac{\partial \ln f_3}{\partial \mu}\right) = 0,\\
&I - \sum_{i=1}^{I} Z_{i1} = b_s y_s + b_b y_b.
\end{aligned}$$

By substitute $b_b y_b$ and $b_s y_s$ from equations (i) and (ii),

$$\sum_{i=1}^{I} Z_{i1} + I = a_s x_s + a_b x_b, \ \ldots \text{(iii)}.$$

We can rewrite x_b, y_b, and y_s to

$$x_b = \frac{\sum_{i=1}^{I} Z_{i1} + I - a_s x_s}{a_b}, \ y_b = \frac{a_s x_s - \sum_{i=1}^{I} Z_{i1}}{b_b}, \text{ and } y_s = \frac{I - a_s x_s}{b_s}.$$

To solve four variables, we need four equations. We hence artificially develop equation (iv), substitute x_b, y_b, and y_s to

$$\frac{1}{y_b} - \frac{1}{x_b} = \frac{1}{y_s} - \frac{1}{x_s} \ldots \text{(iv)}$$

and organize equation (iv) into

$$\frac{b_b}{a_s x_s - \sum_{i=1}^{I} \mathrm{Z}_{\mathrm{i}1}} - \frac{a_b}{\sum_{i=1}^{I} \mathrm{Z}_{\mathrm{i}1} + \mathrm{I} - a_s x_s} = \frac{b_s}{\mathrm{I} - a_s x_s} - \frac{1}{x_s}.$$

Finally, we mathematically process fractions to a common denominator and achieve a function:

$$f(x) = a_3 x_s^3 - a_2 x_s^2 + a_1 x_s - a_0$$

where

$$\begin{aligned}
a_3 &= a_b a_s^2 + a_s^3 + a_s^2 b_s + a_s^2 b_s, \\
a_2 &= a_b a_s I + 2a_s^2 I + 2a_s b_b I + a_s b_s I + a_b a_s \sum_{i=1}^{I} Z_{i1} + 2a_s^2 \sum_{i=1}^{I} Z_{i1} + a_s b_b \sum_{i=1}^{I} Z_{i1} \\
&\quad + 2a_s b_s \sum_{i=1}^{I} Z_{i1}, \\
a_1 &= a_s I^2 + a_b I \sum_{i=1}^{I} Z_{i1} + 3a_s I \sum_{i=1}^{I} Z_{i1} + b_s I \sum_{i=1}^{I} Z_{i1} + a_s \Big(\sum_{i=1}^{I} Z_{i1}\Big)^2 + b_s \left(\sum_{i=1}^{I} Z_{i1}\right)^2 \\
&\quad + b_b I^2 + b_b I \sum_{i=1}^{I} Z_{i1}, \\
a_0 &= I^2 \sum_{i=1}^{I} Z_{i1} + I\Big(\sum_{i=1}^{I} Z_{i1}\Big)^2.
\end{aligned} \tag{96.11}$$

Numerical Analysis Method

To solve x_s in Eq. 96.11, we employ Newton's method. In numerical analysis, Newton's method (also known as the Newton–Raphson method) is a method for finding successively superior approximations to the roots of a real-valued function.[4] The Newton–Raphson method for our case is implemented as follows.

Given our function $f(x)$ defined over the real x_s and its derivative $f'(x_s)$, we begin with a first guess $x_s^{(0)}$ for a root of the function $f(x_s)$. Provided that the function is reasonably well behaved, a better approximation $x_s^{(1)}$ is

$$x_s^{(1)} = x_s^{(0)} - \frac{f\left(x_s^{(0)}\right)}{f'\left(x_s^{(0)}\right)}.$$

Geometrically, $(x_s^{(1)}, 0)$ is the intersection with the x-axis of a line tangent to f at $(x_s^{(0)}, f(x_s^{(0)}))$.

The process is repeated as

$$x_s^{(n+1)} = x_s^{(n)} - \frac{f\left(x_s^{(n)}\right)}{f'\left(x_s^{(n)}\right)}$$

until a sufficiently accurate value is reached.

The Precision of Parameter Estimation

Information on μ, ε_b, ε_s accumulates at a rate approximately equal to the square root of the number of trade outcomes. On the other hand, information on α and δ accumulates at a rate approximately equal to the square root of the number of trading days. The difference in information accumulation rates dictates that the precision of μ, ε_b, ε_s will exceed that of α and δ estimates.

Different μ (μ_b, μ_s) Case

As shown in Eq. 96.2, the arrival rate of informed buyers differs from that of informed sellers. For our model, Eqs. 96.4 and 96.5 are rewritten as

$$f_2 = e^{-\varepsilon_b} \frac{\varepsilon_b^B}{B_i!} e^{-(\mu_s + \varepsilon_s)} \frac{(\mu_s + \varepsilon_s)^S}{S_i!} \tag{96.4*}$$

[4]Newton's method can also be extended to complex functions and to systems of equations.

$$f_3 = e^{-(\mu_b + \varepsilon_b)} \frac{(\mu_b + \varepsilon_b)^B}{B_i!} e^{-\varepsilon_s} \frac{\varepsilon_s^S}{S_i!} \tag{96.5*}$$

The parameters are estimated jointly. The empirical work of our case is discussed in Chang and Lin (2014). We estimate the rate of informed and uninformed trading on a particular day for index futures and spot markets, as well as of the information event structure.

The logarithm of f_1 remains the same, but those of f_2 and f_3 change to

$$\begin{aligned}
\ln f_2 &= -\varepsilon_b + B_i \ln \varepsilon_b - \ln(B_i!) - (\mu_s + \varepsilon_s) + S_i \ln(\mu_s + \varepsilon_s) - \ln(S_i!),\\
\ln f_3 &= -(\mu_b + \varepsilon_b) + B_i \ln(\mu_b + \varepsilon_b) - \ln(B_i!) - \varepsilon_s + S_i \ln \varepsilon_S - \ln(S_i!),
\end{aligned}$$

We use the same method for same μ case to estimate ε_b, ε_s, μ_b, and μ_s. We find it easier to estimate $\theta^* = (\alpha, \delta, \mu_b, \mu_s, \varepsilon_b, \varepsilon_s)$ than $\theta = (\alpha, \delta, \mu, \varepsilon_b, \varepsilon_s)$ because there is one more condition for the case of different μ. The details of the estimation are shown as

$$\begin{aligned}
&\frac{\partial Q}{\partial \varepsilon_b} = \sum_{i=1}^{I} \left(Z_{i1} \frac{\partial \ln f_1}{\partial \varepsilon_b} + Z_{i2} \frac{\partial \ln f_2}{\partial \varepsilon_b} + Z_{i3} \frac{\partial \ln f_3}{\partial \varepsilon_b} \right) = 0\\
&I = a_b x_b + b_b y_b, \ \ldots \text{(v)}\\
&\text{where} \quad x_b = \frac{1}{\varepsilon_b}, y_b = \frac{1}{\mu_b + \varepsilon_b}
\end{aligned}$$

$$\begin{aligned}
&\frac{\partial Q}{\partial \varepsilon_s} = \sum_{i=1}^{I} \left(Z_{i1} \frac{\partial \ln f_1}{\partial \varepsilon_s} + Z_{i2} \frac{\partial \ln f_2}{\partial \varepsilon_s} + Z_{i3} \frac{\partial \ln f_3}{\partial \varepsilon_s} \right) = 0\\
&I = a_s x_s + b_s y_s, \ \ldots \text{(vi)}\\
&\text{where} \quad x_s = \frac{1}{\varepsilon_s}, y_s = \frac{1}{\mu_s + \varepsilon_s}
\end{aligned}$$

$$\begin{aligned}
&\frac{\partial Q}{\partial \mu_b} = \sum_{i=1}^{I} \left(Z_{i1} \frac{\partial \ln f_1}{\partial \mu_b} + Z_{i2} \frac{\partial \ln f_2}{\partial \mu_b} + Z_{i3} \frac{\partial \ln f_3}{\partial \mu_b} \right) = 0\\
&\sum_{i=1}^{I} Z_{i3} = b_b y_b, \ \ldots \text{(vii)}\\
&\text{where} \quad y_b = \frac{1}{\mu_b + \varepsilon_b}
\end{aligned}$$

$$\begin{aligned}
&\frac{\partial Q}{\partial \mu_s} = \sum_{i=1}^{I} \left(Z_{i1} \frac{\partial \ln f_1}{\partial \mu_s} + Z_{i2} \frac{\partial \ln f_2}{\partial \mu_s} + Z_{i3} \frac{\partial \ln f_3}{\partial \mu_s} \right) = 0\\
&\sum_{i=1}^{I} Z_{i3} = b_s y_s; \ldots \text{(viii)}\\
&\text{where} \quad y_s = \frac{1}{\mu_s + \varepsilon_s}
\end{aligned}$$

We can solve x_b, x_s, y_b, and y_s from the following equations:

$$\begin{aligned} &I = a_b x_b + b_b y_b, \ldots (\mathrm{v}) \\ &I = a_s x_s + b_s y_s, \ldots (\mathrm{vi}) \\ &\sum_{i=1}^{I} Z_{i3} = b_b y_b, \ldots (\mathrm{vii}) \\ &\sum_{i=1}^{I} Z_{i2} = b_s y_s. \ldots (\mathrm{viii}) \end{aligned}$$

From (vii), $y_b = \frac{\sum_{i=1}^{I} Z_{i3}}{b_b}, x_b = \frac{I - \sum_{i=1}^{I} Z_{i3}}{a_b}$, then $\varepsilon_b = \frac{a_b}{I - \sum_{i=1}^{I} Z_{i3}}, \mu_b = \frac{b_b}{\sum_{i=1}^{I} Z_{i3}} - \varepsilon_b.$

From (viii), $y_s = \frac{\sum_{i=1}^{I} Z_{i2}}{b_s}, x_s = \frac{I - \sum_{i=1}^{I} Z_{i2}}{a_s}$, then $\varepsilon_s = \frac{a_s}{I - \sum_{i=1}^{I} Z_{i2}}, \mu_s = \frac{b_s}{\sum_{i=1}^{I} Z_{i2}} - \varepsilon_s.$

Based on these parameters, we are ready to calculate Easley et al.'s probability of informed trading, $PIN = \frac{\alpha\mu}{\alpha\mu + \varepsilon_{st} + \varepsilon_b}$ and our revised $PIN^* = \frac{\alpha(\delta|\mu_s| + (1-\delta)|\mu_b|)}{\alpha(\delta|\mu_s| + (1-\delta)|\mu_b|) + \varepsilon_s + \varepsilon_b}$.

Acknowledgments Special thanks to professors Lii-Yuh Leu, Ren-Her Wang, and Charles Chang for their helpful comments and consultation. All errors are our own.

References

Bagehot, W. (1971). The only game in town. *Financial Analysis Journal, 27*, 12–14.

Boehmer, E., Gramming, J., & Theissen, E. (2007). Estimating the probability of informed trading-does trade misclassification matter? *Journal of Financial Markets, 10*, 26–47.

Bollen, N. P., Smith, T., & Whaley, R. E. (2004). Modeling the bid/ask spread: Measuring the inventory-holding premium. *Journal of Financial Economics, 72*, 97–141.

Bookstein, A., & Swanson, D. (1974). Probabilistic models for automatic indexing. *Journal of the American Society for Information Science, 25*(5), 312–318.

Brennan, M. J., & Subrahmanyam, A. (1996). Market microstructure and asset pricing: On the compensation for illiquidity in stock returns. *Journal of Financial Economics, 41*, 441–464.

Chang, C., & Lin, E. (2014). On the determinants of basis spread for Taiwan index futures and the role of speculators. *Review of Pacific Basin Financial Markets and Policies, 17*(1), 1450002-1–1450002-30.

Church, K. W., & Gale, W. A. (1995). Poisson mixtures. *Natural Language Engineering, 1*(2), 163–190.

Copeland, T. E., & Galai, D. (1983). Information effects on the bid-ask spread. *Journal of Finance, 38*, 1457–1469.

Duarte, J., & Young, L. (2009). Why is PIN priced? *Journal of Financial Economics, 91*(2), 119–138.

Easley, D., Kiefer, N., O'Hara, M., & Paperman, J. B. (1996). Liquidity, information and infrequently traded stocks. *Journal of Finance, 51*, 1405–1436.

Easley, D., Hvidkjaer, S., & O'Hara, M. (2002). Is information risk a determinant of asset returns? *Journal of Finance, 57*, 2185–2221.

Easley, D., Engle, R. F., O'Hara, M., & Wu, L. (2008). Time-varying arrival rates of informed and uninformed traders. *Journal of Financial Econometrics, 6*, 171–207.

Easley, D., de Prado López, M., & O'Hara, M. (2012). Flow toxicity and liquidity in a high frequency world. *Review of Financial Studies, 25*(5), 1457–1493.

Ellis, K., Michaely, R., & O'Hara, M. (2000). The accuracy of trade classification rules: Evidence from Nasdaq. *Journal of Financial and Quantitative Analysis, 35*, 529–551.

Foster, F. D., & Viswanathan, S. (1993). Variations in trading volume, return volatility, and trading costs: Evidence on recent price formation models. *Journal of Finance, 48*, 187–211.

Glosten, L. R., & Harris, L. E. (1988). Estimating the components of the bid/ask spread. *Journal of Financial Economics, 21*, 123–142.

Hasbrouck, J. (1991). Measuring the information content of stock trades. *Journal of Finance, 46*, 179–207.

Johnson, N., & Kotz, S. (1969). *Discrete distributions*. Boston: Houghton Mifflin.

Lee, C. F., & Chen, H. Y. (2012). Alternative errors-in-variable models and their applications in finance research. Rutgers University working paper.

Lee, C., & Ready, M. (1991). Inferring trade direction from intraday data. *Journal of Finance, 46*, 733–747.

Lei, Q., & Wu, G. (2005). Time-varying informed and uninformed trading activities. *Journal of Financial Markets, 8*, 153–181.

Lin, E., Lee, C. F., & Wang, K. (2013). Futures mispricing, order imbalance, and short-selling constraints. *International Review of Economics and Finance*, *25*, 408–423.

MacKinlay, A. C., & Ramaswamy, K. (1988). Index futures arbitrage and the behavior of stock index futures prices. *Review of Financial Studies*, *1*, 137–158.

Madhavan, A., & Smidt, S. (1991). A Bayesian Model of intraday specialist pricing. *Journal of Financial Economics, 30*, 99–134.

McLachlan, G. J., & Krishnan, T. (2008). *The EM algorithm and extensions* (Wiley series in Probability and Statistics; 2nd ed.). Hoboken, NJ: Wiley.

Odders-White, E. R. (2000). On the occurrence and consequences of inaccurate trade classification. *Journal of Financial Markets, 3*, 259–286.

Popescu, M., & Kumar, R. (2008). An ex-ante measure of the probability of informed trading. http://ssrn.com/abstract=891717

CEO Stock Options and Analysts' Forecast Accuracy and Bias

97

Kiridaran Kanagaretnam, Gerald J. Lobo, and Robert Mathieu

Contents

Abstract

In this study, we investigate the relations between CEO stock options and analysts' earnings forecast accuracy and bias. We argue that a higher level of stock options may induce managers to undertake riskier projects, to change

K. Kanagaretnam (✉)
Schulich School of Business, York University, Toronto, ON, Canada
e-mail: KKanagaretnam@schulich.yorku.ca

G.J. Lobo
C.T. Bauer College of Business, University of Houston, Houston, TX, USA
e-mail: gjlobo@uh.edu

R. Mathieu
School of Business and Economics, Wilfrid Laurier University, Waterloo, ON, Canada
e-mail: rmathieu@wlu.ca

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_97,

and/or reallocate their effort, and to possibly engage in gaming (such as opportunistic earnings and disclosure management) and hypothesize that these managerial behaviors will result in an increase in the complexity of forecasting and, hence, in less accurate analysts' forecasts. We also posit that analysts' optimistic forecast bias will increase as the level of stock options pay increases. We reason that as forecast complexity increases with stock options pay, analysts, needing greater access to management's information to produce accurate forecasts, have incentives to increase the optimistic bias in their forecasts. Alternatively, a higher level of stock options pay may lead to improved disclosure because it better aligns managers' and shareholders' interests. The improved disclosure, in turn, may result in more accurate and less biased analysts' forecasts.

Using ordinary least squares estimation, we test these hypotheses relating the level of CEO stock options pay to analysts' forecast accuracy and bias on a sample of firms from the Standard & Poor's ExecuComp database over the period 1993–2003. Our OLS models relate forecast accuracy and forecast bias (the dependent variables) to CEO stock options (the independent variable) and controls for earnings characteristics, firm characteristics, and forecast characteristics. We measure forecast accuracy as negative one times the absolute value of the difference between forecasted and actual earnings scaled by beginning of period stock price and forecast bias as forecasted minus actual earnings scaled by beginning of period stock price. We control for differences in earnings characteristics by including earnings volatility, whether the firm has a loss, and earnings surprise; for differences in firm characteristics by including firm size, growth (measured as book-to-market ratio, percentage change in total assets, and percentage change in annual sales), and corporate governance quality (measured as percentage of shares outstanding owned by the CEO, whether the CEO is also chairman of the board of directors, number of annual board meetings, and whether directors are awarded stock options); and for differences in forecast characteristics by including analyst following and analyst forecast dispersion. In addition, the models include controls for industry and year. We use four measures of options: new options, existing exercisable options, existing unexercisable options, and total options (sum of the previous three), all scaled by total number of shares outstanding, and estimate two models for each dependent variable, one including total options and the other including new options, existing exercisable options, and existing unexercisable options. We also use both contemporaneous as well as lagged values of options in our main tests.

Our results indicate that analysts' earnings forecast accuracy decreases and forecast optimism increases as the level of stock options (particularly new options and existing exercisable options) in CEO pay increases. These findings suggest that the incentive alignment effects of stock options are more than offset by the investment, effort allocation, and gaming incentives induced by stock options grants to CEOs. Given that analysts' forecasts are an important source of information to capital markets, our finding of a decline in the quality of

the information provided by analysts has implications for the level and variability of stock prices. It also has implications for information asymmetry and cost of capital, as well as for valuation models that rely on analysts' earnings forecasts.

Keywords
CEO stock options • Analysts' forecast accuracy • Analysts' forecast bias • CEO compensation • Agency costs • Investment risk taking • Effort allocation • Opportunistic earnings management • Opportunistic disclosure management • Forecasting complexity

97.1 Introduction

Firms grant equity incentives such as stock pay, restricted stock, and stock options to create incentives for executives to make decisions that benefit shareholders.[1] By linking executive compensation to shareholder wealth, stock options purportedly help reduce agency costs that arise from the separation of ownership and control in corporations. Over the last decade, stock options have become an increasingly larger component of executive compensation. Most large firms compensate their top executives through stock options which, on a Black-Scholes valuation basis, now represent the largest single component of managerial pay (Murphy 1999; Hall and Murphy 2003).

Several studies have examined the economic implications of stock options and other forms of equity compensation. Much of that research has focused on the relation between stock options compensation and firm performance (Core et al. 1999; Guay 1999; Hanlon et al. 2003; Lam and Chng 2006; Bauman and Shaw 2006), on the link between stock options and investment decisions (Smith and Watts 1992; Bizjak et al. 1993), and on the relation between stock options and dividend policy and dividend yield (Lambert et al. 1989; Atan et al. 2010). Other research has examined whether stock options and other equity-based compensation induce managers to increase short-term stock price through earnings management (Bartov and Mohanram 2004; Cheng and Warfield 2005; Cao and Laksmana 2010).[2] Efendi et al. (2007) provide evidence that the amount of stock options in-the-money is the most influential factor affecting the likelihood of a misstatement. Bergstresser and Philippon (2006) document that managers with a large proportion of stock and options holdings are more likely to use discretionary accruals to manipulate earnings. Chen (2002) finds a negative relationship between incentive compensation and stock ownership held by outside directors.

[1] Other reasons for granting stock options are to attract and retain executives, to conserve cash, to reduce reported accounting expense, and to defer taxes.

[2] Although there is evidence relating earnings management to stock options compensation, little is known about whether this earnings management actually results in higher payouts or about its effect on other goals of the firm.

The purpose of this study is to examine the implications of executive stock options compensation for the accuracy of and bias in analysts' earnings forecasts. Prior research has studied the links between stock options compensation and earnings/disclosure management and between disclosure quality/forecast complexity and analysts' forecast properties. However, it has not directly examined the link between stock options compensation and analysts' forecast accuracy and bias. Our study provides a direct test of this link. One contribution of our study is that it provides a triangulation of the relationships observed in prior research. If financial intermediaries such as analysts can see through the incentives for manipulation by managers with large stock options grants, then the quality of the information will not be affected. However, if analysts cannot see through this manipulation, it will affect the accuracy and bias of their forecasts, and, hence, it becomes an important issue. In this regard, it validates those findings by empirically documenting the relation between stock options compensation and the quality of analysts' earnings forecasts (a common proxy for a firm's information environment).

Higher levels of stock options may induce managers to undertake riskier projects, to change and/or reallocate their effort, and to possibly engage in gaming (such as opportunistic earnings and disclosure management). Consequently, the forecasting task will be more complex as the proportion of stock options compensation increases, leading to less accurate forecasts. While the above reasoning suggests that forecast accuracy may decrease as the level of stock options in CEO pay increases, higher stock options pay may also result in increased accuracy if it better aligns managers' and shareholders' interests. Hanlon et al. (2003) report an increase in future operating earnings associated with past stock options grants, providing empirical support for improved incentive alignment. Better incentive alignment likely improves managers' disclosures which, in turn, may lead to more accurate forecasts.

Analysts' compensation and reputation are, to a large extent, dependent on the accuracy of their forecasts. They can improve the accuracy of their forecasts if they have access to management's private information.[3] Such access becomes even more valuable as the difficulty of the forecasting task increases. Analysts can increase access to management's private information by developing better relations with management. One way of accomplishing this is by making optimistic forecasts. In so doing, analysts trade off forecast bias for improved accuracy. If forecasting difficulty increases with the level of stock options in CEO pay, then so will the optimistic bias in analysts' earnings forecasts. Alternatively, the improved management disclosures resulting from the increased incentive alignment effects of stock options may lead to less biased forecasts.

We present empirical evidence on the relation between the level of stock options in CEO pay and the accuracy and bias of analysts' earnings forecasts for firms in the Standard & Poor's ExecuComp database over the period 1993–2003.

[3]Ke and Yu (2006) and Chen and Matsumoto (2006) are examples of recent research on analysts' incentives for access to management, i.e., the management relations hypothesis.

We estimate the relations between the level of CEO stock options pay and analysts' forecast accuracy and bias using ordinary least squares estimation. We use four measures of options: new options, existing exercisable options, existing unexercisable options, and total options (sum of the previous three) all scaled by total number of shares outstanding. We estimate two models, one including total options and the other including new options, existing exercisable options, and existing unexercisable options. We also use both contemporaneous as well as lagged values of options in our main tests. Our results indicate that analysts' earnings forecast accuracy decreases and forecast optimism increases as the level of stock options (particularly for new options and existing exercisable options) in CEO pay increases. These findings suggest that the incentive alignment effects of stock options are more than offset by the investment, effort allocation, and gaming incentives induced by stock options grants to CEOs.

While stock options may help companies attract, retain, and motivate executives, they also have associated costs. Hall and Murphy (2003) indicate that stock options may be an inefficient form of compensation because the value to recipients who are undiversified and risk averse, and who neither can sell nor hedge against their risk, is less than the cost to the firm. Consequently, options are a costly form of compensation relative to cash or stock compensation. We document a decrease in the quality of analysts' earnings forecasts as the level of stock options in CEO pay increases. Given that analysts' forecasts are an important source of information to capital markets, a decline in the quality of the information provided by analysts has implications for the level and variability of stock prices. It also has implications for information asymmetry and cost of capital.

Our findings also have implications for investors, academics, and other users of financial analysts' forecasts. Because analysts' forecasts serve as expectations of a firm's future prospects, they play an important role in firm valuation. Analysts' forecasts are also commonly used as measures of the market's earnings expectations in studies that investigate the relation between earnings and stock returns and changes in analysts' forecasts are related to stock returns (Givoly and Lakonishok 1979; Imhoff and Lobo 1984; Stickel 1991; Barber et al. 2001; Jegadeesh et al. 2004). Furthermore, earnings forecasts serve as inputs to other research outputs such as stock recommendations (Loh and Mian 2006), target price forecasts (Bandyopadhyay et al. 1995), valuation models (Frankel and Lee 1998), and growth and return on equity investment models (Easton et al. 2002). By identifying the relations between the level of stock options in CEO pay and forecast accuracy and bias, our study provides investors, researchers, and other users with an ex ante indicator of the accuracy of analysts' forecasts.

The rest of this paper is organized as follows. Section 97.2 develops the research hypotheses, and Sect. 97.3 describes the sample selection and research design. Section 97.4 presents the results of the empirical analysis. Section 97.5 reports the results of sensitivity tests, and Sect. 97.6 contains the conclusions of the study.

97.2 Hypotheses

97.2.1 CEO Stock Options and Forecast Accuracy

An increase in stock options pay to CEOs may increase the complexity of the forecasting task for several reasons. *First*, the convex payoffs from options may induce otherwise risk-averse managers to undertake riskier projects going forward (Murphy 2003; Ross 2004). Theoretical models (Smith and Stulz 1985; Smith and Watts 1992; Bizjak et al. 1993) demonstrate that, because managers' investment decisions are particularly difficult to monitor, firms with substantial investment opportunities tend to encourage higher investment activities by aligning the interests of managers and shareholders through stock options grants. Consistent with this notion, Rajgopal and Shevlin (2002) find empirical evidence linking executive stock options and exploration risks in the oil and gas industry. Relatedly, Guay (1999) documents that equity risk is positively related to the convexity in executives' compensation schemes. These results are consistent with stock options providing managers with incentives to mitigate risk-related incentive problems.

Second, the incentive effect of stock options is to motivate managers to exert higher effort. Such contributions will translate into higher performance in both the current and future periods. Managers exert effort in multiple dimensions (Banker and Datar 1989; Holmstrom and Milgrom 1991). With new incentives, managers may also reallocate their effort contributions in addition to exerting higher effort. Stock options grants may induce managers to reallocate their effort mix from short-term effort that focuses more on improving current performance to long-term (or strategic) effort that places more emphasis on improving future performance (Bushman and Indjejikian 1993; Feltham and Xie 1994). To the extent that managers were short-term oriented prior to receiving stock options grants, the reallocation of effort mix could have a negative impact on current performance. The extent of this potential effect on current performance cannot be easily gauged because the reallocation of effort mix is not directly observable or predictable. This, in turn, may result in an increase in forecasting complexity.[4]

Third, managers with higher levels of stock options may engage in higher levels of gaming (Hall 2003). This gaming behavior can take many forms including opportunistic earnings and disclosure management. Lambert (2001) points out that earnings management strategies will be influenced by the shape of the compensation contract (i.e., whether it is linear, concave, or convex in a given region). An increase in stock options pay relative to cash compensation will increase the convexity of the compensation contract and result in increased incentives for earnings management that is anti-smoothing in nature. Consequently, managers

[4]Furthermore, Feltham and Xie (1994) show that, if there are multiple tasks and multiple public signals that are influenced by the manager's action, it is unlikely that the market price provides an efficient single performance measure. Therefore, overly relying on stock-based compensation may lead to incongruent behavior by CEOs, further increasing the difficulty of the forecasting task.

may increase their current earnings when earnings are high and reduce their current earnings when earnings are low. Such behavior will directly contribute to an increase in earnings volatility. Bartov and Mohanram (2004) provide evidence consistent with the above argument. They document that managers inflate earnings through accruals management in the period leading up to abnormally large stock options exercises. Other research has examined whether stock options and other equity-based compensation induce managers to engage in high levels of earnings management. For example, Cheng and Warfield (2005) document that managers with high equity incentives are more likely to manage earnings in order to boost stock price, and that this earnings management, in turn, increases personal gains from executives' insider trading. Bergstresser and Philippon (2006) also report that managers with a large proportion of stock and options holdings are more likely to use discretionary accruals to manipulate earnings.[5]

Another form of gaming could be voluntary disclosure management. Prior research (e.g., Yermack 1997; Aboody and Kaznik 2000) shows that the timing of corporate voluntary disclosures is related to the granting of stock options. Aboody and Kaznik (2000) observe that CEOs who receive their options before earnings are announced are significantly more likely to issue bad news forecasts and less likely to issue good news forecasts than are CEOs who receive their awards after the earnings announcement.[6] They also find that management forecasts issued during the 3 months prior to scheduled awards are significantly less optimistically biased than forecasts issued for the same firms during other months. Because managers receiving stock options employ such opportunistic disclosure strategies, the complexity of the forecasting task increases as stock options pay increases. This is especially true when CEOs get multiple stock options grants during the same fiscal year.

The above arguments suggest that an increase in the level of stock options in CEO compensation likely increases the difficulty of forecasting, resulting in less accurate forecasts. However, recent research provides evidence suggesting that stock options can improve the alignment between managers' and shareholders' interests. Hanlon et al. (2003) examine whether stock options granted to the top five executives are related to future operating earnings. Their results indicate that each

[5]However, in related research, Hribar and Nichols (2007) provide evidence that not controlling for operating volatility increases the risk of over-rejecting the null hypothesis of no earnings management.

[6]Although Aboody and Kaznik (2000) study only fixed schedule awards, we argue that the incentive to maximize the stock options pay by manipulating the stock price at the grant date is present for all stock options awards, and that the incentive is especially strong for firms that make multiple grants in a given year. We note that a large number of our sample firms made multiple grants in the same fiscal year thus increasing this incentive. It is also interesting to note that the stock options award dates are generally not publicly known until the issue of proxy statements which are available only 2–3 months after the fiscal year-end (Yermack 1997).

dollar of stock options granted is associated with more than one dollar of future operating earnings over the next 5 years.

Cheng and Warfield (2005) examine the link between stock options and stock ownership and earnings management. They provide evidence that managers with high equity incentives (stock options and stock ownership) are more likely to provide reported earnings meeting or just beating analysts' forecasts and are less likely to report large positive earnings surprises, suggesting that stock options motivate types of earnings management that might increase forecast accuracy.

Given that the literature provides mixed evidence, we present our hypothesis in null form:

H1 Analysts' earnings forecast accuracy is unrelated to the level of stock options in CEO pay.

The alternate hypothesis is that earnings forecast accuracy is either positively or negatively related to the level of stock options in CEO pay. Given that the alternate hypothesis is nondirectional, we test H1 using a two-tailed test.

97.2.2 CEO Stock Options and Forecast Bias

Lim (2001) presents a model demonstrating that statistically optimal forecasts, in terms of mean squared error, may be positively and predictably biased. In this model, analysts trade off forecast bias for forecast accuracy. Analysts have incentives to provide accurate forecasts in order to increase their compensation and market value. They need access to management's private information to improve their forecast accuracy. Consequently, they have to maintain favorable relations with management to ensure that they have access to such information. One way of accomplishing this is by issuing optimistically biased forecasts. Although forecast bias by itself is not desirable, analysts can increase their forecast accuracy by incorporating in their forecasts the private information that management makes available to them as a reward for their forecast optimism. And, because access to management's private information is more valuable when firms' earnings are less predictable, analysts have greater incentives to issue optimistic forecasts for such firms.

Lim (2001) and Das et al. (1998) provide empirical evidence that analysts' forecasts are more optimistically biased when earnings are less predictable and the forecasting task is more complex. Recent examples of research supporting the management relations hypothesis include Chen and Matsumoto (2006) and Ke and Yu (2006). Chen and Matsumoto (2006) find that analysts issuing more favorable recommendations experience a greater increase in their relative forecast accuracy compared with analysts who issue less favorable recommendations. Ke and Yu (2006) document that analysts produce more accurate forecasts and are less likely to be fired by their employers when they issue initial optimistic earnings forecasts.

A higher level of CEO stock options pay may increase the forecasting complexity because it may induce managers to undertake riskier investments, to change/reallocate their effort, to manipulate accounting earnings, and to issue opportunistic voluntary disclosures. Therefore, we expect analysts' optimistic forecast bias to increase as the level of stock options in CEO pay increases. However, as discussed in the previous section, there is also evidence suggesting that stock options can align the incentives of management and shareholders (see, e.g., Hanlon et al. 2003), which, in turn, may lead to better disclosure and less optimistic bias. Accordingly, we provide the following hypothesis stated in null form:

H2 Analysts' optimistic earnings forecast bias is unrelated to the level of stock options in CEO pay.

The alternate hypothesis is that earnings forecast bias either increases or decreases with the level of stock options in CEO pay. Given that the alternate hypothesis is nondirectional, we test H2 using a two-tailed test.

97.3 Data Description and Research Design

97.3.1 Sample Selection

Our sample comprises firms with data available in the ExecuComp, I/B/E/S, and Compustat databases for the period 1993–2003. We exclude financial institutions and agricultural firms[7] and observations for which the CEOs are not identified in ExecuComp. These selection criteria result in an initial sample of 6,272 firm-year observations.

We obtain compensation data from ExecuComp, earnings forecasts, actual earnings, and stock prices from I/B/E/S. I/B/E/S forecasts generally exclude extraordinary and other special items. By using forecast and actual earnings from I/B/E/S, we ensure greater consistency between these two variables (Philbrick and Ricks 1991). We test our hypotheses using forecasts made 9 months before the earnings announcement.[8] We obtain other required financial data from Compustat.

[7]To be consistent with the prior literature on executive pay (Core et al. 1999; Hanlon et al. 2003), we omit financial institutions and agricultural companies. However, for completeness, we also conducted the analysis with these companies included in the sample. Our main conclusions are unaffected by this inclusion.

[8]We repeat all our analyses using 3-month-ahead forecasts to examine the robustness of our results to the forecast horizon. Most results for the 3-month forecasts mirror those presented for the 9-month forecasts.

Table 97.1 Sample selection procedure

Total firm-year observations on ExecuComp data for years 1993–2003, where the CEO is identified and without financial and agricultural firms	11,016
Less: Observations lost when merging with I/B/E/S data	(4,744)
Subtotal	6,272
Less: Observations with total compensation less than $1 million	(718)
Subtotal	5,554
Less: Observations with missing financial data	(553)
Less: Deletion of extreme values and other restrictions	(568)
Final sample for tests on forecast accuracy	**4,433**

Additionally, we require firms to have stock prices greater than one dollar to avoid the small deflator problem, to have at least three analysts' forecasts available to obtain a reliable estimate of forecast dispersion, and to have total CEO pay of at least one million dollars to avoid including firms that have low CEO incentives. These restrictions, along with deletion of observations with values in the top and bottom 1 % of the variables used in the regressions, result in 4,433 firm-year observations for the 9-month-ahead forecast accuracy tests and 4,279 firm-year observations for the 9-month-ahead forecast bias tests.[9] Table 97.1 summarizes our sample selection criteria.

97.3.2 Variable Measurement

97.3.2.1 CEO Options (OPTIONS)

We measure the level of CEO options pay (OPTIONS) as the ratio of the number of options to total number of shares outstanding. This is consistent with the measure employed by Cheng and Warfield (2005). It is a simple parsimonious variable that measures the relative proportion of stock options in a CEO's compensation for a particular year. We use four measures of CEO options: new options, existing exercisable options, existing unexercisable options, and total options (sum of the previous three), all scaled by total number of shares outstanding. We also use both contemporaneous as well as lagged values of OPTIONS in our main tests.[10]

97.3.2.2 Forecast Accuracy (ACCURACY)

We measure forecast accuracy for the 9-month-ahead forecast as minus one times the absolute value of the deviation of the mean EPS forecast from the actual EPS for

[9]The sample firms represent a variety of industries, with the largest representation being retail (8 %), electronic equipment (6.7 %), business services (5.3 %), and telecommunications (6 %).

[10]As a sensitivity check, when using lagged OPTIONS, we delete observations that have a new CEO in the current year. Our main results are robust to deletion of these observations.

that year divided by stock price at the forecast date. This measure increases with forecast accuracy and is defined as

$$\mathrm{ACCURACY_{it}} = (-1) * \frac{\left|\mathrm{FEPS_{it}^{t-1}} - \mathrm{AEPS_{it}}\right|}{\mathrm{P_{i,t-1}}} \quad (97.1)$$

where, for firm i, $\mathrm{ACCURACY_t}$ is minus one times the absolute forecast error at time t, $\mathrm{FEPS_t^{t-1}}$ is the mean EPS forecast from I/B/E/S for year t made at time t − 1 (i.e., 9 months prior to the fiscal year-end), $\mathrm{AEPS_t}$ is the actual earnings per share obtained from I/B/E/S, and $\mathrm{P_{t-1}}$ is the stock price at the time of the forecast obtained from I/B/E/S.

97.3.2.3 Forecast Bias (BIAS)

We measure forecast bias for firm i at time t as the difference between the mean EPS forecast made 9 months prior to the fiscal year-end and the actual EPS, divided by stock price at the forecast date:

$$\mathrm{BIAS_{it}} = \frac{\mathrm{FEPS_{it}^{t-1}} - \mathrm{AEPS_{it}}}{\mathrm{P_{i,t-1}}} \quad (97.2)$$

The variable BIAS increases as the level of forecast optimism increases.

97.3.2.4 Other Factors Affecting Forecast Accuracy and Bias

In our empirical tests, we control for previously identified determinants of forecast accuracy and bias. These include earnings characteristics, firm characteristics, and forecast characteristics.

97.3.2.5 Earnings Characteristics

Prior research (Lang and Lundholm 1996; Das et al. 1998; Brown 2001; Duru and Reeb 2002) identifies earnings volatility (VOLROA), whether firms have accounting losses (LOSS), and absolute earnings surprise (ABSESUP) as earnings characteristics that negatively affect forecast accuracy. Kross et al. (1990) and Lim (2001) provide evidence that long-term earnings volatility is associated with less accurate forecasts. This is because the forecasting task is more difficult for firms with historically volatile earnings compared to firms with historically more stable earnings. Prior studies (Brown 2001) also document that analysts' forecasts of firms with losses are generally less accurate than those of firms with profits, partly due to the increased difficulty with estimating losses stemming from other managerial incentives such as "big baths." Lang and Lundholm (1996) and Duru and Reeb (2002), among others, find that larger earnings surprises are associated with less accurate forecasts. This may be due to the effect of anchoring to previously reported earnings.

Consistent with prior research, we measure VOLROA as the standard deviation of return on assets estimated using data from the prior 5 years, LOSS as an indicator variable representing firm-years with reported losses, and ABSESUP as the absolute value of the difference between the current year's EPS and the previous year's EPS, divided by stock price at the beginning of the year.

The same earnings characteristics also affect analysts' optimistic bias. In addition, Gu and Wu (2003) provide evidence that optimistic forecast bias is negatively associated with earnings skewness (SKEW) primarily due to mean-median differences in skewed earnings distributions. Prior research reports that earnings predictability variables, such as earnings volatility and whether the firm is a loss firm, are positively related to optimistic bias (Duru and Reeb 2002). However, after controlling for the level of earnings, Eames and Glover (2003) do not find a significant relation between forecast error and earnings predictability. Given this finding, we include level of earnings (LEVEARN), measured as annual earnings scaled by year-end market value of equity, as a control variable.[11] We include earnings surprise (ESUP) in place of absolute earnings surprise as a control variable for bias. This variable is measured as the difference between the current year's EPS and the previous year's EPS, divided by the price at the beginning of the year. We also include negative earnings surprise (NEGESUP), where NEGESUP equals 0 if ESUP is positive and equals ESUP if ESUP is negative. We include ESUP and NEGESUP to control for the anchoring behavior of analysts who tend to anchor their forecasts closely to previous period's actual results.

97.3.2.6 Firm Characteristics

We include size (SIZE), growth (GROWTH, CHASSETS, and CHSALES), and governance variables (SHROWN, CEOCHAIR, NUMMTGS, and DIROPT) as firm-specific control variables that are likely to be related to forecast accuracy and bias. We measure SIZE as the natural log of assets at the beginning of the year. Previous studies report that SIZE is related to analyst forecast accuracy and bias (Duru and Reeb 2002; Gu and Wu 2003; Ho and Tsay 2004). Because large firms have a richer information environment, we expect a positive relation between accuracy and size. From a strategic reporting bias standpoint, analysts have stronger incentives to issue optimistic forecasts for smaller firms to facilitate management communication since there is less public information available for these firms (Lim 2001; Gu and Wu 2003). Therefore, we expect the coefficient on SIZE to be negatively related to optimistic forecast bias.

We include the following three proxies for growth, GROWTH, CHASSETS, and CHSALES, where GROWTH is the ratio of the book value of equity at the beginning of the year to the market value of equity at the beginning of the year (i.e., the book-to-market ratio), CHASSETS is the percentage change in total assets at the beginning of the year, and CHSALES is the percentage

[11] Gu (2003) argues that inclusion of earnings level as a control variable will induce spurious relationships between the variable capturing forecast efficiency and other control variables. Our main results are stronger when we exclude earnings level as a control variable.

change in annual sales at the beginning of the year.[12] We include these measures of growth as control variables because firm growth is an important driver of forecast complexity; however, we do not offer directional predictions for these proxies for growth.[13]

Prior research (Byard et al. 2006) shows that analyst forecast accuracy is related to a firm's corporate governance quality. We include SHROWN, CEOCHAIR, NUMMTGS, and DIROPT as proxies for a firm's governance quality. We measure SHROWN as the number of shares owned by the CEO divided by the total number of shares outstanding. Extant evidence indicates that insider ownership generally serves to align insiders' interests with those of shareholders (e.g., McConnell and Servaes 1990). Additionally, higher insider ownership is negatively associated with earnings management (Warfield et al. 1995). Higher insider ownership induces insiders, including the CEO, to maximize shareholder wealth, thereby mitigating the agency problem. Therefore, we expect SHROWN to be positively related to accuracy. The variable CEOCHAIR is an indicator variable which equals one if the CEO is also the chairman of the board and zero otherwise. A CEO who is also chairman of the board of directors could undermine the effectiveness of the board by dissuading directors from expressing alternative viewpoints. Separation of the positions of Chairman and CEO is also an important indicator of board independence (Jensen 1993; Daily and Dalton 1997). Combining these positions leads to a conflict of interest and impairs the board's independence and effectiveness in executing its oversight and governance responsibilities.

The variable NUMMTGS indicates the number of board meetings held in a year. Boards that meet more frequently should be more effective monitors of management (Conger et al. 1998). In addition, Xie et al. (2003) find that the level of earnings management is lower for companies whose boards meet more frequently. The variable DIROPT is an indicator variable which equals one if the directors are awarded stock options in the year and zero otherwise. Directors receiving higher pay may be less vigilant in monitoring the management. This is especially true if the directors were appointed to the board by the same CEO. Although CEOCHAIR, NUMMTGS, and DIROPT are important proxies of a firm's governance quality, we do not offer directional predictions for these variables.

97.3.2.7 Forecast Characteristics

Prior research has identified two forecast characteristics that are related to forecast accuracy and bias – number of analysts following a firm (Duru and Reeb 2002; Das et al. 1998) and dispersion in analysts' forecasts (Lang and Lundholm 1996; Gu and Wu 2003). This research finds that forecast accuracy is higher for firms with larger analyst following (FOLLOW) and lower for firms with

[12]Prior research on forecast accuracy and bias (e.g., Duru and Reeb 2002) does not control for differences in growth. Our results are robust to the exclusion of GROWTH as a control variable in the regressions.

[13]The variance inflation factors for variables in our main regressions are all below three, indicating that there are no severe multicollinearity problems.

higher analysts' forecast dispersion (DISP). It also documents that optimistic forecast bias decreases with analyst following and increases with forecast dispersion.

Forecast dispersion may also proxy for the degree of difficulty in forecasting earnings, with high analysts' forecast dispersion firms exhibiting lower levels of earnings predictability. Therefore, we also expect forecast dispersion to be negatively related to forecast accuracy and positively related to forecast bias.

97.3.3 Empirical Models

We estimate the following regression model to test our first hypothesis on forecast accuracy:

$$\begin{aligned}\mathrm{ACCURACY}_{it} = {} & a_0 + a_1\mathrm{OPTIONS}_{it} + a_2\mathrm{VOLROA}_{i,t-1} + a_3\mathrm{LOSS}_{it} + a_4\mathrm{ABSESUP}_{it} \\ & + a_5\mathrm{GROWTH}_{i,t-1} + a_6\mathrm{SIZE}_{i,t-1} + a_7\mathrm{FOLLOW}_{i,t-1} + a_8\mathrm{DISP}_{i,t-1} \\ & + a_9\mathrm{SHROWN}_{it} + a_{10}\mathrm{CEOCHAIR}_{it} + a_{11}\mathrm{NUMMTGS}_{it} \\ & + a_{12}\mathrm{DIROPT}_{it} + a_{13}\mathrm{GROWTH}_{i,t-1} + a_{14}\mathrm{CHASSETS}_{i,t-1} \\ & + a_{15}\mathrm{CHSALES}_{i,t-1} + \langle\text{Industry controls}\rangle + \langle\text{Year controls}\rangle \\ & + \varepsilon_{it}\end{aligned} \tag{97.3}$$

We include industry and year indicator variables to control for industry and year fixed effects. We employ the 48 industries (other than financial and agricultural) identified by Fama and French (1997) as our industry categories.

Hypothesis 1 predicts that forecast accuracy changes with the level of CEO stock options pay. Therefore, we expect a_1, the coefficient on OPTIONS, to be different from zero. Recall that we are using four measures of OPTIONS: new options, existing exercisable options, existing unexercisable options, and total options (sum of the previous three). We use two models to test our hypothesis. We use total options as the measure of OPTIONS in the first model, and we use new options, existing exercisable options, and existing unexercisable options as measures of OPTIONS in the second model. In addition, we use both contemporaneous as well as lagged values of OPTIONS in our tests.

We include VOLROA, LOSS, and ABSESUP in the model to control for cross-sectional differences in earnings characteristics because prior research documents that these variables affect forecast accuracy. We include SIZE, GROWTH, CHASSETS, and CHSALES to control for differences in firm characteristics, and we include FOLLOW and DISP to account for the effects of differences in forecast characteristics on forecast accuracy. We include SHROWN, CEOCHAIR, NUMMTGS, and DIROPT to control for cross-sectional differences in corporate governance. We discuss the expected signs on the control variables and present their definitions in Sect. 97.3.2.

We estimate the following regression model to test our second hypothesis on forecast bias:

$$\begin{aligned}\mathrm{BIAS}_{it} = {} & a_0 + a_1\mathrm{OPTIONS}_{it} + a_2\mathrm{SKEW}_{i,t-1} + a_3\mathrm{VOLROA}_{i,t-1} + a_4\mathrm{LOSS}_{it}\\ & + a_5\mathrm{LEVEARN}_{it} + a_6\mathrm{ESUP}_{it} + a_7\mathrm{NEGESUP}_{it} + a_8\mathrm{SIZE}_{i,t-1}\\ & + a_9\mathrm{FOLLOW}_{it} + a_{10}\mathrm{DISP}_{it} + a_{11}\mathrm{SHROWN}_{it} + a_{12}\mathrm{CEOCHAIR}_{it}\\ & + a_{13}\mathrm{NUMMTGS}_{it} + a_{14}\mathrm{DIROPT}_{it} + a_{15}\mathrm{GROWTH}_{i,t-1} + a_{16}\mathrm{CHASSETS}_{i,t-1}\\ & + a_{17}\mathrm{CHSALES}_{i,t-1} + \langle\text{Industry controls}\rangle + \langle\text{Year controls}\rangle + \varepsilon_{it}\end{aligned} \tag{97.4}$$

Hypothesis 2 predicts that forecast bias increases or decreases as the level of CEO stock options pay changes. Therefore, we expect a_1, the coefficient on OPTIONS, to be different from zero. Once again, we use two models to test hypothesis 2. The first model uses total options as the measure of OPTIONS. The second model uses new options, existing exercisable options, and existing unexercisable options as measures of OPTIONS. We also include both contemporaneous and lagged values of these four measures.

We include SKEW, VOLROA, LOSS, LEVEARN, ESUP, and NEGESUP in the model to control for cross-sectional differences in earnings characteristics that have been shown in prior research to affect forecast bias. We include SIZE, GROWTH, CHASSETS, and CHSALES to control for differences in firm characteristics, and we include FOLLOW and DISP to account for differences in forecast characteristics that affect forecast bias. We include SHROWN, CEOCHAIR, NUMMTGS, and DIROPT to control for cross-sectional differences in corporate governance. We discuss the expected signs on the control variables and present their definitions in Sect. 97.3.2.

97.4 Empirical Analysis

97.4.1 Descriptive Statistics

Table 97.2 presents descriptive statistics for the key variables. Consistent with prior research, the mean forecast bias (BIAS) is positive and 0.64 % of stock price. The median value of forecast bias is 0.37 %, which is consistent with prior research. The mean and median values of forecast accuracy (ACCURACY) are negative by construction. The mean value of forecast accuracy is −1.30 % of stock price. The level of new CEO stock options (NOPT) has a mean value of 0.21 % and a median value of 0.11 % indicating that, on average, a significant amount of CEO stock options are awarded relative to the number of shares outstanding.

Table 97.3 presents the correlation matrix for the stock options variables used in the regression analysis. Forecast accuracy is significantly negatively related to all

Table 97.2 Descriptive statistics on CEO options, forecast accuracy, forecast bias, and control variables

Variable	Obs.	Mean	Median	Standard deviation	25 %	75 %
1. CEO options						
NOPT (%)	4,433	0.2146	0.1120	0.3644	0.0363	0.2535
NOPT, $t-1$ (%)	4,433	0.2070	0.0974	0.3831	0.0238	0.2342
EOPT (%)	4,433	0.5839	0.3205	0.7980	0.1019	0.7451
EOPT, $t-1$ (%)	4,433	0.5320	0.2801	0.7535	0.0822	0.6798
UEOPT (%)	4,433	0.4369	0.2474	0.6246	0.0856	0.5623
UEOPT, $t-1$ (%)	4,433	0.4177	0.2229	0.6042	0.0701	0.5402
TOPT (%)	4,433	1.2356	0.8285	1.3963	0.3394	1.6435
TOPT, $t-1$ (%)	4,433	1.1561	0.7411	1.3698	0.2912	1.5276
2. Forecast variables						
ACCURACY	4,433	−0.0129	−0.0067	0.0186	−0.0142	−0.0026
BIAS	4,279	0.0064	0.0037	0.0216	−0.0017	0.0100
FOLLOW	4,433	17.2673	15.0000	9.7937	9.0000	23.0000
DISP	4,433	0.0040	0.0021	0.0059	0.0010	0.0044
SKEW	4,279	0.0000	0.0000	0.0011	−0.0004	0.0004
3. Control variables						
VOLROA	4,433	0.0365	0.0252	0.0392	0.0115	0.0480
ABSESUP	4,433	0.0319	0.0156	0.0480	0.0069	0.0360
ESUP	4,279	0.0007	0.0034	0.0468	−0.0150	0.0150
LOSS	4,433	0.1066	0.0000	0.3087	0.0000	0.0000
LEVEARN	4,279	0.0418	0.0486	0.0604	0.0285	0.0690
SIZE	4,433	7.8174	7.755	1.6550	6.6395	8.9370
GROWTH	4,433	0.2645	0.1126	0.5428	0.0456	0.2907
CHASSETS (%)	4,433	16.9140	8.8190	37.04481	1.7000	21.0000

CHSALES (%)	4,433	14.1356	8.9430	30.5571	1.8000	20.0000
SHROWN (%)	4,433	1.5556	0.2335	3.9541	0.0763	0.9024
CEOCHAIR	4,433	0.6932	1.000	0.4612	0.0000	1.000
NUMMTGS	4,433	7.0746	7.0000	2.8362	5.0000	9.0000
DIROPT	4,433	0.5736	1.000	0.4946	0.000	1.000

Variable definitions:
NOPT = new stock options to CEO in the current year scaled by total number of shares outstanding
EOPT = existing exercisable stock options of CEO scaled by total number of shares outstanding
UEOPT = existing unexercisable stock options of CEO scaled by total number of shares outstanding
TOPT = sum of *NOPT*, *EOPT*, and *UEOPT*
ACCURACY = (−1) * absolute value of [mean EPS forecast − actual EPS]/price at forecast date
BIAS = signed forecast error measured as [mean EPS forecast − actual EPS]/price at forecast date
FOLLOW = number of analysts following the firm
DISP = forecast dispersion, measured as the standard deviation of analysts' forecasts deflated by price at the forecast date
SKEW = difference between the mean and the median forecast scaled by price at the forecast date
VOLROA = earnings volatility measured as the standard deviation of return on assets for the previous 5-year period
ABSESUP = earnings surprise measured as the absolute value of the difference between the current year's EPS and the last year's EPS, divided by price at the beginning of the year
ESUP = change in earnings (CHG_EPS) measured as the difference between the current year's EPS and the last year's EPS, divided by price at the beginning of the year
LOSS = a dummy variable which equals to 1 when earnings are negative and 0 otherwise
LEVEARN = annual earnings scaled by the year-end market value of equity
SIZE = firm size measured as the natural log of beginning assets
GROWTH = beginning book value of equity divided by the beginning market value of equity
CHASSETS = annual percentage change in total assets at the beginning of the year
CHSALES = annual percentage change in total sales at the beginning of the year
SHROWN = number of shares owned by the CEO divided by the total number of shares outstanding
CEOCHAIR = an indicator variable which equals "1" if the CEO is also the chairman of the board and "0" otherwise
NUMMTGS = the number of board meetings held in a year
DIROPT = an indicator variable which equals "1" if the directors are awarded stock options in the year and "0" otherwise

Table 97.3 Correlation matrix and Pearson correlation coefficients

	NOPT	NOPT, $t-1$	EOPT	EOPT, $t-1$	UEOPT	UEOPT, $t-1$	TOPT	TOPT, $t-1$
ACCURACY	−0.173 (0.00)	−0.125 (0.00)	−0.128 (0.00)	−0.106 (0.00)	−0.132 (0.00)	−0.116 (0.00)	−0.177 (0.00)	−0.144 (0.00)
BIAS	0.131 (0.00)	0.095 (0.00)	0.102 (0.00)	0.080 (0.00)	0.089 (0.00)	0.077 (0.00)	0.132 (0.00)	0.104 (0.00)
NOPT	1.000	0.337 (0.00)	0.373 (0.00)	0.301 (0.00)	0.546 (0.00)	0.346 (0.00)	0.717 (0.00)	0.412 (0.00)
NOPT, $t-1$		1.000	0.358 (0.00)	0.320 (0.00)	0.528 (0.00)	0.676 (0.00)	0.529 (0.00)	0.752 (0.00)
EOPT			1.000	0.871 (0.00)	0.328 (0.00)	0.431 (0.00)	0.815 (0.00)	0.767 (0.00)
EOPT, $t-1$				1.000	0.287 (0.00)	0.341 (0.00)	0.704 (0.00)	0.786 (0.00)
UEOPT					1.000	0.771 (0.00)	0.778 (0.00)	0.644 (0.00)
UEOPT, $t-1$						1.000	0.682 (0.00)	0.817 (0.00)
TOPT							1.000	0.833 (0.00)

p-values are reported in parentheses
Variable definitions:
ACCURACY = (−1) * absolute value of [mean EPS forecast − actual EPS]/price at forecast date
BIAS = signed forecast error measured as [mean EPS forecast − actual EPS]/price at forecast date
NOPT = new stock options to CEO in the current year scaled by total number of shares outstanding
EOPT = existing exercisable stock options of CEO scaled by total number of shares outstanding
UEOPT = existing unexercisable stock options of CEO scaled by total number of shares outstanding
TOPT = sum of *NOPT*, *EOPT*, and *UEOPT*

four components of CEO options, and forecast bias is significantly positively related to all four components.

97.4.2 Estimation Results

Tables 97.4 and 97.5 present estimation results of our empirical models on forecast accuracy and bias respectively. In each table, we present the results for our two models, one using total options and the other using new options, existing exercisable options, and existing unexercisable options. We present the results using contemporaneous measures of OPTIONS in Panel A and the results using lagged values in Panel B.

Table 97.4 presents estimation results for Eq. 97.3. This specification is used to test our first hypothesis that analyst forecast accuracy is unrelated to CEO stock options.[14] For the total options measure (model 1 in Table 97.4), the coefficient on TOPT is negative and statistically significant at the 1 % level for both contemporaneous and lagged values, indicating rejection of hypothesis H1. Model 2 results also indicate rejection of hypothesis H1 as both new options (NOPT) and existing exercisable options (EOPT) have strong negative relations (significant at the 5 % level or better) with forecast accuracy. Additionally, the relations between forecast accuracy and the control variables are generally consistent with expectations and with prior research. In particular, negative earnings (LOSS), absolute earnings surprise (ABSESUP), and dispersion in analysts' forecasts (DISP) are all negatively related to forecast accuracy as predicted, and their coefficients are statistically significant at the 1 % level in both Panel A and Panel B. Earnings volatility (VOLROA) is also negative, as expected, and significant at the 5 % level in all regressions.

Analyst following (FOLLOW) and firm size (SIZE) are positively related to forecast accuracy, as expected, but neither is significant. Among the variables without predicted signs, number of board meetings (NUMMTGS) is negative and significant at 1 % when contemporaneous values are used, and change in sales (CHSALES) is negative and significant for all specifications.

The above findings are consistent with the notion that the increased forecasting complexity accompanying the increase in CEO stock options compensation adversely affects the forecasting ability of financial analysts. This adverse effect leads to a decline in the accuracy of their forecasts. The above result strongly holds for total options in model 1 and new options and existing exercisable options in model 2.

Table 97.5 presents estimation results for Eq. 97.4. This specification is used to test hypothesis H2. If a_1, the coefficient on OPTIONS, is significantly different

[14]We also estimate the regression without the absolute value of earnings surprise (ABSESUP) that might be mechanically related to accuracy. The main results are not affected by the exclusion of that variable. We note that inclusion of ABSESUP can only weaken the hypothesized relationship between ACCURACY and OPTIONS because ABSESUP is closely related to ACCURACY.

Table 97.4 Stock options and forecast accuracy

		Model 1		Model 2	
	Sign	Coefficient	t-Statistics	Coefficient	t-Statistics
Panel A: Contemporaneous options					
TOPT	+/−	**−0.1162**	**−5.74*****		
NOPT	+/−			**−0.1336**	**−2.63*****
EOPT	+/−			**−0.1300**	**−3.79*****
UEOPT	+/−			**−0.0882**	**−1.85***
VOLROA, t − 1	−	−0.0073	−2.33**	−0.0169	−2.33**
LOSS	−	−0.0098	−11.45***	−0.0098	−11.42***
ABSESUP	−	−0.0988	−17.67***	−0.0987	−17.65***
SIZE, t − 1	+	0.0004	1.39	0.0004	1.40
FOLLOW	+	0.0000	0.58	0.0000	0.55
DISP	−	−0.6203	−13.67***	−0.6191	−13.63***
SHROWN	+	−0.0031	−0.49	−0.0029	−0.46
CEOCHAIR	?	−0.0006	−1.04	−0.0005	−1.00
NUMMTGS	?	−0.0002	−2.78***	−0.0002	−2.79***
DIROPT	?	0.0007	1.41	0.0007	1.39
GROWTH, t − 1	?	0.0752	1.45	0.0741	1.43
CHASSETS, t − 1	?	−0.0000	−0.37	−0.0000	−0.38
CHSALES, t − 1	?	−0.0000	−1.98**	−0.0000	−1.99**
INTERCEPT	?	−0.0042	−1.52	−0.0042	−1.52
Industry control		Yes		Yes	
Year control		Yes		Yes	
Observations		4,433		4,433	
F-value		30.84		29.93	
Adjusted R^2		30.77 %		30.74 %	
Panel B: Lagged (previous year's) options					
TOPT, t − 1	+/−	**−0.1058**	**−5.21*****		
NOPT, t − 1	+/−			**−0.1308**	**−3.64*****
EOPT, t − 1	+/−			**−0.1223**	**2.18****
UEOPT, t − 1	+/−			**−0.0307**	**−0.36**
VOLROA, t − 1	−	−0.0164	−2.27**	−0.0166	−2.30**
LOSS	−	−0.0102	−11.91***	−0.0103	−11.94***
ABSESUP	−	−0.1011	−17.97***	−0.1011	−17.95***
SIZE, t − 1	+	0.0002	0.66	0.0002	0.64
FOLLOW	+	0.0000	0.78	0.0000	0.73
DISP	−	−0.6114	−13.46***	−0.6131	−13.42***
SHROWN	+	−0.0062	−1.07	−0.0063	−1.08
CEOCHAIR	?	−0.0002	−0.42	−0.0002	−0.37
NUMMTGS	?	0.0001	0.75	0.0001	0.70
DIROPT	?	0.0005	0.91	0.0005	0.94
GROWTH, t − 1	?	0.0564	1.09	0.0556	1.07
CHASSETS, t − 1	?	−0.0000	−0.09	−0.0000	−0.13

(*continued*)

Table 97.4 (continued)

		Model 1		Model 2	
	Sign	Coefficient	t-Statistics	Coefficient	t-Statistics
***CHSALES, t* − 1**	?	−0.0000	−2.50**	−0.0000	−2.48**
INTERCEPT	?	−0.0045	−1.62	−0.0043	−1.53
Industry control		Yes		Yes	
Year control		Yes		Yes	
Observations		4,425		4,425	
F-value		30.29		29.41	
Adjusted R^2		30.45 %		30.40 %	

Variable definitions:
NOPT = new stock options to CEO in the current year scaled by total number of shares outstanding
EOPT = existing exercisable stock options of CEO scaled by total number of shares outstanding
UEOPT = existing unexercisable stock options of CEO scaled by total number of shares outstanding
TOPT = sum of *NOPT*, *EOPT*, and *UEOPT*
ACCURACY = (−1) * absolute value of [mean EPS forecast − actual EPS]/price at forecast date
FOLLOW = number of analysts following the firm
DISP = forecast dispersion, measured as the standard deviation of analysts' forecasts deflated by price at the forecast date
VOLROA = earnings volatility measured as the standard deviation of return on assets for the previous 5-year period
ABSESUP = earnings surprise measured as the absolute value of the difference between the current year's EPS and the last year's EPS, divided by price at the beginning of the year
LOSS = a dummy variable which equals to 1 when earnings are negative and 0 otherwise
SIZE = firm size measured as the natural log of beginning assets
GROWTH = beginning book value of equity divided by the beginning market value of equity
CHASSETS = annual percentage change in total assets at the beginning of the year
CHSALES = annual percentage change in total sales at the beginning of the year
SHROWN = number of shares owned by the CEO divided by the total number of shares outstanding
CEOCHAIR = an indicator variable which equals "1" if the CEO is also the chairman of the board and "0" otherwise
NUMMTGS = the number of board meetings held in a year
DIROPT = an indicator variable which equals "1" if the directors are awarded stock options in the year and "0" otherwise
***Significant at the 0.01 level, **Significant at the 0.05 level, *Significant at the 0.10 level. Significance levels are based on two-tailed tests

from zero, the empirical results will reject our hypothesis H2 stated in the null. Once again, to test H2, we use two models, one with total options for OPTIONS and the second with new options, existing exercisable options, and existing unexercisable options for OPTIONS. As before, we present the results for both contemporaneous (Panel A) and lagged (Panel B) values of OPTIONS.[15]

[15]We also estimate the model without earnings surprise (ESUP), negative earnings surprise (NEGESUP), and level of earnings (LEVEARN) that might be mechanically related to bias. The relationship between level of options and bias is not affected by the omission of those variables.

Table 97.5 Stock options and forecast bias (optimism)

	Sign	Model 1 Coefficient	Model 1 t-Statistics	Model 2 Coefficient	Model 2 t-Statistics
Panel A: Contemporaneous options					
TOPT	+/−	**0.1245**	**4.78*** **		
NOPT	+/−			**0.1061**	**3.56*** **
EOPT	+/−			**0.1007**	**2.28** **
UEOPT	+/−			**0.0223**	**0.37**
SKEW, t − 1	−	−0.3056	−1.14	−0.3105	−1.16
VOLROA, t − 1	?	0.0382	3.95***	0.0378	3.91***
LOSS	?	0.0107	7.56***	0.0104	7.39***
LEVEARN	?	−0.0615	−8.55***	−0.0615	−8.55***
ESUP	?	−0.0542	−6.23***	−0.0541	−6.22***
NEGESUP	?	0.0026	3.39***	0.0026	3.36***
SIZE, t − 1	−	−0.0005	−1.47	−0.0005	−1.43
FOLLOW	−	−0.0000	−0.11	−0.0000	−0.09
DISP	+	0.0000	1.46	0.0000	1.49
SHROWN	−	−0.0099	−1.23	−0.0097	−1.21
CEOCHAIR	?	0.0007	1.03	0.0007	1.05
NUMMTGS	?	0.0000	0.13	0.0000	0.14
DIROPT	?	0.0001	0.11	0.0001	0.10
GROWTH, t − 1	?	−0.2095	−3.18***	−0.2149	−3.26***
CHASSETS, t − 1	?	−0.0000	−1.07	−0.0000	−1.09
CHSALES, t − 1	?	0.0000	0.35	0.0000	0.34
INTERCEPT	?	0.0080	2.25**	0.0081	2.27**
Industry control		Yes		Yes	
Year control		Yes		Yes	
Observations		4,279		4,279	
F-value		14.85		14.54	
Adjusted R^2		18.26 %		18.35 %	
Panel B: Lagged (previous year's) options					
TOPT, t − 1	+/−	**0.1283**	**4.95*** **		
NOPT, t − 1	+/−			**0.1273**	**3.95*** **
EOPT, t − 1	+/−			**0.1038**	**2.24** **
UEOPT, t − 1	+/−			**0.0294**	**0.41**
SKEW, t − 1	−	−0.2593	−0.96	−0.2639	−0.98
VOLROA, t − 1	?	0.0337	3.46***	0.0328	3.37***
LOSS	?	0.0115	8.11***	0.0113	8.01***
LEVEARN	?	−0.0607	−8.41***	−0.0605	−8.38***
ESUP	?	−0.0546	−6.16***	−0.0545	−6.15***
NEGESUP	?	0.0026	3.38***	0.0027	3.40***
SIZE, t − 1	−	−0.0005	−1.46	−0.0005	−1.55
FOLLOW	−	−0.0000	−0.19	−0.0000	−0.21
DISP	+	0.0000	1.51	0.0000	1.56

(*continued*)

Table 97.5 (continued)

		Model 1		Model 2	
	Sign	Coefficient	t-Statistics	Coefficient	t-Statistics
SHROWN	–	−0.0031	−0.42	−0.0035	−0.47
CEOCHAIR	?	0.0008	1.23	0.0009	1.32
NUMMTGS	?	0.0001	0.86	0.0001	0.71
DIROPT	?	−0.0000	−0.01	−0.0000	−0.07
GROWTH, t − **1**	?	−0.1978	−2.98***	−0.1927	−2.90***
CHASSETS, t − **1**	?	−0.0000	−1.24	−0.0000	−1.25
CHSALES, t − **1**	?	−0.0000	−0.01	0.0000	0.05
INTERCEPT	?	0.0072	2.01**	0.0078	2.18**
Industry control		Yes		Yes	
Year control		Yes		Yes	
Observations		4,264		4,264	
F-value		14.64		14.36	
Adjusted R^2		18.09 %		18.21 %	

Variable definitions:
NOPT = new stock options to CEO in the current year scaled by total number of shares outstanding
EOPT = existing exercisable stock options of CEO scaled by total number of shares outstanding
UEOPT = existing unexercisable stock options of CEO scaled by total number of shares outstanding
TOPT = sum of *NOPT*, *EOPT*, and *UEOPT*
BIAS = signed forecast error measured as [mean EPS forecast − actual EPS]/price at forecast date
FOLLOW = number of analysts following the firm
DISP = forecast dispersion, measured as the standard deviation of analysts' forecasts deflated by price at the forecast date
SKEW = difference between the mean and the median forecast scaled by price at the forecast date
VOLROA = earnings volatility measured as the standard deviation of return on assets for the previous 5-year period
ABSESUP = earnings surprise measured as the absolute value of the difference between the current year's EPS and the last year's EPS, divided by price at the beginning of the year
ESUP = change in earnings (CHG_EPS) measured as the difference between the current year's EPS and the last year's EPS, divided by price at the beginning of the year
LOSS = a dummy variable which equals to 1 when earnings are negative and 0 otherwise
LEVEARN = annual earnings scaled by the year-end market value of equity
SIZE = firm size measured as the natural log of beginning assets
GROWTH = beginning book value of equity divided by the beginning market value of equity
CHASSETS = annual percentage change in total assets at the beginning of the year
CHSALES = annual percentage change in total sales at the beginning of the year
SHROWN = number of shares owned by the CEO divided by the total number of shares outstanding
CEOCHAIR = an indicator variable which equals "1" if the CEO is also the chairman of the board and "0" otherwise
NUMMTGS = the number of board meetings held in a year
DIROPT = an indicator variable which equals "1" if the directors are awarded stock options in the year and "0" otherwise
***Significant at the 0.01 level, **Significant at the 0.05 level, *Significant at the 0.10 level. Significance levels are based on two-tailed tests

For the total options measure (model 1 in Table 97.5), the coefficient on TOPT is positive and statistically significant at the 1 % level for both contemporaneous and lagged values, indicating rejection of hypothesis H2. The model 2 results also indicate rejection of hypothesis H2 as both new options (NOPT) and existing exercisable options (EOPT) have strong positive relations (significant at the 5 % level or better) with forecast bias. These results are consistent with the management relations hypothesis.

Additionally, the relations between forecast bias and the control variables are generally consistent with the prior literature. We document a positive relation between forecast bias and volatility of return on assets (VOLROA) and loss firms (LOSS). We furthermore document a negative relation between forecast bias and growth (GROWTH) for all the regressions. These results are consistent with the results reported in prior research (Das et al. 1998; Duru and Reeb 2002; Eames and Glover 2003).

The evidence presented in Tables 97.4 and 97.5 supports the predictions that analysts' earnings forecast accuracy decreases and forecast optimism increases as the level of stock options (in particular, new options and exercisable options) in CEO pay increases. These results are robust to the time period (current versus prior year) when the options are granted.

97.4.3 Additional Analysis

As an additional test, we examine the relations between an alternate measure of CEO stock options and forecast accuracy and bias while controlling for endogeneity in the model. Since the results reported in Tables 97.4 and 97.5 are most pronounced for new options, we focus on an alternate proxy for new options, the proportion of stock options pay in CEO total compensation (compensation mix, COMPMIX). We measure this variable as the ratio of the Black-Scholes value of the new options granted to the CEO in a given year to the total compensation granted to the CEO in that year, where total compensation is the value of stock options plus cash compensation (i.e., salary plus bonus). This is consistent with the measure employed by Klassen and Mawani (2000) and Ittner et al. (2003). We also control for endogeneity as this may be a potential problem if both the level of CEO stock options pay and analysts' forecast accuracy/bias are determined by common variables such as a firm's fundamentals. To address this potential endogeneity problem, we examine the relations between CEO options and analysts' forecast accuracy/bias using two-stage least squares regression analysis (2SLS).

In the first stage, we regress the level of CEO options, the dependent variable, on previously identified, firm-specific determinants of stock options grants as well as control variables (Core et al. 1999; Aggarwal and Samwick 1999). These additional variables, related to cross-sectional differences in options compensation, include the prior year's return on assets (ROA), leverage ratio (LEV), and standard deviation of monthly returns for the prior 12-month period (STD). Three sets of variables are common to both regressions: growth (GROWTH), firm size (SIZE), and share

Table 97.6 2SLS estimation of compensation mix and forecast accuracy and bias

		Column 1		Column 2	
	Sign	Coefficient	t-Statistics	Coefficient	t-Statistics
Panel A: Compensation mix and forecast accuracy					
COMPMIX	NA/+/−			**−0.3776**	**−8.69*****
VOLROA, t − 1	NA/−			−0.0378	−0.89
LOSS	NA/−			−0.0084	−8.31***
ABSESUP	NA/−			−0.1094	−16.52***
SIZE, t − 1	+/+	0.0271	7.25***	0.0002	0.55
FOLLOW	NA/+			0.0001	1.80*
DISP	NA/−			−0.5852	−10.91***
SHROWN	−/+	−0.7432	−7.35***	−0.0030	−0.40
CEOCHAIR	NA/?			−0.0009	−1.37
NUMMTGS	NA/?			−0.0002	−2.03**
DIROPT	NA/?			0.0009	1.46
GROWTH, t − 1	+/?	3.0900	3.69***	0.0738	1.21
CHASSETS, t − 1	NA/?			0.0000	0.29
CHSALES, t − 1	NA/?			0.0000	−1.85
STD	+/NA	1.1769	15.36***		
ROA	+/NA	0.1501	2.54**		
LEV	−/NA	−0.1513	−5.19***		
INTERCEPT	?/?	0.1972	5.40***	0.0097	2.74***
Industry control		Yes		Yes	
Year control		Yes		Yes	
Observations		3,973		3,973	
F-value		19.46		25.52	
Adjusted R^2		16.75 %		23.76 %	
Panel B: Compensation mix and forecast bias					
COMPMIX	NA/+			0.01343	2.88***
SKEW, t − 1	NA/−			−0.3156	−1.16
VOLROA, t − 1	NA/?			0.0381	3.83***
LOSS	NA/?			0.0116	7.80***
LEVEARN	NA/?			−0.0564	−7.79***
ESUP	NA/−			−0.0560	−6.33***
NEGESUP	NA/−			0.0028	3.52***
SIZE, t − 1	+/−	0.0242	6.40***	−0.0008	−2.35**
FOLLOW	NA/−			−0.0000	−0.44
DISP	NA/+			0.0000	1.72*
SHROWN	−/−	−0.7714	−7.58***	−0.0104	−1.26
CEOCHAIR	NA/?			0.0010	1.37
NUMMTGS	NA/?			0.0000	0.03
DIROPT	NA/?			0.0001	0.12
GROWTH, t − 1	+/?	3.0284	3.61***	−0.1895	−2.83***

(*continued*)

Table 97.6 (continued)

	Sign	Column 1 Coefficient	Column 1 t-Statistics	Column 2 Coefficient	Column 2 t-Statistics
CHASSETS, t − 1	NA/?			0.0000	0.68
CHSALES, t − 1	NA/?			−0.0000	−1.63
STD	+/NA	1.2514	15.72***		
ROA	+/NA	0.1228	1.96**		
LEV	−/NA	−0.1327	−4.44***		
INTERCEPT	?/?	0.1972	5.40***	0.0075	1.91*
Industry control		Yes		Yes	
Year control		Yes		Yes	
Observations		3,742		3,742	
F-value		19.31		15.95	
Adjusted R^2		17.12 %		17.17 %	

Variable definitions:
COMPMIX = Black-Scholes value of stock options in CEO compensation divided by CEO total compensation (i.e., cash salary + bonus + options)
ACCURACY = (−1) * absolute value of [mean EPS forecast − actual EPS]/price at forecast date
BIAS = the signed forecast error measured as the [mean EPS forecast − actual EPS]/price at forecast date
FOLLOW = number of analysts following the firm
DISP = forecast dispersion, measured as the standard deviation of analysts' forecasts deflated by price at the forecast date
SKEW = difference between the mean and the median forecast scaled by price at the forecast date
VOLROA = earnings volatility measured as the standard deviation of return on assets for the previous 5-year period
ESUP = change in earnings measured as the difference between the current year's EPS and the last year's EPS, divided by price at the beginning of the year
NEGESUP = 0 if *ESUP* is positive and is equal to *ESUP* if *ESUP* is negative
LOSS = a dummy variable which equals 1 when earnings are negative and 0 otherwise
LEVEARN = annual earnings scaled by the year-end market value of equity
GROWTH = beginning book value of equity divided by beginning market value of equity
SIZE = firm size measured as the natural log of beginning assets
SHROWN = number of shares owned by the CEO divided by the total number of shares outstanding
CEOCHAIR = an indicator variable which equals "1" if the CEO is also the chairman of the board and "0" otherwise
NUMMTGS = the number of board meetings held in a year
DIROPT = an indicator variable which equals "1" if the directors are awarded stock options in the year and "0" otherwise
STD = the standard deviation of monthly returns over previous year
ROA = return on assets for year
LEV = financial leverage (liabilities over equity)
***Significant at the 0.01 level, **Significant at the 0.05 level, *Significant at the 0.10 level. Significance levels are based on one-tailed tests when the coefficient sign is predicted and on two-tailed tests otherwise
NA = not applicable

ownership (SHROWN). In the second stage, we use the models described in Eqs. 97.3 and 97.4 with compensation mix (COMPMIX) as the proxy for OPTIONS.

The results reported in Table 97.6 are consistent with the results presented in Tables 97.4 and 97.5. That is, for compensation mix (COMPMIX), we observe a negative association between accuracy and COMPMIX and a positive association between bias and COMPMIX.

97.5 Sensitivity Analyses

We conduct several additional tests to examine the robustness of our findings. First, we control for cross-sectional correlation, since our tests are based on pooled cross-sectional and time series data. Although we include fixed industry and year effects in our primary tests, we conduct Fama-MacBeth estimation with industry controls to account for residual cross-sectional correlation as an additional test. The Fama-MacBeth t-statistics of the coefficient on total options in model 1 and new options and exercisable options in model 2 are statistically significant at the 5 % level or higher for all the tests.

Second, we assess the sensitivity of our results to use of the median analysts' forecast in place of the mean forecast for computing forecast accuracy and bias. We estimate models (3) and (4) using these alternative measures of forecast accuracy and bias as dependent variables and without earnings skewness as an independent variable. We find that the results reported in Tables 97.4 and 97.5 are robust to the choice of median analysts' forecasts in place of mean forecasts.

Finally, since our sample period of years 1993–2003 includes both pre- and post-SOX periods, we delete years 2002 and 2003 and carry out a subsample analysis without the post-SOX periods. For the pre-SOX subsample, consistent with our main results in Tables 97.4 and 97.5, total options in model 1 and new options and exercisable options in model 2 have a significant negative association (at the 5 % level or better) with forecast accuracy and a significant positive association (at the 5 % level or better) with forecast bias.

97.6 Summary and Conclusions

We examine the relation between the level of CEO stock options and the accuracy and bias of analysts' earnings forecasts. We use four different measures of stock options: new options, existing exercisable options, existing unexercisable options, and total options (sum of the previous three). We also use both contemporaneous as well as the lagged values of options compensation in our tests. We hypothesize that forecast accuracy is related to the level of CEO stock options pay. This is because higher levels of stock options may induce managers to undertake riskier projects, to change and/or reallocate their effort, and to possibly engage in gaming (such as opportunistic earnings and disclosure management). However, higher levels of stock options may also better align managers' incentives with those of shareholders and lead to more accurate forecasts. We also hypothesize that forecast bias is related to the level of stock options in CEO pay. The underlying rationale for this is that as the forecast complexity increases with stock options pay, analysts, who need greater access to management's information to produce accurate forecasts, increase the optimistic bias in their forecasts. Alternatively, because higher levels of CEO stock options may better align managers' and shareholders' incentives, they may lead to less biased forecasts.

Our results indicate that analysts' earnings forecast accuracy decreases and forecast optimism increases as the level of stock options in CEO pay increases. Furthermore, our results are robust to the measure of CEO stock options and to the use of current or prior year values of options. These findings suggest that the incentive alignment effects of stock options are more than offset by the investment, effort allocation, and gaming incentives induced by stock options grants to CEOs.

Our study contributes to the current debate on the costs and benefits of the stock options pay to managers. It demonstrates that the level of stock option compensation in CEO pay is an important determinant of analysts' earnings forecast accuracy and bias. Analysts are an important information intermediary in capital markets. The decline in the quality of their forecasts with increased stock options compensation indicates that stock option compensation indirectly affects the quality of the information available to market participants.

Acknowledgments We thank Zhaoyang Gu, Emad Mohd, Mohamed Shehata, K. (Shiva) Sivaramakrishnan, Scott Whisenant, Yoonseok Zang, Jian Zhou, the anonymous reviewers, and workshop participants at the University of Houston, Drexel University, Wilfrid Laurier University, the 2005 American Accounting Association Meeting, the 2005 Canadian Academic Accounting Association Meeting, and the 2005 European Accounting Association Meeting for their helpful comments and Alper Ozesen and Li Wang for their research assistance. We gratefully acknowledge the contribution of Thompson Financial for providing forecast data, available through the Institutional Brokers Estimate System (I/B/E/S). These data have been provided as part of a broad academic program to encourage earnings expectations research. We also thank the Social Sciences and Humanities Research Council of Canada (SSHRC) for their financial support.

References

Aboody, D., & Kaznik, R. (2000). CEO stock option awards and the timing of corporate voluntary disclosures. *Journal of Accounting Economics, 29*(1), 73–100.

Aggarwal, R., & Samwick, A. (1999). The other side of the trade-off: The impact of risk on executive compensation. *Journal of Political Economy, 107*, 65–105.

Atan, R., Jasni, N. R., & Shahwan, Y. (2010). The impact of IFRS 2 "share-based payment" on Malaysian companies. *Review of Pacific Basin Financial Markets and Policies, 13*(3), 449–468.

Bandyopadhyay, S., Brown, L., & Richardson, G. (1995). Analysts' use of earnings forecasts in predicting stock returns: Forecast horizon effects. *International Journal of Forecasting, 11*, 429–445.

Banker, R., & Datar, S. (1989). Sensitivity, precision, and linear aggregation of signals for performance evaluation. *Journal of Accounting Research, 27*(1), 21–39.

Barber, B., Lehavy, R., McNichols, M., & Trueman, B. (2001). Can investors profit from the prophets? Security analyst recommendations and stock returns. *Journal of Finance, 56*, 531–563.

Bartov, E., & Mohanram, P. (2004). Private information, earnings manipulations, and executive stock-option exercises. *The Accounting Review, 79*(4), 889–920.

Bauman, M. P., & Shaw, K. W. (2006). Stock option compensation and the likelihood of meeting analysts' quarterly earnings targets. *Review of Quantitative Finance and Accounting, 26*, 301–319.

Bergstresser, D., & Philippon, T. (2006). CEO incentives and earnings management. *Journal of Financial Economics, 80*, 511–529.

Bizjak, J., Brickley, J., & Coles, J. (1993). Stock-based incentive compensation and investment behavior. *Journal of Accounting and Economics, 16*(1–3), 349–372.

Brown, L. (2001). A temporal analysis of earnings surprises: Profit vs. losses. *Journal of Accounting Research, 39*(2), 221–241.

Bushman, R., & Indjejikian, R. (1993). Accounting income, stock price, and managerial compensation. *Journal of Accounting and Economics, 16*(1–3), 3–23.

Byard, D., Li, Y., & Weintrop, J. (2006). Corporate governance and the quality of financial analysts' information. *Journal of Accounting and Public Policy, 25*, 609–625.

Cao, J., & Laksmana, I. (2010). The effect of capital market on the association between R&D spending and CEO option compensation. *Review of Quantitative Finance and Accounting, 34*, 273–300.

Chen, C. Y. (2002). Additional evidence on the association between director stock ownership and incentive compensation. *Review of Quantitative Finance and Accounting, 19*, 21–44.

Chen, S., & Matsumoto, D. (2006). Favorable versus unfavorable recommendations: The Impact on analyst access to management-provided information. *Journal of Accounting Research, 44*(5), 657–689.

Cheng, Q., & Warfield, T. (2005). Equity incentives and earnings management. *The Accounting Review, 80*(2), 441–476.

Conger, J., Finegold, D., & Lawler, E., III. (1998). Appraising boardroom performance. *Harvard Business Review, 76*, 136–148.

Core, J., Holthausen, R., & Larcker, D. (1999). Corporate governance, chief executive officer compensation, and firm performance. *Journal of Financial Economics, 51*(3), 371–406.

Daily, C. M., & Dalton, D. R. (1997). Separate, but not independent: Board leadership structure in large corporations. *Corporate Governance: An International Review, 5*(3), 126–136.

Das, S., Levine, C., & Sivaramakrishnan, K. (1998). Earnings predictability and bias in analysts' earnings forecast. *The Accounting Review, 73*(2), 277–294.

Duru, A., & Reeb, D. (2002). International diversification and analysts' forecast accuracy and bias. *The Accounting Review, 77*(2), 415–433.

Eames, M., & Glover, S. (2003). Earnings predictability and the direction of analysts' earnings forecast errors. *The Accounting Review, 78*(3), 707–724.

Easton, P., Taylor, G., Shroff, P., & Sougiannis, T. (2002). Using forecasts of earnings to simultaneously estimate growth and the rate of return on equity investment. *Journal of Accounting Research, 40*, 657–676.

Efendi, J., Srivastava, A., & Swanson, E. (2007). Why do corporate managers misstate financial statements? The role of option compensation and other factors. *Journal of Financial Economics, 85*(3), 667–708.

Fama, E., & French, K. (1997). Industry costs of equity. *Journal of Financial Economics, 43*, 153–193.

Fama, E., & MacBeth, J. (1973). Risk, return and equilibrium: Empirical tests. *Journal of Political Economy, 81*, 607–636.

Feltham, G., & Xie, J. (1994). Performance measure congruity and diversity in multi-task principal/agent relations. *The Accounting Review, 69*(3), 429–453.

Frankel, R., & Lee, C. (1998). Accounting valuation, market expectation, and cross- sectional stock returns. *Journal of Accounting & Economics, 25*, 283–319.

Givoly, D., & Lakonishok, J. (1979). The information content of financial analysts' forecasts of earnings: Some evidence on semi-strong efficiency. *Journal of Accounting and Economics, 1*, 165–185.

Gu, Z. (2003). *Controlling for actual earnings: Does it mitigate or create spurious relations in analyst forecast efficiency?* (Working paper). Carnegie Mellon University.

Gu, Z., & Wu, J. S. (2003). Earnings skewness and analyst forecast bias. *Journal of Accounting and Economics, 35*(1), 5–29.

Guay, W. (1999). The sensitivity of CEO wealth to equity risk: An analysis of the magnitude and determinants. *Journal of Financial Economics, 53*, 43–71.

Hall, B. (2003). Six challenges in designing equity-based pay. *Journal of Applied Corporate Finance, 15*, 21–33.

Hall, B., & Murphy, K. (2003). The trouble with stock options. *Journal of Economic Perspectives, 17*(3), 49–70.

Hanlon, M., Rajgopal, S., & Shevlin, T. (2003). Are executive stock options associated with future earnings? *Journal of Accounting and Economics, 36*, 3–43.

Ho, L. C. J., & Tsay, J. (2004). Analysts' forecasts of Taiwanese firms' earnings: Some empirical evidence. *Review of Pacific Basin Financial Markets and Policies, 7*(4), 571–597.

Holmstrom, B., & Milgrom, P. (1991). Multi-task principal–agent analyses: Incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organization, 7*, 24–52.

Hribar, P., & Nichols, C. (2007). The use of unsigned earnings quality measures in tests of earnings management. *Journal of Accounting Research, 45*, 1017–1053.

Imhoff, E., & Lobo, G. (1984). Information content of analysts' composite forecast revisions. *Journal of Accounting Research, 22*(2), 541–554.

Ittner, C., Lambert, R., & Larcker, D. (2003). The structure and performance consequences of equity grants to employees of new economy firms. *Journal of Accounting and Economics, 34*(1–3), 89–127.

Jegadeesh, N., Kim, J., Krische, S., & Lee, C. (2004). Analyzing the analysts: When do recommendations add value? *Journal of Finance, 59*, 1083–1124.

Jensen, M. C. (1993). The modern industrial revolution, exit, and the failure of internal control systems. *The Journal of Finance, 48*(3), 831–880.

Ke, B., & Yu, Y. (2006). The effect of issuing biased earnings forecasts on analysts' access to management and survival. *Journal of Accounting Research, 44*(5), 965–999.

Klassen, K., & Mawani, A. (2000). The impact of financial and tax reporting incentives on option grants to Canadian CEOs. *Contemporary Accounting Research, 17*(2), 227–262.

Kross, W., Ro, B., & Schroeder, D. (1990). Earnings expectations: The analysts' information advantage. *The Accounting Review, 65*(2), 1124–1131.

Lam, S. S., & Chng, B. F. (2006). Do executive stock option grants have value implications for firm performance? *Review of Quantitative Finance and Accounting, 26*, 249–274.

Lambert, R. (2001). Contracting theory and accounting. *Journal of Accounting and Economics, 32*(3), 3–87.

Lambert, R. A,, Lanen, W. N., & Larcker, D. F. (1989). Executive stock option plans and corporate dividend policy. *Journal of Financial and Quantitative Analysis, 24*, 409–425.

Lang, M., & Lundholm, R. (1996). Corporate disclosure policy and analyst behavior. *The Accounting Review, 71*(4), 467–493.

Lim, T. (2001). Rationality and analysts' forecast bias. *Journal of Finance, 56*(1), 369–385.

Loh, R., & Mian, G. (2006). Do accurate earnings forecasts facilitate superior investment recommendations? *Journal of Financial Economics, 80*, 455–483.

McConnell, J., & Servaes, H. (1990). Additional evidence on equity ownership and corporate value. *Journal of Financial Economics, 27*, 595–612.

Murphy, K. (1999). Executive compensation. In O. Ashenfleter & D. Card (Eds.), *Handbook of labor economics*. Amsterdam: North Holland.

Murphy, K. (2003). Stock-based pay in new economy firms. *Journal of Accounting and Economics, 34*(1–3), 129–147.

Philbrick, D., & Ricks, W. (1991). Using value line and IBES analyst forecasts in accounting research. *Journal of Accounting Research, 29*(2), 397–417.

Rajgopal, S., & Shevlin, T. (2002). Empirical evidence on the relation between stock option compensation and risk taking. *Journal of Accounting and Economics, 33*(2), 145–171.

Ross, S. A. (2004). Compensation, incentives, and the duality of risk aversion and riskiness. *Journal of Finance, 59*(1), 207–225.

Smith, C., & Stulz, R. (1985). The determinants of firms' hedging policies. *Journal of Financial and Quantitative Analysis, 20*, 391–405.

Smith, C., & Watts, R. (1992). The investment opportunity set and corporate financing, dividend, and compensation policies. *Journal of Financial Economics, 32*(3), 117–161.

Stickel, S. (1991). Common stock returns surrounding earnings forecast revisions: More puzzling evidence. *The Accounting Review, 66*, 402–416.

Warfield, T., Wild, J., & Wild, K. (1995). Managerial ownership, accounting choices, and informativeness of earnings. *Journal of Accounting and Economics, 20*, 61–91.

Xie, B., Davidson, W., & DaDalt, P. (2003). Earnings management and corporate governance: The role of the board and the audit committee. *Journal of Corporate Finance, 9*, 295–316.

Yermack, D. (1997). Good timing: CEO stock option awards and company news announcements. *Journal of Finance, 52*(2), 449–476.

Option Pricing and Hedging Performance Under Stochastic Volatility and Stochastic Interest Rates

98

Charles Cao, Gurdip S. Bakshi, and Zhiwu Chen

Contents

Abstract

Recent studies have extended the Black–Scholes model to incorporate either stochastic interest rates or stochastic volatility. But, there is not yet any

C. Cao (✉)
Department of Finance, Smeal College of Business, Penn State University, University Park, PA, USA
e-mail: qxc2@psu.edu

G.S. Bakshi (✉)
Department of Finance, College of Business, University of Maryland, College Park, MD, USA
e-mail: gbakshi@rhsmith.umd.edu

Z. Chen (✉)
School of Management, Yale University, New Haven, USA
e-mail: zhiwu.chen@yale.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_98,

comprehensive empirical study demonstrating whether and by how much each generalized feature will improve option pricing and hedging performance.

This chapter fills this gap by first developing an implementable option model in closed form that admits both stochastic volatility and stochastic interest rates and that is parsimonious in the number of parameters. The model includes many known ones as special cases. Based on the model, both delta-neutral and single-instrument minimum-variance hedging strategies are derived analytically. Using S&P 500 option prices, we then compare the pricing and hedging performance of this model with that of three existing ones that respectively allow for (i) constant volatility and constant interest rates (the Black–Scholes), (ii) constant volatility but stochastic interest rates, and (iii) stochastic volatility but constant interest rates. Overall, incorporating stochastic volatility and stochastic interest rates produces the best performance in pricing and hedging, with the remaining pricing and hedging errors no longer systematically related to contract features. The second performer in the horse race is the stochastic volatility model, followed by the stochastic interest rate model and then by the Black–Scholes.

Keywords

Stock option pricing • Stochastic volatility • Stochastic interest rates • Hedge ratios • Hedging • Pricing performance and Hedging performance

98.1 Introduction

Option pricing has, in the last two decades, witnessed an explosion of new models that each relaxes some of the restrictive assumptions underlying the seminal Black and Scholes (1973) model. In doing so, most of the focus has been on the counter-factual constant volatility and constant interest-rate assumptions. For example, Merton's (1973) option pricing model allows interest rates to be stochastic but keeps a constant volatility for the underlying asset, while Amin and Jarrow (1992) develop a similar model where, unlike in Merton's, interest rate risk is also priced. A second class of option models admits stochastic conditional volatility for the underlying asset but maintains the constant interest-rate assumption. These include the Cox and Ross (1976) constant elasticity of variance model and the stochastic volatility models of Bailey and Stulz (1989), Bates (1996b, 2000), Heston (1993), Hull and White (1987a), Scott (1987), Stein and Stein (1991), and Wiggins (1987). Recently, Bakshi and Chen (1997) and Scott (1997) have developed closed-form equity option formulas that admit both stochastic volatility and stochastic interest rates.[1] Their efforts have, in some sense, helped reach the ultimate possibility of

[1] Amin and Ng (1993), Bailey and Stulz (1989), and Heston (1993) also incorporate both stochastic volatility and stochastic interest rates, but their option pricing formulas are not given in closed form, which makes applications difficult. Consequently, comparative statics and hedge ratios are difficult to obtain in their cases.

completely relaxing the Black–Scholes assumptions of constant volatility and constant interest rates. As a practical matter, these sufficiently general pricing formulas should in principle result in significant improvement in pricing and hedging performance over the Black–Scholes model. While option pricing theory has made such impressive progress, the empirical front is nonetheless far behind.[2] Will incorporating these general features improve both pricing and hedging effectiveness? If so, by how much? Can these relaxed assumptions help resolve the well-known empirical biases associated with the Black–Scholes formula, such as the volatility smiles [e.g., Rubinstein (1985, 1994)]? These empirical questions must be answered before the potential of the general models can be fully realized in practical applications.

In this chapter, we first develop a practically implementable version of the general equity option pricing models in Bakshi and Chen (1997) and Scott (1997) that admits stochastic interest rates and stochastic volatility, yet resembles to the extent possible the Black–Scholes model in its implementability. We present procedures for applying the resulting model to price and hedge option-like derivative products. Next, we conduct a complete analysis of the relative empirical performance, in both pricing and hedging, of the four classes of models that respectively allow for (i) constant volatility and constant interest rates (*the BS model*), (ii) constant volatility but stochastic interest rates (*the SI model*), (iii) stochastic volatility but constant interest rates (*the SV model*), and (iv) stochastic volatility and stochastic interest rates (*the SVSI model*). As the SVSI model has all the other three models nested, one should expect its static pricing and dynamic hedging performance to surpass that of the other classes. But, this performance improvement must come at the cost of potentially more complex implementation steps. In this sense, conducting such a horse race study can at least offer a clear picture of possible trade-offs between costs and benefits that each model may present.

Specifically, the SVSI option pricing formula is expressed in terms of the underlying stock price, the stock's volatility, and the short-term interest rate. The spot volatility and the short interest rate are each assumed to follow a Markov mean-reverting square-root process. Consequently, seven structural parameters need to be estimated as input to the model. These parameters can be estimated using the Generalized Method of Moments (GMM) of Hansen (1982), as is done in, for instance, Andersen and Lund (1997), Chan et al. (1992).

[2]There have been a few empirical studies that investigate the pricing, but not the hedging, performance of versions of the stochastic volatility model, relative to the Black–Scholes model. These include Bates (1996b, 2000), Dumas et al. (1998), Madan et al. (1998), Nandi (1996), and Rubinstein (1985). In Bates' work, currency and equity index options data are used to test a stochastic volatility model with Poisson jumps included. Nandi does investigate the pricing and hedging performance of Heston's stochastic volatility model, but he focuses exclusively on a single-instrument minimum-variance hedge that involves only the S&P 500 futures. As will be clear shortly, we address in this chapter both the pricing and the hedging effectiveness issues from different perspectives and for four distinct classes of option models.

Or, they can be backed out from the pricing model itself by using observed option prices, as is similarly done for the BS model both in the existing literature and in Wall Street practice.

In our empirical investigation, we will adopt this implied parameter approach to implement the four models. In this regard, it is important to realize that the BS model is implemented as if the spot volatility and the spot interest rates were assumed to be time varying within the model, that is, the spot volatility is backed out from option prices each day and used, together with the current yield curve, to price the following day's options. The SI and the SV models are implemented with a similarly internally inconsistent treatment, though to a lesser degree. Since this implementation is how one would expect each model to be applied, we chose to follow this convention in order to give the alternatives to the standard BS model the "toughest hurdle." Clearly, such a treatment works in the strongest favor of the BS model and is especially biased against the SVSI model.

Based on 38,749 S&P 500 call (and put) option prices for the sample period from June 1988 to May 1991, our empirical investigation leads to the following conclusions. First, on the basis of two out-of-sample pricing error measures, the SVSI model is found to perform slightly better than the SV model, while they both perform substantially better than the SI (the third-place performer) and the BS model. That is, when volatility is kept constant, allowing interest rates to vary stochastically can produce respectable pricing improvement over the BS model. However, in the presence of stochastic volatility, doing so no longer seems to improve pricing performance much further. Thus, modeling stochastic volatility is far more important than stochastic interest rates, at least for the purpose of pricing options. It is nonetheless encouraging to know that based on our sample, both the SVSI and the SV models typically reduce the BS model's pricing errors by more than half, whereas the SI model helps reduce the BS pricing errors by 20 % or more. While all four models inherit moneyness- and maturity-related pricing biases, the severity of these types of bias is increasingly reduced by the SI, the SV, and the SVSI models. In other words, the SVSI model produces pricing errors that are the least moneyness or maturity related. This conclusion is also confirmed when the Rubinstein (1985) implied-volatility-smile diagnostic is adopted to examine each model.

Two types of hedging strategy are employed in this study to gauge the relative hedging effectiveness. The first type is the conventional delta-neutral hedge, in which as many distinct hedging instruments as the number of risk sources affecting the hedging target's value are used so as to make the net position completely risk immunized (locally). Take the SVSI model as an example. The call option value is driven by three risk sources: the underlying price shocks, volatility shocks, and shocks to interest rates. Accordingly, we employ the underlying stock, a different call option, and a position in a discount bond to create a delta-neutral hedge for a target call option. That closed-form expressions are derived for each hedge ratio is of great value for devising hedging strategies analytically. Similarly, for the SV model, we only need to rely on the underlying stock and an option contract to design a delta-neutral hedge. Based on the delta-neutral hedging errors, the same performance ranking of the four models obtains as that determined by their static pricing performance, except that now the SVSI and the SV models and the SI and the BS models are respectively pairwise virtually indistinguishable. This reenforces the view that adding stochastic interest rates may not affect performance much. However, it is found that the average hedging errors by the SVSI and the SV models are typically less than one-third of the corresponding BS model's hedging errors. Furthermore, reducing the frequency of hedge rebalancing does not tend to reduce the SV and the SVSI models' hedging effectiveness, whereas the BS and the SI models' hedging errors are often doubled when rebalancing frequency changes from daily to once every 5 days. Therefore, after stochastic volatility is controlled for, the frequency of hedge rebalancing will have relatively little impact on hedging effectiveness. This finding is in accord with Galai's (1983a) results that in any hedging scheme, it is probably more important to control for stochastic volatility than for discrete hedging [see Hull and White (1987b) for a similar, simulation-based result for currency options].

To see how the models perform under different hedging schemes, we also look at minimum-variance hedges involving only a position in the underlying asset. As argued by Ross (1995), the need for this type of hedges may arise in contexts where a perfect delta-neutral hedge may not be feasible, either because some of the underlying risks are not traded or even reflected in any traded financial instruments or because model misspecifications and transaction costs render it undesirable to use as many instruments to create a perfect hedge. In the present context, both volatility risk and interest rate risk are, of course, traded and hence can, as indicated above, be controlled for by employing an option and a bond. But, a point can be made that it is sometimes more preferable to adopt a single-instrument minimum-variance hedge. To study this type of hedges, we again calculate the absolute and the dollar-value hedging errors for each model. Results from this exercise indicate that the SV model performs the best among all four, while the BS and the SV models outperform their respective stochastic interest rate counterparts, the SI and the SVSI models. Therefore, under the single-instrument hedges, incorporating stochastic interest rates actually worsens hedging performance. It is also true that hedging errors under this type of hedges are always significantly higher than those under the conventional delta-neutral hedges, for each given moneyness and maturity option category. Thus, whenever possible, including more instruments in a hedge will in general produce better hedging effectiveness.

While our discussion is mainly focused on results obtained using the entire sample period and under specific model implementation designs, robustness of these empirical results is also checked by examining alternative implementation designs, different subperiods, as well as option transaction price data. Especially, given the popularity of the "implied-volatility matrix" method among practitioners, we will also implement each of the four models and compare their pricing and hedging performance, by using only option contracts from a given moneyness–maturity category. It turns out that this alternative implementation scheme does not change the rankings of the four models.

The rest of the chapter proceeds as follows. Section 98.2 develops the SVSI option pricing formula. It discusses issues pertaining to the implementation of the formula and derives the hedge ratios analytically. Section 98.3 provides a description of the S&P 500 option data. In Sect. 98.4 we evaluate the static pricing and the dynamic hedging performance of the four models. Concluding remarks are offered in Sect. 98.5.

98.2 The Option Pricing Model

Consider an economy in which the instantaneous interest rate at time t, denoted $R(t)$, follows a Markov diffusion process:

$$dR(t) = [\theta_R - \overline{\kappa}_R R(t)]dt + \sigma_R \sqrt{R(t)} d\omega_R(t) \quad t \in [0, T], \tag{98.1}$$

where $\overline{\kappa}_R$ regulates the speed at which the interest rate adjusts to its long-run stationary value $\frac{\theta_R}{\overline{\kappa}_R}$ and $\omega_R = \{\omega_R(t) : t \in [0, T]\}$ is a standard Brownian motion.[3] This single-factor interest rate structure of Cox et al. (1985) is adopted as it requires the estimation of only three structural parameters. Adding more factors to the term structure model will of course lead to more plausible formulas for bond prices, but it can make the resulting option formula harder to implement.

Take a generic non-dividend-paying stock whose price dynamics are described by

$$\frac{dS(t)}{S(t)} = \mu(S, t)dt + \sqrt{V(t)} d\omega_S(t) \quad t \in [0, T], \tag{98.2}$$

where $\mu(S, t)$, which is left unspecified, is the instantaneous expected return and ω_S a standard Brownian motion. The instantaneous stock return variance, $V(t)$, is assumed to follow a Markov process:

$$dV(t) = [\theta_v - \overline{\kappa}_v V(t)]dt + \sigma_v \sqrt{V(t)} d\omega_v(t) \quad t \in [0, T], \tag{98.3}$$

where again ω_v is a standard Brownian motion and the structural parameters have the usual interpretation. We refer to $V(t)$ as the spot volatility or, simply, volatility. This process is also frugal in the number of parameters to be estimated and is similar to the one in Heston (1993). Letting ρ denote the correlation coefficient between ω_S and ω_v, the covariance between changes in $S(t)$ and in $V(t)$ is

[3]Here we follow a common practice to assume from the outset a structure for the underlying price and rate processes, rather than derive them from a full-blown general equilibrium. See Bates (1996a), Heston (1993), Melino and Turnbull (1990, 1995), and Scott (1987, 1997). The simple structure assumed in this section can, however, be derived from the general equilibrium model of Bakshi and Chen (1997).

$Cov_t[dS(t), dV(t)] = \rho\sigma_S\sigma_v S(t)V(t)dt$ which can take either sign and is time varying. According to Bakshi et al. (1997, 2000), Bakshi and Chen (1997), Bates (1996a), Cao and Huang (2008), and Rubinstein (1985), this additional feature is important for explaining the skewness and kurtosis-related biases associated with the BS formula. Finally, for ease of presentation, assume that the equity-related shocks and the interest rate shocks are uncorrelated[4]: $Cov_t(d\omega_S, d\omega_R) = Cov_t(d\omega_v, d\omega_R) = 0$.

By a result from Harrison and Kreps (1979), there are no free lunches in the economy if and only if there exists an equivalent martingale measure with which one can value claims as if the economy were risk neutral. For instance, the time t price $B(t, \tau)$ of a zero-coupon bond that pays \$1 in τ periods can be determined via

$$B(t,\tau) = E_Q\left\{\exp\left(-\int_t^{t+\tau} R(S)ds\right)\right\}, \tag{98.4}$$

where E_Q denotes the expectation with respect to an equivalent martingale measure and conditional on the information generated by $R(t)$ and $V(t)$. Assume that the factor risk premiums for $R(t)$ and $V(t)$ are, respectively, given by $\lambda_R R(t)$ and $\lambda_v V(t)$, for two constants λ_R and λ_v. Bakshi and Chen (1997) provide a general equilibrium model in which risk premiums have precisely this form and in which the interest rate and stock price processes are as assumed here. Under this assumption, we obtain the risk-neutralized processes for $R(t)$ and $V(t)$ below:

$$dR(t) = [\theta_R - \kappa_R R(t)]dt + \sigma_R\sqrt{R(t)}d\omega_R(t) \tag{98.5}$$

[4]This assumption on the correlation between stock returns and interest rates is somewhat severe and likely counterfactual. To gauge the potential impact of this assumption on the resulting option model's performance, we initially adopted the following stock price dynamics:

$$\frac{dS(t)}{S(t)} = \mu(S,t)dt + \sqrt{V(t)}d\omega_S(t) + \sigma_{S,R}\sqrt{R(t)}d\omega_R(t) \quad t \in [0,T],$$

with the rest of the stochastic structure remaining the same as given above. Under this more realistic structure, the covariance between stock price changes and interest rate shocks is $Cov_t[dS(t), dR(t)] = \sigma_{S,R}\sigma_R R(t)S(t)dt$, so bond market innovations can be transmitted to the stock market and vice versa. The obtained closed-form option pricing formula under this scenario would have one more parameter $\sigma_{S,R}$ than the one presented shortly, but when we implemented this slightly more general model, we found its pricing and hedging performance to be indistinguishable from that of the SVSI model studied in this chapter. For this reason, we chose to present the more parsimonious SVSI model derived under the stock price process in Eq. 98.2. We could also make both the drift and the diffusion terms of $V(t)$ a linear function of $R(t)$ and $\omega_R(t)$. In such cases, the stock returns, volatility, and interest rates would all be correlated with each other (at least globally), and we could still derive the desired equity option valuation formula. But, that would again make the resulting formula more complex while not improving its performance.

$$dV(t) = [\theta_v - \kappa_v V(t)]dt + \sigma_v \sqrt{V(t)} d\omega_v(t), \tag{98.6}$$

where $\kappa_R \equiv \overline{\kappa}_R + \lambda_R$ and $\kappa_v \equiv \overline{\kappa}_v + \lambda_v$. The risk-neutralized stock price process becomes

$$\frac{dS(t)}{S(t)} = R(t)dt + \sqrt{V(t)} d\omega_S(t), \tag{98.7}$$

that is, under the martingale measure, the stock should earn no more and no less than the risk-free rate. With these adjustments, we solve the conditional expectation in Eq. 98.4 and obtain the familiar bond price equation below:

$$B(t,\tau) = \exp[-\varphi(\tau) - \varrho(\tau)R(t)], \tag{98.8}$$

where $\varphi(\tau) = \frac{\theta_R}{\sigma_R^2}\left\{(\varsigma - \kappa_R)\tau + 2\ln\left[1 - \frac{(1-e^{-\varsigma\tau})(\varsigma - \kappa_R)}{2\varsigma}\right]\right\}, \varrho(\tau) = \frac{2(1-e^{-\varsigma\tau})}{2\varsigma - [\varsigma - \kappa_R](1-e^{-\varsigma\tau})}$, and $\varsigma \equiv \sqrt{\kappa_R^2 + 2\sigma_R^2}$. See Cox et al. (1985) for an analysis of this class of term structure models.

98.2.1 Pricing Formula for European Options

Now, consider a European call option written on the stock, with strike price K and term to expiration τ. Let its time t price be denoted by $C(t, \tau)$. As (S, R, V) form a joint Markov process, the price $C(t, \tau)$ must be a function of $S(t)$, $R(t)$, and $V(t)$ (in addition to τ). By a standard argument, the option price must solve

$$\begin{aligned}&\frac{1}{2}VS^2\frac{\partial^2 C}{\partial S^2} + RS\frac{\partial C}{\partial S} + \rho\sigma_v VS\frac{\partial^2 C}{\partial S \partial V} + \frac{1}{2}\sigma_v^2 V\frac{\partial^2 C}{\partial V^2} + [\theta_v - \kappa_v V]\frac{\partial C}{\partial V}\\&+\frac{1}{2}\sigma_R^2 R\frac{\partial^2 C}{\partial R^2}[\theta_R - \kappa_R R]\frac{\partial C}{\partial R} - \frac{\partial C}{\partial \tau} - RC = 0,\end{aligned} \tag{98.9}$$

subject to $C(t + \tau, 0) = \max\{S(t + \tau) - K, 0\}$. In the Appendix 1 it is shown that

$$C(t,\tau) = S(t)\Pi_1(t,\tau;S,R,V) - KB(t,\tau)\Pi_2(t,\tau,S,R,V), \tag{98.10}$$

where the risk-neutral probabilities, $\prod_1$ and $\prod_2$, are recovered from inverting the respective characteristic functions [see Heston (1993) and Scott (1997) for similar treatments]:

$$\prod\nolimits_j(t,\tau;S(t),R(t),V(t)) = \frac{1}{2} + \frac{1}{\pi}\int_0^\infty Re\left[\frac{e^{-i\phi\,\ln[K]}f_j(t,\tau,S(t),R(t),V(t);\phi)}{i\phi}\right]d\phi, \tag{98.11}$$

for $j = 1, 2$. The characteristic functions f_j are given by

$$\begin{aligned} f_1(t,\tau) \quad &= \exp\left\{ -\frac{\theta_R}{\sigma_R^2}\left[2\ln\left(1 - \frac{[\xi_R - \kappa_R]\left(1 - e^{-\xi_R \tau}\right)}{2\xi_R} \right) + [\xi_R - \kappa_R]\tau \right] \right. \\ &- \frac{\theta_v}{\sigma_v^2}\left[2\ln\left(1 - \frac{[\xi_v - \kappa_v + (1 + i\phi)\rho\sigma_v]\left(1 - e^{-\xi_v \tau}\right)}{2\xi_v} \right) \right] \\ &- \frac{\theta_v}{\sigma_v^2}[\xi_v - \kappa_v + (1 + i\phi)\rho\sigma_v]\tau + i\phi \ln[S(t)] + \frac{2i\phi\left(1 - e^{-\xi_R \tau}\right)}{2\xi_R - [\xi_R - \kappa_R](1 - e^{-\xi_R \tau})} R(t) \\ &\left. + \frac{i\phi(i\phi + 1)\left(1 - e^{-\xi_v \tau}\right)}{2\xi_v - [\xi_v - \kappa_v + (1 + i\phi)\rho\sigma_v](1 - e^{-\xi_v \tau})} V(t) \right\}, \end{aligned} \tag{98.12}$$

and

$$\begin{aligned} f_2(t,\tau) \quad &= \exp\left\{ -\frac{\theta_R}{\sigma_R^2}\left[2\ln\left(1 - \frac{\left[\xi_R^* - \kappa_R\right]\left(1 - e^{-\xi_R^* \tau}\right)}{2\xi_R^*} \right) + \left[\xi_R^* - \kappa_R\right]\tau \right] \right. \\ &- \frac{\theta_v}{\sigma_v^2}\left[2\ln\left(1 - \frac{\left[\xi_v^* + i\phi\rho\sigma_v\right]\left(1 - e^{-\xi_v^* \tau}\right)}{2\xi_v^*} \right) + \left[\xi_v^* - \kappa_v + i\phi\rho\sigma_v\right]\tau \right] \\ &+ i\phi \ln[S(t)] - \ln[B(t,\tau)] + \frac{2(i\phi - 1)\left(1 - e^{-\xi_R^* \tau}\right)}{2\xi_R^* - \left[\xi_R^* - \kappa_R\right](1 - e^{-\xi_R^* \tau}} R(t) \\ &\left. + \frac{i\phi(i\phi - 1)\left(1 - e^{-\xi_v^* \tau}\right)}{2\xi_v^* - \left[\xi_v^* - \kappa_v + i\phi\rho\sigma_v\right]\left(1 - e^{-\xi_v^* \tau}\right)} V(t) \right\}, \end{aligned} \tag{98.13}$$

where $\xi_R = \sqrt{\kappa_R^2 - 2\sigma_R^2 i\phi}$, $\xi_v = \sqrt{[\kappa_v - (1 + i\phi)\rho\sigma_v]^2 - i\phi(i\phi + 1)\sigma_v^2}$, $\xi_R^* = \sqrt{\kappa_R^2 - 2\sigma_R^2(i\phi - 1)}$, and $\xi_v^* = \sqrt{[\kappa_v - i\phi\rho\sigma_v]^2 - i\phi(i\phi - 1)\sigma_v^2}$. The price of a European put on the same stock can be determined from the put–call parity.

The option valuation model in Eq. 98.10 has several distinctive features. First, it applies to cases with stochastically varying interest rates and volatility. It contains as special cases most existing models, such as the SV models, the SI models, and clearly the BS model. Second, as mentioned earlier, it allows for a flexible correlation structure between the stock return and its volatility, as opposed to the perfect correlation assumed in, for instance, Heston's (1993) model. Furthermore, the volatility risk premium is time varying and state dependent. This is a departure from Hull and White (1987), Scott (1987), Stein and Stein (1991), and Wiggins (1987) where the volatility risk premium is either a constant or zero. Third, when compared to the general models in Bakshi and Chen (1997) and Scott (1997), the formula in Eq. 98.10 is parsimonious in the number of parameters; especially, it is given only as a function of identifiable variables such that all parameters can be estimated based on available financial market data.

The pricing formula in Eq. 98.10 applies to European equity options. But, in reality most of the traded option contracts are American in nature. While it is beyond the scope of the chapter to derive a model for American options, it is nevertheless possible to capture the first-order effect of early exercise in the following manner. For options with early-exercise potential, compute the Barone-Adesi and Whaley (1987) or Kim (1990) early-exercise premium, treating it as if the stock volatility and the yield curve were time invariant. Adding this early-exercise adjustment component to the European option price in Eq. 98.10 should deliver a reasonable approximation of the corresponding American option price [e.g., Bates (1996b)].

98.2.2 Hedging and Hedge Ratios

One appealing feature of a closed-form option pricing formula, such as the one in Eq. 98.10, is the possibility of deriving comparative statics and hedge ratios analytically. In the present context, there are three sources of stochastic variations over time, price risk $S(t)$, volatility risk $V(t)$, and interest rate risk $R(t)$. Consequently, there are three deltas:

$$\Delta_S(t,\tau;K) \equiv \frac{\partial C(t,\tau)}{\partial S} = \Pi_1 > 0 \tag{98.14}$$

$$\Delta_V(t,\tau;K) \equiv \frac{\partial C(t,\tau)}{\partial V} = S(t)\frac{\partial \Pi_1}{\partial V} - KB(t,\tau)\frac{\partial \Pi_2}{\partial V} > 0 \tag{98.15}$$

$$\Delta_R(t,\tau;K) \equiv \frac{\partial C(t,\tau)}{\partial R} = S(t)\frac{\partial \Pi_1}{\partial V} - KB(t,\tau)\left\{\frac{\partial \Pi_2}{\partial V} - \varrho(\tau)\Pi_2\right\} > 0, \tag{98.16}$$

where, for $g = V, R$ and $j = 1, 2$,

$$\frac{\partial \Pi_j}{\partial g} = \frac{1}{\pi}\int_0^\infty Re\left[(i\phi)^{-1}e^{-i\ \phi \ln[K]}\ \frac{\partial f_j}{\partial g}\right]d\phi. \tag{98.17}$$

The second-order partial derivatives with respect to these variables are provided in the Appendix.

As $V(t)$ and $R(t)$ are both stochastic in our model, these deltas will in general differ from their Black–Scholes counterpart. To see how they may differ, let us resort to an example in which we set $R\,(t) = 6.27\ \%$, $S\,(t) = 279$, $\sqrt{V(t)} = 22.12\ \ \%$, $K_R = 0.481$, $\theta_R = 0.037$, $\sigma_R = 0.049$, $K_v = 1.072$, $\theta_v = 0.041$, $\sigma_v = 0.284$, and $\rho = -0.60$. These values are backed out from the S&P 500 option prices as of July 5, 1988. Fix $K = \$270$ and $\tau = 45$ days. Let Δ_S be as given in Eq. 98.14 for the SVSI model and Δ_S^{bs} its BS counterpart, with Δ_S^{bs} calculated using the same implied volatility.

In unreported results, we plot the difference between Δ_S and Δ_S^{bs}, across different spot price levels and different correlation values. The correlation coefficient ρ is chosen to be the focus as it is known to play a crucial role in determining the skewness of the stock return distribution. When ρ is at -0.50 and -1.0, the difference between the deltas is W shaped, and it reaches the highest value when the option is at the money. The reverse is true when ρ is positive. Thus, Δ_S is generally different from Δ_S^{bs}. Analogous difference patterns emerge when the other option deltas are compared with their respective BS counterpart.

We also plot $Delta_V$ and $Delta_R$ and observe the following. (i) The volatility hedge ratio Δ_V from the SVSI model is, at each spot price, lower than its BS counterpart (except for deep in-the-money options when $\rho < 0$ and for deep out-of-the-money options when $\rho > 0$).[5] (ii) The interest rate delta, Δ_R, and its BS counterpart, Δ_S^{bs}, are almost not different from each other for slightly out-of-the-money options but can be dramatically different for at-the-money options as well as for sufficiently deep in-the-money or deep out-of-the-money calls. For example, pick $\rho = -1.0$. When $S = \$315$, we have $\Delta_R = 30.94$ and $\Delta_R^{bs} = 32.35$; when $S = \$226$, we have $\Delta_R = 0.003$ and $\Delta_R^{bs} = 0.430$. (iii) As expected, out-of-the-money options are overall less sensitive to changes in the spot interest rate, regardless of the model used. In summary, if a portfolio manager/trader relies, in an environment with stochastic interest rates and stochastic volatility, on the BS model to design a hedge for option positions, the manager/trader will likely fail.

Analytical expressions for the deltas are useful for constructing hedges based on an option formula. Below, we present two types of hedges by using the SVSI model as an example.

98.2.2.1 Delta-Neutral Hedges

To demonstrate how the deltas may be used to construct a delta-neutral hedge, consider an example in which a financial institution intends to hedge a short position in a call option with τ periods to expiration and strike price K. In the stochastic interest rate–stochastic volatility environment, a perfectly delta-neutral hedge can be achieved by taking a long position in the replicating portfolio of the call. As three traded assets are needed to control the three sources of uncertainty, the replicating portfolio will involve a position in (i) some $X_S(t)$ shares of the underlying stock (to control for the $S(t)$ risk), (ii) some $X_B(t)$ units of a τ-period discount bond (to control for the $R(t)$ risk), and (iii) some $X_C(t)$ units of another call option with strike price $\overline{K}$ (or any option on the stock with a different maturity) in order to control for the volatility risk $V(t)$. Denote the time t price of the replicating portfolio by $G(t) : G(t) = X_0(t) + X_S(t)S(t) + X_B(t)B(t,\tau) + X_C(t)C\left(t,\tau;\overline{K}\right)$, where $X_0(t)$ denotes the amount put into the instantaneously maturing risk-free bond and it serves as a residual "cash position." Deriving the dynamics for $G(t)$ and comparing them with those of $C(t, \tau; K)$, we find the following solution for the delta-neutral hedge:

$$X_C(t) = \frac{\Delta_V(t,\tau;K)}{\Delta_V\left(t,\tau;\overline{K}\right)} \tag{98.18}$$

$$X_S(t) = \Delta_S(t,\tau;K) - \Delta_S\left(t,\tau;\overline{K}\right)X_C(t) \tag{98.19}$$

$$X_B(t) = \frac{1}{B(t,\tau)\varrho(\tau)}\left\{\Delta_R\left(t,\tau;\overline{K}\right)X_C(t) - \Delta_R(t,\tau;K)\right\} \tag{98.20}$$

and the residual amount put into the instantaneously maturing bond is

$$X_0(t) = C(t,\tau;K) - X_S(t)S(t) - X_C(t)C\left(t,\tau;\overline{K}\right) - X_B(t)B(t,\tau), \tag{98.21}$$

where all the primitive deltas, Δ_S, Δ_R, and Δ_V, are as determined in Eqs. 98.14, 98.14, and 98.16. Like the option prices, these hedge ratios all depend on the values taken by $S(t)$, $V(t)$, and $R(t)$ and those by the structural parameters. Such a hedge created using the general option pricing model should in principle perform better than using the BS model. In the latter case, only the underlying price uncertainty is controlled for, but not the uncertainties associated with volatility and interest rate fluctuations.

In theory this delta-neutral hedge requires continuous rebalancing to reflect the changing market conditions. In practice, of course, only discrete rebalancing is possible. To derive a hedging effectiveness measure, suppose that portfolio rebalancing takes place at intervals of length Δt. Then, precisely as described above, at time t the short call option goes long in (i) $X_S(t)$ shares of the underlying asset, (ii) $X_B(t)$ units of the τ-period bond, and (iii) $X_C(t)$ contracts of a call option with the same term to expiration but a different strike price $\overline{K}$, and invests the residual, X_0, in an instantaneously maturing risk-free bond. After the next interval, compute the hedging error according to

$$\begin{aligned} H(t+\Delta t) = {} & X_0 e^{R(t)\Delta t} + X_S(t)S(t+\Delta t) + X_B(t)B(t+\Delta t,\tau-\Delta t) \\ & + X_C(t)C\left(t+\Delta t,\tau-\Delta t;\overline{K}\right) - C(t+\Delta t,\tau-\Delta t;K). \end{aligned} \tag{98.22}$$

Then, at time $t + \Delta t$, reconstruct the self-financed portfolio, repeat the hedging error calculation at time $t + 2\Delta t$, and so on. Record the hedging errors $H(t + j\Delta t)$, for $j = 1, \cdots, J \equiv \frac{\tau - t}{\Delta t}$. Finally, compute the average absolute hedging error as a function of rebalancing frequency $\Delta t : H(\Delta t) = \frac{1}{J}\sum_{j=1}^{J}|H(t+j\Delta t)|$ and the average dollar-value hedging error: $\overline{H}(\Delta t) = \frac{1}{J}\sum_{j=1}^{J}H(t+j\Delta t)$.

In comparison, if one relies on the BS model to construct a delta-neutral hedge, the hedging error measures can be similarly defined as in Eq. 98.22, except that

$X_B(t)$ and $X_C(t)$ must be restricted to zero and $X_S(t)$ must be the BS delta. Likewise, if the SI model is applied, the only change is to set $X_C(t)$ to zero with Δ_S and Δ_R determined by the SI model; in the case of the SV model, set $X_B(t) = 0$ and let Δ_S and Δ_V be as determined in the SV model. The Appendix provides in closed form an SI option pricing formula and an SV option formula.

98.2.2.2 Single-Instrument Minimum-Variance Hedges

As discussed before, consideration of such factors as model misspecification and transaction costs may render it more practical to use only the underlying asset of the target option as the hedging instrument. Under this single-instrument constraint, a standard design is to choose a position in the underlying stock so as to minimize the variance of instantaneous changes in the value of the hedge. Letting $X_S(t)$ again be the number of shares of the stock to be purchased, solving the standard minimum-variance hedging problem under the SVSI model gives

$$X_S(t) = \frac{Cov_t[ds(t), dC(t,\tau)]}{Var[dS(t)]} = \Delta_S + \rho\sigma_v \frac{\Delta_V(t,\tau)}{S(t)}, \tag{98.23}$$

and the resulting residual cash position for the replicating portfolio is

$$X_0(t) = C(t,\tau) - X_S(t)S(t). \tag{98.24}$$

This minimum-variance hedge solution is quite intuitive, as it says that if stock volatility is deterministic (i.e., $\sigma_v = 0$) or if stock returns are not correlated with volatility changes (i.e., $\rho = 0$), one only needs to long $\Delta_S(t)$ shares of the stock and no other adjustment is necessary. However, if volatility is stochastic and correlated with stock returns, the position to be taken in the stock must control not only for the direct impact of underlying stock price changes on the target option value but also for the indirect impact of that part of volatility changes which is correlated with stock price fluctuations. This effect is reflected in the last term in Eq. 98.23, which shows that the additional number of shares needed besides Δ_S is increasing in ρ (assuming $\sigma_v > 0$).

As for the previous case, suppose that the target call is shorted and that $X_S(t)$ shares are bought and $X_0(t)$ dollars are put into the instantaneous risk-free bond, at time t. The combined position is a self-financed portfolio. At time $t + \Delta t$, the hedging error of this minimum-variance hedge is calculated as

$$H(t+\Delta t) = X_S(t)S(t+\Delta t) + X_0(t)e^{R(t)\Delta t} - C(t+\Delta t, \tau - \Delta t). \tag{98.25}$$

Unlike in Nandi (1996) where he uses the remaining variance of the hedge as a hedging effectiveness gauge, we compute, based on the entire sample period, the average absolute and the average dollar hedging errors to measure the effectiveness of the hedge.

Minimum-variance hedging errors under the SV model as well as under the SI model can be similarly determined accounting for their modeling differences. In the case of the SV model, there is still an adjustment term for the single stock position as in Eq. 98.23. But, for the SI model, the corresponding $X_S(t)$ is the same as its Δ_S. For the BS model, this single-instrument minimum-variance hedge is the same as the delta-neutral hedge. Both types of hedging strategy will be examined under each of the four alternative models.

98.2.3 Implementation

In addition to the strike price and the term to expiration (which are specified in the contract), the SVSI pricing formula in Eq. 98.10 requires the following values as input:

- The spot stock price. If the stock pays dividends, the stock price must be adjusted by the present value of future dividends.
- The spot volatility.
- The spot interest rate.
- The matching τ-period yield to maturity (or the bond price).
- The seven structural parameters: $\{\kappa_R, \theta_R, \sigma_R, \kappa_v, \theta_v, \sigma_v, \rho\}$

For computing the price of a European option, we offer two alternative two-step procedures below. One can implement these steps on any personal computer:

Procedure A

Step 1. Obtain a time series each for the short rate, the stock return, and the stock volatility. Jointly estimate the structural parameters, $\{\kappa_R, \theta_R, \sigma_R, \kappa_v, \theta_v, \sigma_v, \rho\}$, using Hansen's (1982) GMM.

Step 2. Determine the risk-neutral probabilities, Π_1 and Π_2, from the characteristic functions in Eqs. 98.12 and 98.13. Substitute (i) the two probabilities, (ii) the stock price, and (iii) the yield to maturity into (98.10) to compute the option price.

While offering an econometrically rigorous method to estimate the structural parameters, Step 1 in Procedure A may not be as practical or convenient, because of its requirement on historical data. A further difficulty with this approach is its dependence on the measurement of stock volatility. In implementing the BS model, practitioners predominantly use the implied volatility from the model itself rather than relying on historical data. This practice has not only reduced data requirement dramatically but also resulted in significant performance improvement [e.g., Bates (2000) and Melino and Turnbull (1990, 1995)]. Clearly, one can also follow this practice to implement the SVSI model.

Procedure B

Step 1. Collect N option prices on the same stock and taken from the same point in time (or same day), for any $N \geq 8$. Let $\hat{C}_n(t, \tau_n, K_n)$ be the observed price and $C_n(t, \tau_n, K_n)$ the model price as determined by Eq. 98.10 with $S(t)$ and $R(t)$ taken from the market, for the n-th option with τ_n periods to expiration and strike price K_n and for each $n = 1, \ldots, N$. Clearly, the difference between $\hat{C}_n$ and C_n is a function of the values taken by $V(t)$ and by $\Phi \equiv \{\kappa_R, \theta_R, \sigma_R, \kappa_v, \theta_v, \sigma_v, \rho\}$. Define

$$\epsilon_n[V(t), \Phi] \equiv \hat{C}_n(t, \tau_n, K_n) - C_n(t, \tau_n, K_n), \tag{98.26}$$

for each n. Then, find $V(t)$ and parameter vector Φ (a total of eight), so as to minimize the sum of squared errors:

$$\sum_{n=1}^{N} \left| \epsilon_n[V(t), \Phi]^2 \right|. \tag{98.27}$$

The result from this step is an estimate of the implied spot variance and seven structural parameter values, for date t. See Bates (1996b, 2000), Dumas et al. (1998), Longstaff (1995), Madan et al. (1998), and Nandi (1996) where they adopt this technique for similar purposes.

Step 2. Based on the estimate from the first step, follow Step 2 of Procedure A to compute date-(t + 1)'s option prices on the same stock.

In the existing literature, the performance of a new option pricing model is often judged relative to that of the BS model when the latter is implemented using the model's own implied volatility and the time-varying interest rates. Since volatility and interest rates in the BS are assumed to be constant over time, this internally inconsistent practice will clearly and significantly bias the application results in favor of the BS model. But, as this is the current standard in judging performance, we will follow Procedure B to implement the SVSI model and similar procedures to implement the BS, the SV, and the SI models. Then, the models will be ranked relative to each other according to their performance so determined.

98.3 Data Description

For all the tests to follow, we use, based on the following considerations, S&P 500 call option prices as the basis. First, options written on this index are the most actively traded European-style contracts. Recall that like the BS model, formula (98.10) applies to European options. Second, the daily dividend distributions are available for the index (from the S&P 500 Information Bulletin). Harvey and Whaley (1992a, b), for instance, emphasize that critical pricing errors can result when dividends are omitted from empirical tests of any option valuation model. Furthermore, S&P 500 options and options on S&P 500 futures have been the focus of many existing empirical investigations including Bates (2000), Dumas et al. (1998), Madan et al. (1998), Nandi (1996), and Rubinstein (1994). Finally, we also used S&P 500 put option prices to estimate the pricing and hedging errors of all four models and found the results to be similar, both qualitatively and quantitatively, to those reported in the chapter. To save space, we chose to focus on the results based on the call option prices.

The sample period extends from June 1, 1988 through May 31, 1991. The intradaily transaction prices and bid–ask quotes for S&P 500 options are obtained from the Berkeley Option Database. Note that the recorded S&P 500 index values are not the daily closing index levels. Rather, they were the corresponding index levels at the moment when the recorded option transaction took place or when an option price quote was recorded. Thus, there is no nonsynchronous price issue here, except that the S&P 500 index level itself may contain stale component stock prices at each point in time.

The data on the daily Treasury bill bid and ask discounts with maturities up to 1 year are hand collected from the *Wall Street Journal* and provided to us by Hyuk Choe and Steve Freund. By convention, the average of the bid and ask Treasury bill discounts is used and converted to an annualized interest rate. Careful attention is given to this construction since Treasury bills mature on Thursdays, while index options expire on the third Friday of the month. In such cases, we utilize the two Treasury bill rates straddling the option's expiration date to obtain the interest rate of that maturity, which is done for each contract and each day in the sample. The Treasury bill rate with 30 days to maturity is the surrogate used for the short rate in Eq. 98.1 (and in the determination of the probabilities in Eq. 98.10).

For European options, the spot stock price must be adjusted for discrete dividends. For each option contract with τ periods to expiration from time t, we first obtain the present value of the daily dividends $D(t)$ by computing

$$\overline{D}(t,\tau) = \sum_{S=1}^{\tau-t} e^{-R(t,s)s} D(t+s), \tag{98.28}$$

where $R(t, s)$ is the s-period yield to maturity. This procedure is repeated for all option maturities and for each day in our sample. In the next step, we subtract the present value of future dividends from the time t index level, in order to obtain the dividend-exclusive S&P 500 spot index series that is later used as input into the option models.

Several exclusion filters are applied to construct the option price data set. First, option prices that are time stamped later than 3:00 p.m. Central Daytime are eliminated. This ensures that the spot price is recorded synchronously with its option counterpart. Second, as options with less than 6 days to expiration may induce liquidity-related biases, they are excluded from the sample. Third, to mitigate the impact of price discreteness on option valuation, option prices lower than $\$\frac{3}{8}$ are not included. Finally, quote prices that are less than the intrinsic value of the option are taken out of the sample.

We divide the option data into several categories according to either moneyness or term to expiration. A call option is said to be at the money (ATM) if its $\frac{S}{K} \in (0.97, 1.03)$, where S is the spot price and K the strike; out-of-the-money (OTM) if $\frac{S}{K} \leq 0.97$; and in-the-money (ITM) if $\frac{S}{K} \geq 1.03$. A finer partition resulted in nine moneyness categories. By the term to expiration, each option can be classified as [e.g., Rubinstein (1985)] (i) extremely short term (<30 days), (ii) short term

Table 98.1 Sample properties of S&P 500 index options

Moneyness	Term to expiration (days)						
S/K	<30	30–60	60–90	90–120	120–180	≥180	Subtotal
<0.93	0.78	1.33	1.99	2.84	4.88	7.82	
	{23}	{246}	{266}	{431}	{1,080}	{1,538}	{3,584}
0.93–0.95	1.02	1.91	3.30	5.08	8.14	12.86	
	{121}	{595}	{267}	{319}	{596}	{646}	{2,544}
0.95–0.97	1.35	3.05	5.35	7.45	10.87	15.91	
	{488}	{1,012}	{316}	{351}	{670}	{628}	{3,465}
0.97–0.99	2.47	5.53	8.23	10.83	14.19	19.33	
	{838}	{1,020}	{312}	{336}	{676}	{706}	{3,888}
0.99–1.01	5.27	8.99	11.96	14.55	17.95	23.20	
	{776}	{954}	{285}	{308}	{629}	{631}	{3,583}
1.01–1.03	9.65	13.17	15.99	18.84	22.06	27.74	
	{752}	{906}	{276}	{283}	{607}	{597}	{3,421}
1.03–1.05	14.79	17.80	20.80	23.36	26.39	31.91	
	{675}	{844}	{241}	{264}	{542}	{501}	{3,067}
1.05–1.07	20.20	22.63	25.83	27.83	30.69	35.70	
	{620}	{760}	{224}	{242}	{449}	{473}	{2,818}
≥1.07	41.23	42.28	47.50	49.27	51.34	59.82	
	{2,143}	{2,350}	{1,284}	{1,355}	{2,184}	{3,063}	{12,379}
Subtotal	{6,436}	{8,687}	{3,471}	{3,889}	{7,483}	{8,783}	{38,749}

The reported numbers are respectively the average quoted bid–ask midpoint price and the number of observations. Each option contract is consolidated across moneyness and term-to-expiration categories. The sample period extends from June 1, 1988 through May 31, 1991 for a total of 38,749 calls. Daily information from the last quote of each option contract is used to obtain the summary statistics. *S* denotes the spot S&P 500 index level and K is the exercise price

(30–60 days), (iii) near term (60–120 days), (iv) middle maturity (120–180 days), and (v) long term (>180 days). The proposed moneyness and term-to-expiration classifications resulted in 54 categories for which the empirical results will be reported.

Table 98.1 describes sample properties of the S&P 500 call option prices used in the tests. Summary statistics are reported for the average bid–ask midpoint price and the total number of observations, for each moneyness–maturity category. Note that there is a total of 38,749 call price observations, with deep in-the-money and at-the-money options, respectively, taking up 32 % and 28 % of the total sample and that the average call price ranges from \$0.78 for extremely short term, deep out-of-the-money options to \$59.82 for long-term, deep in-the-money options.

98.4 Empirical Tests

This section examines the relative empirical performance of the four models. The analysis is intended to present a complete picture of what each generalization of the benchmark BS model can really buy in terms of performance improvement and whether each generalization produces a worthy tradeoff between benefits and costs. We will pursue this analysis by using three yardsticks: (i) the size of the out-of-sample cross-sectional pricing errors (*static performance*); (ii) the size of model-based hedging errors (*dynamic performance*); and (iii) the existence of systematic biases across strike prices or across maturities (i.e., *does the implied volatility still smile*?).

Based on Procedure B of Sect. 98.2.3, Table 98.2 reports the summary statistics for the daily estimated structural parameters and the implied spot standard deviation, respectively, for the SVSI, the SV, the SI, and the BS models. Take the SVSI model as an example. Over the entire sample period 06:1988–05:1991, $\kappa_v = 0.906$, $\theta_v = 0.042$, and $\sigma_v = 0.414$. These estimates imply a long-run mean of 21.53 % for the volatility process. The implicit (average) half-life for variance mean reversion is 9.18 months. These estimates are similar in magnitude to those reported in Bates (1996b, 2000) for S&P 500 futures options. The estimated parameters for the (risk neutralized) short-rate process are also reasonable and comparable to those in Chan et al. (1992). The presented correlation estimate for ρ is -0.763. The average implied standard deviation is 19.27 %. As seen from the reported standard errors in Table 98.2, for each given model the daily parameter and spot volatility estimates are quite stable from subperiod to subperiod. Histogram-based inferences (not reported) indicate that the majority of the estimated values are centered around the mean.

In estimating the structural parameters and the implied volatility for a given day, we used all S&P 500 options collected in the sample for that day (regardless of maturity and moneyness). This is the treatment applied to the SI, the SV, and the SVSI models. For the BS model, however, Whaley (1982) makes the point that ATM options may give an implied-volatility estimate which produces the best pricing and hedging results. Based on his justification, we used, for each given day, one ATM option that had at least 15 days to expiration to back out the BS model's implied-volatility value. This estimate was then used to determine the next day's pricing and hedging errors of the BS model. See Bates (1996a) for a review of alternative approaches to estimating the BS model's implied volatility.

Observe in Table 98.2 that for the overall sample period, the average implied standard deviation is 19.27 % by the SVSI model, 19.02 % by the SV, 18.14 % by the SI, and 18.47 % by the BS model, where the difference between the highest and the lowest is only 1.13 %. For each subperiod the implied-volatility estimates are similarly close across the four models. This is somewhat surprising. It should, however, be recognized that this comparison is based only on the average estimates over a given period. When we examined the day-to-day time-series paths of the four models' implied-volatility estimates, we found the difference between the two models' implied standard deviations to be sometimes as high as 6 %. Economically, option prices and hedge ratios are generally quite sensitive to the volatility input (see Figlewski 1989). Even small differences in the implied-volatility estimate can lead to significantly different pricing and hedging results.

Table 98.2 Estimates of the structural parameters for stochastic interest rates (SI), stochastic volatility (SV), and stochastic volatility and stochastic interest rates (SVSI) models

	06:1988–05:1991			06:1988–05:1989			06:1989–05:1990			06:1990–05:1991		
Parameters	SI	SV	SVSI	SI	SV	SVSI	SI	SV	SVSI	SI	SV	SVSI
κ_R	2.35		0.61	2.51		0.49	2.29		0.61	2.03		0.76
	(0.03)		(0.02)	(0.06)		(0.02)	(0.05)		(0.03)	(0.05)		(0.04)
θ_R	0.35		0.02	0.33		0.02	0.34		0.02	0.35		0.02
	(0.00)		(0.00)	(0.01)		(0.00)	(0.01)		(0.00)	(0.01)		(0.00)
σ_R	0.04		0.03	0.04		0.04	0.04		0.03	0.04		0.03
	(0.00)		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.01)		(0.00)
κ_v		1.10	0.91		1.29	1.16		1.05	0.78		0.94	0.76
		(0.02)	(0.03)		(0.05)	(0.07)		(0.03)	(0.05)		(0.04)	(0.04)
θ_v		0.04	0.04		0.04	0.04		0.04	0.04		0.04	0.04
		(0.00)	(0.01)		(0.00)	(0.00)		(0.00)	(0.00)		(0.00)	(0.00)
σ_v		0.38	0.41		0.30	0.35		0.42	0.44		0.43	0.46
		(0.00)	(0.00)		(0.01)	(0.01)		(0.01)	(0.01)		(0.00)	(0.00)
ρ		−0.64	−0.76		−0.54	−0.73		−0.62	−0.74		−0.76	−0.82
		(0.01)	(0.01)		(0.01)	(0.01)		(0.01)	(0.01)		(0.01)	(0.01)
$\sqrt{V(t)}$ (%)	18.14	19.02	19.27	17.22	18.41	18.50	17.43	17.61	18.10	20.14	21.19	21.47
	(0.13)	(0.15)	(0.15)	(0.15)	(0.17)	(0.18)	(0.13)	(0.18)	(0.21)	(0.25)	(0.35)	(0.35)

Each day in the sample, the structural parameters of a given model are estimated by minimizing the sum of squared pricing errors between the market price and the model-determined price for each option. The daily average of the estimated parameters is reported first, followed by its standard error in parentheses. The average implied volatility obtained from inverting the Black–Scholes model (using a short-term at-the-money option) is, respectively, 18.47 %, 17.72 %, 17.41 %, and 20.52 % over the sample periods 06:1988–05:1991, 06:1988–05:1989, 06:1989–05:1990, and 06:1990–05:1991

98.4.1 Static Performance

To examine out-of-sample cross-sectional pricing performance for each model, we use *previous day's* option prices to back out the required parameter values and then use them as input to compute *current day's* model-based option prices. Next, subtract the model-determined price from the observed market price, to compute both the absolute pricing error and the percentage pricing error. This procedure is repeated for each call and each day in the sample, to obtain the average absolute and the average percentage pricing errors and their associated standard errors. These steps are separately followed for each of the BS, the SI, the SV, and the SVSI models. The results from this exercise are reported in Table 98.3.

Let's first examine the relative performance in pricing OTM options. Overpricing of OTM options is often considered a critical problem for the BS model (e.g., McBeth and Merville 1979 and Rubinstein 1985). Panel A of Table 98.3 reports the absolute and the percentage pricing error estimates for OTM options. According to both error measures, the overall ranking of the four models is consistent with our priors: the SVSI model outperforms all others, followed by the SV, the SI, and finally the BS model. For extremely short-term (<30 days) and extremely out-of-the-money ($\frac{S}{K} < 0.93$) options, for example, the average absolute pricing error by the SVSI model is \$0.23 versus \$0.53 by the BS, \$0.28 by the SI, and \$0.25 by the SV model. For this category, the BS model's absolute pricing error is cut by more than a half by each of the other three models. Fix the moneyness category at $\frac{S}{K} \in (0.93, 0.95)$. Then, for medium-term (120–180 days) options, the SVSI model produces an average absolute pricing error of \$0.44 versus \$1.38 by the BS, \$0.72 by the SI, and \$0.39 by the SV model. For short-term (30–60 day) calls, the absolute pricing errors are \$0.44 by the SVSI, \$0.48 by the SV, \$0.73 by the SI, and \$0.90 by the BS model. Clearly, the performance improvement is significant for each moneyness and maturity category in Panel A, from the BS to the SI, to the SV, and to the SVSI model. This pricing performance ranking of the four models can also be seen using the average percentage pricing errors, as given in the same table. Here, the SVSI model produces percentage pricing errors that are the lowest in magnitude. As an example, take OTM options with term to expiration of 30–60 days and with $\frac{S}{K} \in (0.93, 0.95)$ In this category the BS, the SI, the SV, and the SVSI models, respectively, have average percentage pricing errors of -54.50 %, -46.20 %, -26.16 %, and -18.85 %. For long-term options with $\frac{S}{K} \in (0.93, 0.95)$ and with $\frac{S}{K} \in (0.95, 0.97)$, the SVSI model results in a percentage pricing error that is as low as 0.71 % and 0.30 %, respectively.

For ATM calls, recall that the BS model's implied-volatility input is backed out from the (previous day's) short-term ATM options, which should give the BS model a relative advantage in pricing ATM options. In contrast, the implied spot variance for the other models is obtained by minimizing the sum of squared errors for all options of the previous day. Thus, for ATM options, one would expect the BS model to perform relatively better. As seen from Panel B of Table 98.3, except for the shortest-term ATM calls, the SVSI model typically generates the lowest absolute and percentage pricing errors (especially for longer-term options), followed by the SV, by the SI,

Table 98.3 Out-of-sample pricing errors

Moneyness S/K	Model	Percentage pricing error						Absolute pricing error					
		Term to expiration (days)						Term to expiration (days)					
		<30	30–60	60–90	90–120	120–180	≥180	<30	30–60	60–90	90–120	120–180	≥180
Panel A: out-of-the-money options													
<0.93	BS	−65.99	−86.80	−62.45	−57.63	−47.71	−33.72	0.53	1.00	1.14	1.50	1.96	2.36
		(12.02)	(4.51)	(2.96)	(2.92)	(1.37)	(1.05)	(0.10)	(0.04)	(0.05)	(0.06)	(0.05)	(0.06)
	SI	−24.53	−58.13	−40.04	−28.43	−16.70	−3.92	0.24	0.66	0.72	0.80	0.91	0.96
		(6.59)	(3.81)	(2.60)	(1.67)	(0.95)	(0.63)	(0.04)	(0.03)	(0.04)	(0.03)	(0.05)	(0.05)
	SV	−22.08	−30.38	−12.43	−4.02	0.89	6.08	0.25	0.44	0.34	0.33	0.43	0.62
		(6.90)	(3.07)	(1.54)	(0.90)	(0.47)	(0.39)	(0.04)	(0.03)	(0.02)	(0.02)	(0.04)	(0.05)
	SVSI	−16.29	−21.96	−5.68	−1.68	0.92	0.18	0.23	0.38	0.29	0.33	0.46	0.66
		(7.79)	(2.64)	(1.40)	(0.93)	(0.51)	(0.64)	(0.05)	(0.02)	(0.02)	(0.02)	(0.04)	(0.04)
0.93–0.95	BS	−53.68	−54.50	−33.82	−21.88	−16.43	−11.25	0.56	0.90	1.05	1.24	1.38	1.80
		(5.31)	(2.08)	(1.79)	(1.25)	(0.61)	(0.56)	(0.04)	(0.03)	(0.04)	(0.06)	(0.04)	(0.06)
	SI	−42.06	−49.30	−32.22	−15.78	−10.18	−5.91	0.42	0.77	0.92	0.83	0.85	0.98
		(5.32)	(2.18)	(2.07)	(1.07)	(0.55)	(0.43)	(0.03)	(0.02)	(0.05)	(0.05)	(0.03)	(0.05)
	SV	−25.68	−26.16	−8.83	−3.39	−0.55	1.23	0.40	0.48	0.35	0.39	0.39	0.52
		(4.61)	(1.43)	(0.81)	(0.61)	(0.30)	(0.24)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
	SVSI	−22.50	−18.85	−4.84	−2.39	0.66	0.71	0.38	0.44	0.31	0.42	0.44	0.58
		(4.53)	(1.43)	(0.85)	(0.74)	(0.32)	(0.26)	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)	(0.03)

(*continued*)

Table 98.3 (continued)

Moneyness S/K	Model	Percentage pricing error						Absolute pricing error					
		Term to expiration (days)						Term to expiration (days)					
		<30	30–60	60–90	90–120	120–180	≥180	<30	30–60	60–90	90–120	120–180	≥180
0.95–0.97	BS	−36.61	−28.83	−16.21	−9.91	−7.75	−5.77	0.55	0.81	0.87	1.03	1.05	1.44
		(2.33)	(0.93)	(0.95)	(0.84)	(0.41)	(0.45)	(0.03)	(0.02)	(0.04)	(0.05)	(0.04)	(0.06)
	SI	−35.83	−30.09	−18.97	−7.44	−5.70	−3.62	0.51	0.81	0.92	0.69	0.79	0.86
		(2.45)	(1.09)	(1.30)	(0.68)	(0.41)	(0.32)	(0.04)	(0.02)	(0.05)	(0.04)	(0.03)	(0.04)
	SV	−23.68	−16.94	−5.63	−1.63	−0.26	0.56	0.45	0.51	0.40	0.40	0.41	0.49
		(2.06)	(0.68)	(0.58)	(0.42)	(0.22)	(0.20)	(0.03)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
	SVSI	−16.90	−13.53	−3.59	−1.80	0.05	0.30	0.42	0.49	0.38	0.47	0.45	0.56
		(2.01)	(0.72)	(0.60)	(0.51)	(0.23)	(0.21)	(0.03)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)
Panel B: at-the-money options													
0.97–0.99	BS	−19.94	−10.16	−4.83	−2.66	−2.21	−1.42	0.51	0.66	0.63	0.77	0.82	1.18
		(1.03)	(0.49)	(0.58)	(0.56)	(0.30)	(0.35)	(0.02)	(0.02)	(0.03)	(0.05)	(0.03)	(0.05)
	SI	−22.85	−11.38	−7.76	−1.69	−2.29	−1.76	0.53	0.68	0.71	0.58	0.67	0.74
		(1.16)	(0.59)	(0.70)	(0.46)	(0.30)	(0.23)	(0.02)	(0.02)	(0.04)	(0.03)	(0.03)	(0.03)
	SV	−18.93	−8.37	−2.76	−0.44	−0.16	−0.04	0.50	0.52	0.37	0.37	0.40	0.48
		(1.01)	(0.40)	(0.37)	(0.28)	(0.17)	(0.15)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
	SVSI	−17.10	−7.74	−2.36	−1.04	−0.42	−0.29	0.49	0.53	0.38	0.46	0.44	0.52
		(1.05)	(0.45)	(0.39)	(0.37)	(0.17)	(0.15)	(0.02)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)

0.99–1.01	BS	−4.42	−1.29	1.14	1.44	1.06	0.86	0.44	0.57	0.54	0.74	0.77	0.97
		(0.47)	(0.30)	(0.39)	(0.45)	(0.24)	(0.26)	(0.01)	(0.02)	(0.03)	(0.04)	(0.03)	(0.04)
	SI	−5.97	−2.28	−1.08	1.81	0.29	−0.63	0.50	0.63	0.59	0.68	0.70	0.75
		(0.57)	(0.38)	(0.48)	(0.39)	(0.25)	(0.22)	(0.02)	(0.02)	(0.04)	(0.05)	(0.03)	(0.04)
	SV	−7.93	−3.72	−0.76	0.51	0.13	0.05	0.50	0.50	0.38	0.43	0.42	0.51
		(0.47)	(0.26)	(0.27)	(0.25)	(0.14)	(0.13)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
	SVSI	−8.22	−3.93	−0.74	−0.40	−0.31	−0.25	0.51	0.55	0.42	0.51	0.46	0.59
		(0.49)	(0.29)	(0.30)	(0.31)	(0.15)	(0.15)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)
1.01–1.03	BS	2.43	3.01	4.15	3.76	3.36	2.47	0.46	0.63	0.77	0.96	1.01	1.42
		(0.23)	(0.19)	(0.29)	(0.38)	(0.19)	(0.28)	(0.01)	(0.02)	(0.04)	(0.05)	(0.03)	(0.05)
	SI	1.65	2.29	2.08	3.52	1.76	−0.34	0.47	0.65	0.70	0.88	0.76	0.81
		(0.27)	(0.23)	(0.38)	(0.30)	(0.19)	(0.20)	(0.02)	(0.02)	(0.04)	(0.05)	(0.03)	(0.04)
	SV	−0.99	−0.69	0.46	1.04	0.38	−0.21	0.39	0.42	0.37	0.44	0.42	0.49
		(0.23)	(0.17)	(0.22)	(0.20)	(0.11)	(0.11)	(0.01)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)
	SVSI	−1.49	−1.16	0.12	0.27	−0.19	−0.55	0.41	0.46	0.39	0.48	0.46	0.54
		(0.23)	(0.18)	(0.24)	(0.24)	(0.12)	(0.11)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)
Panel C: in-the-money options													
1.03–1.05	BS	3.69	4.45	5.31	4.76	4.38	2.98	0.59	0.85	1.14	1.20	1.29	1.40
		(0.13)	(0.14)	(0.24)	(0.29)	(0.17)	(0.24)	(0.02)	(0.03)	(0.05)	(0.07)	(0.04)	(0.06)
	SI	3.37	3.83	3.64	4.08	2.51	0.17	0.57	0.82	0.92	1.04	0.90	0.83
		(0.15)	(0.18)	(0.28)	(0.25)	(0.17)	(0.19)	(0.02)	(0.03)	(0.05)	(0.06)	(0.04)	(0.04)
	SV	1.27	0.79	1.09	1.11	0.43	−0.20	0.38	0.42	0.42	0.43	0.42	0.50
		(0.13)	(0.13)	(0.17)	(0.15)	(0.10)	(0.11)	(0.01)	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)
	SVSI	0.84	0.32	0.83	0.29	−0.03	−0.41	0.37	0.45	0.42	0.49	0.45	0.54
		(0.13)	(0.14)	(0.19)	(0.20)	(0.10)	(0.11)	(0.01)	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)

(*continued*)

Table 98.3 (continued)

Moneyness S/K	Model	Percentage pricing error						Absolute pricing error					
		Term to expiration (days)						Term to expiration (days)					
		<30	30–60	60–90	90–120	120–180	≥180	<30	30–60	60–90	90–120	120–180	≥180
1.05–1.07	BS	3.37	4.54	5.57	5.08	4.82	4.27	0.70	1.06	1.46	1.47	1.56	1.73
		(0.10)	(0.11)	(0.22)	(0.27)	(0.14)	(0.22)	(0.02)	(0.02)	(0.05)	(0.06)	(0.04)	(0.06)
	SI	3.28	4.02	4.08	4.47	2.65	0.59	0.69	0.97	1.13	1.29	0.99	0.88
		(0.12)	(0.14)	(0.24)	(0.22)	(0.15)	(0.18)	(0.02)	(0.03)	(0.06)	(0.06)	(0.04)	(0.05)
	SV	1.82	1.41	1.47	1.44	0.54	−0.40	0.45	0.46	0.53	0.50	0.45	0.57
		(0.09)	(0.09)	(0.16)	(0.14)	(0.09)	(0.11)	(0.02)	(0.02)	(0.03)	(0.03)	(0.02)	(0.03)
	SVSI	1.59	1.12	1.35	0.83	0.17	−0.52	0.42	0.46	0.52	0.51	0.44	0.58
		(0.09)	(0.10)	(0.17)	(0.17)	(0.09)	(0.11)	(0.01)	(0.02)	(0.03)	(0.04)	(0.02)	(0.03)
>1.07	BS	1.79	2.65	2.96	3.10	3.36	2.61	0.60	0.95	1.22	1.35	1.56	1.58
		(0.04)	(0.05)	(0.07)	(0.08)	(0.05)	(0.05)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)	(0.02)
	SI	1.86	2.50	2.14	2.45	1.63	−0.76	0.59	0.89	0.88	1.04	0.86	1.09
		(0.05)	(0.06)	(0.07)	(0.08)	(0.05)	(0.05)	(0.01)	(0.02)	(0.02)	(0.03)	(0.02)	(0.02)
	SV	1.36	1.33	1.06	0.92	0.45	−0.64	0.50	0.55	0.52	0.49	0.42	0.64
		(0.13)	(0.13)	(0.17)	(0.15)	(0.10)	(0.11)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
	SVSI	1.29	1.26	1.18	0.81	0.40	−0.37	0.46	0.55	0.58	0.52	0.44	0.57
		(0.03)	(0.03)	(0.04)	(0.04)	(0.03)	(0.03)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)

For a given model, compute the price of each option using previous day's implied parameters and implied stock volatility. The reported percentage pricing error is the sample average of the market price minus the model price divided by the market price. The reported absolute pricing error is the sample average of the absolute difference between the market price and the model price for each call. The corresponding standard errors are recorded in parentheses. The sample period is 06:1988–05:1991, with a total of 38,749 call option prices and finally by the BS model. For the shortest-term options with $\frac{S}{K} \in (0.97, 0.99)$ and $\frac{S}{K} \in (0.99, 1.01)$, the BS and the SI models perform somewhat better than the other two.

Panel C of Table 98.3 reports the average absolute and percentage pricing errors of ITM calls by all four models. While the previous ranking of the models based on OTM and ATM options is preserved by Panel C, it can be noted that the average percentage pricing error is below 1.0 % for 12 out of the 18 categories in the case of the SVSI model, for 8 out of the 18 categories in the case of the SV model, for three categories out of 18 for the SI model, and for none of the 18 categories in the case of the BS model. The pricing improvement by the SV and the SVSI models over the BS and the SI is quite substantial for ITM options, especially for long-term options.

Some patterns of mispricing can, however, be noted across all moneyness–maturity categories. First, all four models produce negative percentage pricing errors for options with moneyness $\frac{S}{K} \leq 0.99$ and positive percentage pricing errors for options with $\frac{S}{K} \geq 1.03$, subject to their time to expiration not exceeding 120 days. This means that the models systematically overprice OTM call options while underprice ITM calls. But the magnitude of such mispricing varies dramatically across the models, with the BS producing the strongest and the SVSI model the weakest systematic biases. Next, according to the absolute pricing error measure, the SV model seems to perform slightly better than the SVSI in pricing calls with more than 90 days to expiration. This pattern is, however, not supported by the percentage pricing errors reported in Table 98.3, possibly because for these relatively long-term calls the two models produce pricing errors that have mixed signs, in which case taking the average absolute value of the pricing errors can sometimes distort the picture. According to the percentage pricing errors, the SVSI model does slightly better than the SV in pricing those longer-term options. Finally, for the BS model, its absolute pricing error has a U-shaped relationship (i.e., "smile") with moneyness, and the magnitude of its percentage pricing error increases as the call goes from deep in the money to deep out of the money, regardless of time to expiration. These patterns are reduced by each relaxation of the BS model assumptions.

98.4.2 Dynamic Hedging Performance

Recall that in implementing a hedge using any of the four models, we follow three basic steps. First, based on Procedure B of Sect. 98.2.3, estimate the structural parameters and spot variance by using day 1's option prices. Next, on day 2, use previous day's parameter and spot volatility estimates and current day's spot price and interest rates, to construct the desired hedge as given in Sect. 98.2.2. Finally, rely on either Eqs. 98.22 or 98.25 to calculate the hedging error as of day 3. We then compute both the average absolute and the average dollar hedging errors of all call options in a given moneyness–maturity category, to gauge the relative hedging performance of each model.

It should be recognized that in both the delta-neutral and the minimum-variance hedging exercises conducted in the two subsections below, the spot S&P 500 index, rather than an S&P 500 futures contract, is used in place of the "spot asset" for the hedges devised in Sect. 98.2.2. This is done out of two considerations. First, the spot S&P 500 and the immediate-expiration-month S&P 500 futures price generally have a correlation coefficient close to one. This means that whether the spot index or the futures price is used in the hedging exercises, the qualitative as well as the quantitative conclusions are most likely the same. In other words, if it is demonstrated using the spot index that one model results in better hedging performance than another, the same hedging performance ranking of the two models will likely be achieved by using an S&P 500 futures contract. After all, our main interest here lies in the relative performance of the models. Second, when a futures contract is used in constructing a hedge, a futures pricing formula has to be adopted. That will introduce another dimension of model misspecification (due to stochastic interest rates), which can in turn produce a compounded effect on the hedging results. For these reasons, using the spot index may lead to a cleaner comparison among the four option models.

98.4.2.1 Effectiveness of Delta-Neutral Hedges

Observe that the construction and the execution of the hedging strategy in Eq. 98.22 requires, in the cases of the SV and the SVSI models, (i) the availability of prices for four time-matched target and hedging-instrumental options, $C(t,\tau;K)$, $C(t,\tau;\overline{K})$, $C(t+\Delta t,\tau-\Delta t;K)$, and $C(t+\Delta t,\tau-\Delta t;\overline{K})$, and (ii) the computation of Δ_S, Δ_V, and Δ_R for the target and the instrumental option. Due to this requirement, it is important to match as closely as possible the time points at which the target and the instrumental option prices were respectively taken, in order to ensure that the hedge ratios are properly determined. For this reason, we use as hedging instruments only options whose prices on both the hedge construction day and the following liquidation day were quoted no more than 15 s apart from the times when the respective prices for the target option were quoted. This requirement makes the overall sample for the hedging exercise smaller than that used for the preceding pricing exercise, but it nonetheless guarantees that the deltas for the target and instrumental options on the same day are computed based on the same spot price. The remaining sample contains 15,041 matched pairs when hedging revision occurs at 1-day intervals and 11,704 matched pairs when rebalancing takes place at 5-day intervals. In addition, we partition the target options into three maturity classes, less than 60 days, 60–180 days, and greater than 180 days, and report hedging results accordingly.

In theory, a call option with any expiration date and any strike price can be chosen as a hedging instrument for any given target option. In practice, however, different choices can mean different hedging effectiveness, even for the same option pricing model. Out of this consideration, we employ as a hedging instrument the call option which has the same expiration date as the target option and whose strike price is the closest, but not identical, to the target option's.

Table 98.4 presents delta-neutral hedging results for the four models. Several patterns emerge from this Table. First, the BS model produces the worst hedging performance by most measures, the SI shows noticeable improvement according to the average dollar hedging errors (especially in the 5-day hedging revision categories) but not so according to the average absolute hedging errors, while the SV and the SVSI models have average absolute and average dollar hedging errors that are typically one-third of the corresponding BS hedging errors, or lower. The improvement by the SV and the SVSI is thus remarkable. Second, as portfolio adjustment frequency decreases from daily to once every 5 days, hedging effectiveness deteriorates, regardless of the model used. The deterioration is especially apparent for OTM and ATM options with $\frac{S}{K} \leq 1.05$. It should, however, be noted with emphasis that for both the SV and the SVSI models, their hedging effectiveness is relatively stable, whether the hedges are rebalanced each day or once every 5 days. For the BS and the SI models, such a change in revision frequency can mean doubling their hedging errors. This finding is strong evidence in support of the SV and the SVSI models for hedging.

Third, the BS model-based delta-neutral hedging strategy always overhedges a target call option, as its average dollar hedging error is negative for each moneyness–maturity category and at either frequency of portfolio rebalancing. In contrast, the dollar hedging errors based on the SV and the SVSI models are more random and can take either sign. Therefore, the BS formula has a systematic hedging bias pattern, whereas the SV and the SVSI do not.

Fourth, the SVSI model is indistinguishable from the SV according to their absolute hedging errors, but is slightly better than the latter when judged using their average dollar hedging errors. Similarly, the SI model has worse hedging performance than the BS according to their absolute hedging error values, but the reverse is true according to their dollar hedging errors. This phenomenon exists possibly because with stochastic interest rates there are larger hedging errors of opposite signs, so that when added together, these errors cancel out, but the sum of their absolute values is nonetheless large.

Finally, no matter which model is used, there do not appear to be moneyness- or maturity-related bias patterns in the hedging errors. In other words, hedging errors do not seem to "smile" across exercise prices or times to expiration, as pricing errors do. This is a striking disparity between pricing and hedging results.

98.4.2.2 Effectiveness of Single-Instrument Minimum-Variance Hedges

If one is, for reasons given before, constrained to using only the underlying stock to hedge a target call option, dimensions of uncertainty that move the target option value but are uncorrelated with the underlying stock price cannot be hedged by any position in the stock and will necessarily be uncontrolled for in such a single-instrument minimum-variance hedge. Based on the sample option data, the average absolute and the average dollar hedging errors, with either a daily or a 5-day rebalancing frequency, are given in Table 98.5 for each of the four models and each of the moneyness–maturity categories. With this type of hedges, the relative performance of the models is no longer clear-cut. For OTM options with $\frac{S}{K} \leq 1.97$,

Table 98.4 Delta-neutral hedging errors

		Dollar hedging error						Absolute hedging error					
		1-day revision			5-day revision			1-day revision			5-day revision		
Moneyness	Model	Term to expiration (days)						Term to expiration (days)					
S/K		<60	60–180	≥180	<60	60–180	≥180	<60	60–180	≥180	<60	60–180	≥180
Panel A: out-of-the-money options													
<0.93	BS	NA	−0.06	−0.04	NA	−0.33	−0.21	NA	0.37	0.45	NA	0.91	0.87
			(0.03)	(0.02)		(0.09)	(0.06)		(0.01)	(0.01)		(0.06)	(0.04)
	SI		−0.08	−0.05		−0.40	−0.36		0.35	0.40		0.69	0.81
			(0.02)	(0.02)		(0.05)	(0.06)		(0.01)	(0.01)		(0.03)	(0.04)
	SV		0.02	0.01		0.03	0.03		0.15	0.14		0.17	0.21
			(0.02)	(0.02)		(0.02)	(0.02)		(0.01)	(0.01)		(0.02)	(0.02)
	SVSI		0.02	0.00		0.03	0.00		0.16	0.14		0.17	0.22
			(0.02)	(0.02)		(0.02)	(0.03)		(0.01)	(0.01)		(0.01)	(0.02)
0.93–0.95	BS	NA	−0.06	−0.01	NA	−0.24	−0.02	NA	0.32	0.46	NA	0.79	0.82
			(0.02)	(0.03)		(0.07)	(0.07)		(0.01)	(0.02)		(0.04)	(0.05)
	SI		−0.08	0.00		−0.29	−0.22		0.33	0.43		0.67	0.66
			(0.02)	(0.04)		(0.05)	(0.07)		(0.01)	(0.02)		(0.03)	(0.05)
	SV		−0.01	−0.01		−0.00	0.00		0.12	0.23		0.13	0.18
			(0.01)	(0.02)		(0.01)	(0.03)		(0.01)	(0.01)		(0.01)	(0.02)
	SVSI		−0.01	−0.01		−0.00	0.00		0.12	0.24		0.13	0.18
			(0.01)	(0.02)		(0.01)	(0.03)		(0.01)	(0.02)		(0.01)	(0.02)

0.95–0.97	BS	−0.08	−0.06	−0.01	−0.55	−0.21	−0.12	0.23	0.33	0.45	0.66	0.77	0.85
		(0.06)	(0.02)	(0.03)	(0.16)	(0.06)	(0.08)	(0.04)	(0.02)	(0.03)	(0.13)	(0.04)	(0.05)
	SI	−0.06	−0.06	−0.06	−0.22	−0.31	−0.34	0.27	0.34	0.41	0.61	0.71	0.89
		(0.02)	(0.02)	(0.04)	(0.05)	(0.05)	(0.09)	(0.01)	(0.01)	(0.02)	(0.03)	(0.03)	(0.06)
	SV	0.03	−0.01	−0.01	−0.03	−0.00	−0.02	0.10	0.13	0.19	0.10	0.14	0.26
		(0.03)	(0.01)	(0.02)	(0.03)	(0.01)	(0.03)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)
	SVSI	0.02	−0.01	−0.01	−0.02	−0.01	−0.01	0.09	0.13	0.20	0.08	0.14	0.27
		(0.03)	(0.01)	(0.02)	(0.02)	(0.01)	(0.03)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)
Panel B: at-the-money options													
0.97–0.99	BS	−0.01	−0.04	−0.04	−0.36	−0.11	−0.21	0.34	0.34	0.46	0.54	0.75	0.89
		(0.05)	(0.02)	(0.03)	(0.08)	(0.05)	(0.07)	(0.03)	(0.01)	(0.02)	(0.06)	(0.03)	(0.05)
	SI	0.08	−0.05	−0.07	−0.30	−0.25	−0.44	0.29	0.36	0.41	0.73	0.70	0.79
		(0.01)	(0.02)	(0.02)	(0.05)	(0.05)	(0.08)	(0.01)	(0.01)	(0.02)	(0.03)	(0.03)	(0.06)
	SV	−0.03	0.01	0.00	−0.04	−0.00	−0.01	0.12	0.13	0.17	0.14	0.14	0.23
		(0.02)	(0.01)	(0.01)	(0.03)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
	SVSI	−0.02	0.01	0.00	0.02	0.00	−0.01	0.12	0.13	0.17	0.13	0.14	0.24
		(0.02)	(0.01)	(0.01)	(0.03)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)
0.99–1.01	BS	−0.10	−0.02	−0.01	−0.43	−0.08	−0.10	0.37	0.37	0.47	0.80	0.77	0.77
		(0.01)	(0.02)	(0.03)	(0.09)	(0.05)	(0.07)	(0.01)	(0.01)	(0.02)	(0.06)	(0.03)	(0.05)
	SI	−0.08	−0.05	−0.03	−0.29	−0.24	−0.18	0.36	0.37	0.42	0.81	0.67	0.61
		(0.02)	(0.02)	(0.03)	(0.05)	(0.05)	(0.08)	(0.01)	(0.01)	(0.02)	(0.03)	(0.03)	(0.05)
	SV	0.01	−0.00	−0.01	0.02	0.01	0.03	0.14	0.13	0.17	0.15	0.15	0.25
		(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)
	SVSI	0.01	0.00	−0.01	0.02	−0.00	0.04	0.14	0.13	0.17	0.16	0.15	0.25
		(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.03)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)

(*continued*)

Table 98.4 (continued)

Moneyness	Model	Dollar hedging error						Absolute hedging error					
		1-day revision			5-day revision			1-day revision			5-day revision		
		Term to expiration (days)						Term to expiration (days)					
S/K		<60	60–180	≥180	<60	60–180	≥180	<60	60–180	≥180	<60	60–180	≥180
1.01–1.03	BS	−0.09	−0.02	−0.01	−0.40	−0.11	−0.09	0.40	0.39	0.46	0.82	0.75	0.82
		(0.03)	(0.02)	(0.03)	(0.08)	(0.05)	(0.07)	(0.02)	(0.01)	(0.02)	(0.05)	(0.03)	(0.05)
	SI	−0.09	−0.05	−0.07	−0.30	−0.25	−0.27	0.38	0.36	0.43	0.75	0.65	0.71
		(0.02)	(0.02)	(0.04)	(0.05)	(0.05)	(0.09)	(0.01)	(0.01)	(0.01)	(0.03)	(0.03)	(0.06)
	SV	0.00	−0.00	0.03	0.01	−0.01	0.05	0.13	0.14	0.17	0.13	0.17	0.24
		(0.03)	(0.01)	(0.01)	(0.02)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
	SVSI	0.00	−0.00	0.03	−0.00	−0.01	0.05	0.14	0.14	0.17	0.13	0.17	0.24
		(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
Panel C: in-the-money options													
1.03–1.05	BS	−0.06	−0.03	−0.05	−0.36	−0.09	−0.23	0.40	0.38	0.47	0.70	0.69	0.90
		(0.02)	(0.02)	(0.03)	(0.05)	(0.04)	(0.08)	(0.02)	(0.01)	(0.02)	(0.03)	(0.03)	(0.05)
	SI	−0.09	−0.05	−0.07	−0.31	−0.19	−0.35	0.39	0.36	0.41	0.65	0.61	0.81
		(0.02)	(0.02)	(0.04)	(0.06)	(0.05)	(0.09)	(0.01)	(0.01)	(0.02)	(0.03)	(0.03)	(0.06)
	SV	0.01	0.01	−0.01	0.00	0.00	−0.03	0.15	0.13	0.15	0.17	0.15	0.24
		(0.03)	(0.01)	(0.02)	(0.01)	(0.01)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
	SVSI	0.01	0.00	−0.00	0.00	0.00	−0.03	0.15	0.12	0.16	0.16	0.14	0.25
		(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)

1.05–1.07	BS	−0.05	−0.02	−0.06	−0.35	−0.06	−0.22	0.41	0.40	0.47	0.68	0.64	0.77
		(0.02)	(0.02)	(0.04)	(0.04)	(0.04)	(0.07)	(0.01)	(0.01)	(0.02)	(0.02)	(0.03)	(0.04)
	SI	−0.07	−0.04	−0.11	−0.26	−0.12	−0.56	0.40	0.37	0.44	0.57	0.51	0.59
		(0.02)	(0.02)	(0.04)	(0.04)	(0.04)	(0.08)	(0.01)	(0.02)	(0.02)	(0.03)	(0.03)	(0.01)
	SV	−0.00	−0.00	−0.01	−0.05	−0.02	0.00	0.16	0.13	0.18	0.18	0.15	0.22
		(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
	SVSI	−0.00	0.00	−0.00	−0.03	−0.02	0.00	0.15	0.12	0.17	0.17	0.15	0.22
		(0.01)	(0.01)	(0.02)	(0.01)	(0.00)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
>1.07	BS	−0.04	−0.03	−0.02	−0.15	−0.07	−0.10	0.36	0.39	0.48	0.51	0.58	0.72
		(0.01)	(0.00)	(0.01)	(0.02)	(0.02)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
	SI	−0.05	−0.04	−0.03	−0.18	0.08	−0.21	0.35	0.37	0.43	0.45	0.51	0.66
		(0.03)	(0.01)	(0.02)	(0.02)	(0.02)	(0.08)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
	SV	−0.01	−0.01	−0.00	−0.03	−0.01	0.01	0.15	0.14	0.20	0.17	0.18	0.27
		(0.01)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)	(0.01)	(0.01)	(0.01)	(0.00)	(0.00)	(0.00)
	SVSI	−0.00	−0.00	0.00	−0.01	0.00	0.01	0.15	0.14	0.20	0.16	0.17	0.27
		(0.00)	(0.00)	(0.01)	(0.01)	(0.00)	(0.01)	(0.00)	(0.00)	(0.00)	(0.00)	(0.00)	(0.01)

For each call option, calculate the hedging error, which is the difference between the market price of the call and the replicating portfolio. The average dollar hedging error and the average absolute hedging error are reported for each model. The standard errors are given in parentheses. The sample period is 06:1988–05:1991. In calculating the hedging errors generated with daily (once every 5 days) hedge rebalancing, 15,041 (11,704) observations are used

Table 98.5 Single-instrument hedging errors

		Dollar hedging error						Absolute hedging error					
		1-day revision			5-day revision			1-day revision			5-day revision		
Moneyness	Model	Term to expiration (days)						Term to expiration (days)					
S/K		<60	60–180	≥180	<60	60–180	≥180	<60	60–180	≥180	<60	60–180	≥180
Panel A: out-of-the-money options													
<0.93	BS	NA	−0.06	−0.04	NA	−0.33	−0.21	NA	0.37	0.45	NA	0.91	0.87
			(0.03)	(0.02)		(0.09)	(0.06)		(0.01)	(0.01)		(0.06)	(0.04)
	SI		−0.09	−0.06		−0.51	−0.45		0.49	0.52		1.33	0.91
			(0.05)	(0.04)		(0.17)	(0.08)		(0.03)	(0.03)		(0.10)	(0.05)
	SV		−0.02	−0.05		−0.03	−0.09		0.30	0.39		0.75	0.75
			(0.03)	(0.02)		(0.08)	(0.05)		(0.02)	(0.02)		(0.05)	(0.04)
	SVSI		0.02	−0.04		0.11	−0.13		0.35	0.43		0.76	0.82
			(0.03)	(0.03)		(0.08)	(0.05)		(0.03)	(0.02)		(0.05)	(0.04)
0.93–0.95	BS	NA	−0.06	−0.01	NA	−0.24	−0.02	NA	0.32	0.46	NA	0.79	0.82
			(0.02)	(0.03)		(0.07)	(0.07)		(0.01)	(0.02)		(0.04)	(0.05)
	SI		−0.10	−0.00		−0.36	−0.38		0.35	0.52		0.97	0.72
			(0.03)	(0.04)		(0.10)	(0.08)		(0.02)	(0.03)		(0.06)	(0.06)
	SV		−0.06	0.00		−0.14	−0.00		0.33	0.43		0.79	0.73
			(0.02)	(0.03)		(0.07)	(0.06)		(0.02)	(0.02)		(0.04)	(0.04)
	SVSI		−0.04	−0.00		−0.17	−0.17		0.36	0.47		0.82	0.78
			(0.03)	(0.03)		(0.07)	(0.07)		(0.02)	(0.02)		(0.04)	(0.04)

0.95–0.97	BS	−0.08	−0.06	−0.01	−0.55	−0.21	−0.12	0.23	0.33	0.45	0.66	0.77	0.85
		(0.06)	(0.02)	(0.03)	(0.16)	(0.06)	(0.08)	(0.04)	(0.02)	(0.03)	(0.13)	(0.04)	(0.05)
	SI	−0.04	−0.09	0.04	−0.44	−0.38	−0.42	0.33	0.39	0.47	0.74	0.94	0.92
		(0.11)	(0.03)	(0.05)	(0.05)	(0.09)	(0.10)	(0.01)	(0.01)	(0.03)	(0.09)	(0.06)	(0.07)
	SV	−0.03	−0.05	−0.01	−0.06	−0.14	−0.16	0.21	0.32	0.42	0.48	0.77	0.74
		(0.06)	(0.02)	(0.03)	(0.16)	(0.06)	(0.07)	(0.04)	(0.02)	(0.02)	(0.09)	(0.04)	(0.04)
	SVSI	−0.04	−0.02	−0.00	−0.21	−0.14	−0.23	0.22	0.35	0.46	0.65	0.83	0.79
		(0.01)	(0.02)	(0.03)	(0.03)	(0.04)	(0.06)	(0.03)	(0.01)	(0.02)	(0.11)	(0.04)	(0.04)
Panel B: at-the-money options													
0.97–0.99	BS	−0.01	−0.04	−0.04	−0.36	−0.11	−0.21	0.34	0.34	0.46	0.54	0.75	0.89
		(0.05)	(0.02)	(0.03)	(0.08)	(0.05)	(0.07)	(0.03)	(0.01)	(0.02)	(0.06)	(0.03)	(0.05)
	SI	0.00	−0.06	−0.06	−0.34	−0.11	−0.42	0.34	0.40	0.51	0.82	0.90	0.82
		(0.06)	(0.03)	(0.04)	(0.05)	(0.09)	(0.08)	(0.05)	(0.02)	(0.03)	(0.09)	(0.06)	(0.06)
	SV	−0.01	−0.05	−0.06	−0.20	−0.11	−0.19	0.35	0.35	0.44	0.56	0.81	0.84
		(0.06)	(0.02)	(0.03)	(0.12)	(0.06)	(0.07)	(0.04)	(0.01)	(0.02)	(0.07)	(0.03)	(0.04)
	SVSI	−0.03	−0.05	−0.03	−0.15	−0.16	−0.27	0.36	0.37	0.45	0.59	0.86	0.89
		(0.06)	(0.02)	(0.03)	(0.11)	(0.06)	(0.07)	(0.04)	(0.01)	(0.02)	(0.07)	(0.03)	(0.05)
0.99–1.01	BS	−0.10	−0.02	−0.01	−0.43	−0.08	−0.10	0.37	0.37	0.47	0.80	0.77	0.77
		(0.01)	(0.02)	(0.03)	(0.09)	(0.05)	(0.07)	(0.01)	(0.01)	(0.02)	(0.06)	(0.03)	(0.05)
	SI	−0.15	−0.04	0.03	−0.34	−0.18	−0.31	0.39	0.41	0.55	0.77	0.82	0.78
		(0.05)	(0.03)	(0.05)	(0.15)	(0.07)	(0.09)	(0.03)	(0.02)	(0.03)	(0.09)	(0.05)	(0.06)
	SV	−0.12	−0.02	−0.00	−0.22	−0.12	−0.06	0.38	0.37	0.45	0.78	0.79	0.69
		(0.04)	(0.02)	(0.03)	(0.11)	(0.05)	(0.06)	(0.03)	(0.01)	(0.02)	(0.06)	(0.03)	(0.04)
	SVSI	−0.04	−0.01	−0.00	−0.27	−0.13	−0.15	0.38	0.40	0.47	0.88	0.89	0.80
		(0.01)	(0.02)	(0.03)	(0.11)	(0.05)	(0.06)	(0.03)	(0.01)	(0.02)	(0.07)	(0.04)	(0.04)

(continued)

Table 98.5 (continued)

		Dollar hedging error						Absolute hedging error					
		1-day revision			5-day revision			1-day revision			5-day revision		
Moneyness	Model	Term to expiration (days)						Term to expiration (days)					
S/K		<60	60–180	≥180	<60	60–180	≥180	<60	60–180	≥180	<60	60–180	≥180
1.01–1.03	BS	−0.09	−0.03	−0.01	−0.40	−0.11	−0.09	0.40	0.39	0.46	0.82	0.75	0.82
		(0.03)	(0.02)	(0.03)	(0.08)	(0.05)	(0.07)	(0.02)	(0.01)	(0.02)	(0.05)	(0.03)	(0.05)
	SI	−0.10	−0.05	−0.04	−0.43	−0.16	−0.34	0.41	0.41	0.51	0.86	0.83	0.89
		(0.04)	(0.03)	(0.05)	(0.11)	(0.07)	(0.10)	(0.02)	(0.01)	(0.03)	(0.07)	(0.05)	(0.07)
	SV	−0.06	−0.04	0.00	−0.21	−0.15	−0.03	0.38	0.39	0.43	0.84	0.81	0.72
		(0.03)	(0.02)	(0.03)	(0.09)	(0.05)	(0.06)	(0.02)	(0.02)	(0.02)	(0.05)	(0.03)	(0.04)
	SVSI	0.06	−0.04	−0.01	−0.29	−0.18	−0.17	0.41	0.42	0.45	0.91	0.85	0.76
		(0.03)	(0.02)	(0.03)	(0.08)	(0.05)	(0.06)	(0.02)	(0.01)	(0.02)	(0.05)	(0.03)	(0.04)
Panel C: in-the-money options													
1.03–1.05	BS	−0.06	−0.03	−0.05	−0.36	−0.09	−0.23	0.40	0.38	0.47	0.70	0.69	0.90
		(0.02)	(0.02)	(0.03)	(0.05)	(0.04)	(0.08)	(0.02)	(0.01)	(0.02)	(0.03)	(0.03)	(0.05)
	SI	−0.08	−0.03	−0.04	−0.47	−0.11	−0.52	0.43	0.40	0.48	0.84	0.77	0.97
		(0.03)	(0.03)	(0.05)	(0.08)	(0.06)	(0.09)	(0.02)	(0.02)	(0.03)	(0.05)	(0.04)	(0.07)
	SV	−0.04	−0.02	−0.07	−0.23	−0.09	−0.23	0.41	0.39	0.46	0.70	0.76	0.88
		(0.03)	(0.02)	(0.04)	(0.06)	(0.05)	(0.08)	(0.02)	(0.01)	(0.02)	(0.03)	(0.03)	(0.05)
	SVSI	−0.05	−0.02	−0.04	−0.27	−0.14	−0.31	0.41	0.40	0.47	0.77	0.80	0.90
		(0.02)	(0.02)	(0.03)	(0.05)	(0.05)	(0.08)	(0.01)	(0.01)	(0.02)	(0.03)	(0.03)	(0.05)

1.05–1.07	BS	−0.05	−0.02	−0.06	−0.35	−0.06	−0.22	0.41	0.40	0.47	0.68	0.64	0.77
		(0.02)	(0.02)	(0.04)	(0.04)	(0.04)	(0.07)	(0.01)	(0.01)	(0.02)	(0.02)	(0.03)	(0.04)
	SI	−0.07	−0.04	−0.07	−0.37	−0.09	−0.55	0.45	0.42	0.50	0.74	0.66	0.81
		(0.03)	(0.03)	(0.05)	(0.05)	(0.06)	(0.08)	(0.02)	(0.02)	(0.03)	(0.04)	(0.04)	(0.06)
	SV	−0.06	−0.03	−0.05	−0.32	−0.09	−0.10	0.43	0.40	0.44	0.71	0.68	0.69
		(0.02)	(0.02)	(0.04)	(0.04)	(0.04)	(0.07)	(0.02)	(0.01)	(0.03)	(0.02)	(0.03)	(0.04)
	SVSI	−0.05	−0.02	−0.02	−0.31	−0.09	−0.09	0.42	0.43	0.46	0.74	0.79	0.75
		(0.02)	(0.02)	(0.04)	(0.04)	(0.04)	(0.07)	(0.01)	(0.01)	(0.02)	(0.03)	(0.03)	(0.05)
>1.07	BS	−0.04	−0.03	−0.02	−0.15	−0.07	−0.10	0.36	0.39	0.48	0.51	0.58	0.72
		(0.01)	(0.00)	(0.01)	(0.02)	(0.02)	(0.03)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
	SI	−0.05	−0.04	−0.04	−0.17	−0.07	−0.26	0.40	0.41	0.47	0.61	0.66	0.79
		(0.01)	(0.01)	(0.02)	(0.03)	(0.03)	(0.03)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.02)
	SV	−0.04	−0.04	−0.02	−0.18	−0.14	−0.09	0.35	0.39	0.44	0.50	0.58	0.64
		(0.04)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)
	SVSI	−0.04	−0.03	−0.01	−0.18	−0.14	−0.10	0.36	0.41	0.46	0.50	0.63	0.70
		(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)

For each call option, calculate the hedging error, which is the difference between the market price of the call and the replicating portfolio. The average dollar hedging error and the average absolute hedging error are reported for each model. The standard errors are shown in parentheses. The sample period is 06:1988–05:1991. In calculating the hedging errors generated with daily (once every 5 days) hedge rebalancing, 15,041 (11,704) observations are used the SV model has, regardless of the hedging error measure used and the hedge revision frequency adopted, the lowest hedging errors, followed by the SVSI, then by the BS, and lastly by the SI model. For ATM options, the hedging performance by the BS and the SV models is almost indistinguishable, but still better, by a small margin, than that by both the SI and the SVSI models, whereas the latter two models' performance is also indistinguishable. Finally, for ITM options, the BS model has the best hedging performance, followed by the SV, the SVSI, and then by the SI model. Having said the above, it should nonetheless be noted that for virtually all cases in Table 98.5, the hedging error differences among the BS, the SV, and the SVSI models are economically insignificant because of their low magnitude. Only the SI model's performance appears to be significantly poorer than the others'.

The fact that the SI model performs worse than the BS and that the SVSI model performs slightly worse than the SV suggests that adding stochastic interest rates to the option pricing framework actually make the single-instrument hedge's performance worse. This can be explained as follows. In the setup of the chapter, interest rate shocks are assumed to be independent of shocks to the stock price and/or to the stochastic volatility. Therefore, in the single-instrument minimum-variance hedges, there is no adjustment in the optimal position in the underlying stock to be taken. The hedging results in Table 98.5 have shown that if interest rate risk is not to be controlled by any position in the hedging instrument, then it is perhaps better to design a single-instrument hedge based on an option model that assumes no interest rate risk. Assuming interest rate risk in an option pricing model and yet not controlling for this risk in a hedge can make the hedging effectiveness worse.

In the case of the SV versus the BS model, the situation is somewhat different from the above. As volatility shocks are assumed to be correlated with stock price shocks, the position to be taken in the underlying stock (i.e., the hedging instrument) needs to be adjusted relative to the BS model-determined hedge, so that this single position not only helps contain the underlying stock's price risk but also neutralize that part of volatility risk which is related to stock price fluctuations (see Eq. 98.23). Thus, by rendering it possible to use the single hedging position to control for both stock price risk and volatility risk, introducing stochastic volatility into the BS framework helps improve the single-instrument hedging performance, albeit by a small amount. Nandi (1996) uses the remaining variance of a hedged position as a hedging effectiveness measure, according to which he finds the SV model performs better than the BS model. Our single-instrument hedging results are hence consistent with his, regarding the SV versus the BS model.

It is useful to recall that all four models are implemented allowing both the spot volatility and the spot interest rates to vary from day to day, which is, except in the sole case of the SVSI model, not consistent with the models' assumptions. Given this practical ad hoc treatment, it may not come as a surprise that when only the underlying asset is used as the hedging instrument, the four models performed virtually indifferently, with the magnitude of their hedging error differences being generally small. As easily seen, if all four models were implemented in a way consistent with the respective model setups, the single-instrument hedges based on the SVSI model would for sure perform the best.

Comparing Tables 98.4 and 98.5, one can conclude that based on a given option model, the conventional delta-neutral hedges perform far better than their single-instrument counterparts, for every moneyness–maturity category. This is not surprising as the former type of hedges involves more hedging instruments (except under the BS model).

98.4.3 Regression Analysis of Option Pricing and Hedging Errors

So far we have examined pricing and hedging performance according to option moneyness–maturity categories. The purpose was to see whether the errors have clear moneyness- and maturity-related biases. By appealing to a regression analysis, we can more rigorously study the association of the errors with factors that are either contract specific or market condition dependent. Fix an option pricing model, and let $\epsilon_n(t)$ denote the n-th call option's percentage pricing error on day t. Then, run the regression below for the entire sample:

$$\begin{aligned}\epsilon_n(t) = \beta_0 + \beta_1 \frac{S(t)}{K_n} + \beta_2\tau_n + \beta_3 SPREAD_n(t)\\ +\beta_5 LAGVOL(t-1) + \beta_4 SLOPE(t) + \eta_n(t),\end{aligned} \tag{98.29}$$

where K_n is the strike price of the call, τ_n the remaining time to expiration, and $SPREAD_n(t)$ the percentage bid–ask spread at date t of the call (constructed by computing $\frac{Ask-Bid}{0.5(Ask+Bid)}$, all of which are contract-specific variables. The variable, $LAGVOL(t-1)$, is the (annualized) standard deviation of the previous day's intraday S&P 500 returns computed over 5-min intervals, and it is included in the regression to see whether the previous day's volatility of the underlying may cause systematic pricing biases. The variable, $SLOPE(t)$, represents the yield differential between 1-year and 30-day Treasury bills. This variable can provide information on whether the single-factor Cox et al. (1985) term structure model assumed in the chapter is sufficient to make the resulting option formula capture all term structure-related effects on the S&P 500 index options. In some sense, the contract-specific variables help detect the existence of cross-sectional pricing biases, whereas $LAGVOL(t-1)$ and $SLOPE(t)$ serve to indicate whether the pricing errors over time are related to the dynamically changing market conditions. Similar regression analyses have been done for the BS pricing errors in, for example, Galai (1983b), George and Longstaff (1993), and Madan et al. (1998). For each given option model, the same regression as in Eq. 98.29 is also run for the conventional delta-neutral hedging errors, with $\epsilon_n(t)$ in Eq. 98.29 replaced by the dollar hedging error for the n-th option on day t.

Table 98.6 reports the regression results based on the entire sample period, where the standard error for each coefficient estimate is adjusted according to the White (1980) heteroskedasticity-consistent estimator and is given in the parentheses. Let us first examine the pricing error regressions. For every option

Table 98.6 Regression analysis of pricing and hedging errors

	Percentage pricing errors				Hedging errors			
Coefficient	BS	SI	SV	SVSI	BS	SI	SV	SVSI
Constant	−0.05	0.28	0.24	0.11	−0.41	−0.30	0.00	−0.03
	(0.03)	(0.03)	(0.02)	(0.02)	(0.11)	(0.10)	(0.05)	(0.05)
S/K	0.22	−0.18	−0.20	−0.09	0.34	0.29	0.00	0.03
	(0.03)	(0.02)	(0.01)	(0.02)	(0.09)	(0.08)	(0.04)	(0.04)
τ	−0.04	0.04	0.08	0.05	0.03	0.08	0.00	0.00
	(0.01)	(0.01)	(0.01)	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
SPREAD	−5.24	−4.48	−2.13	−1.57	2.26	1.04	0.32	0.34
	(0.12)	(0.11)	(0.07)	(0.08)	(0.48)	(0.32)	(0.23)	(0.23)
SLOPE	2.29	1.33	0.32	0.34	−2.09	−2.01	−0.39	−0.39
	(0.16)	(0.13)	(0.08)	(0.11)	(0.58)	(0.65)	(0.26)	(0.25)
LAGVOL	−0.16	0.12	0.06	0.04	−0.31	−0.51	−0.06	−0.05
	(0.02)	(0.02)	(0.01)	(0.01)	(0.07)	(0.05)	(0.02)	(0.02)
Adj. R^2	0.29	0.22	0.12	0.07	0.01	0.01	0.00	0.00

The regression results below are based on the equation:
$\epsilon_n(t) = \beta_0 + \beta_1 \frac{S(t)}{K_n} + \beta_2\tau_n + \beta_3 SPREAD_n(t) + \beta_4 SLOPPE(t) + \beta_4 LAGVOL(t-1) + \eta_n(t)$,
where $\epsilon_n(t)$ denotes either the percentage pricing error or the dollar hedging error of the n-th call on date t; S/K and τ_n, respectively, represent the moneyness and the term to expiration of the option contract; the variable $SPREAD_n(t)$ is the percentage bid–ask spread; $SLOPE(t)$ the yield differential between the 1-year and the 30-day Treasury bill rates; and $LAGVOL(t-1)$ the previous day's (annualized) standard deviation of S&P 500 index returns computed from 5-min intradaily returns. The standard errors, reported in parenthesis, are White's (1980) heteroskedastically consistent estimator. The sample period is 06:1988–05:1991 for a total of 38,749 observations

model, each independent variable has statistically significant explanatory power of the remaining pricing errors. That is, the pricing errors from each model have some moneyness, maturity, intraday volatility, bid–ask spread, and term structure-related biases. The magnitude of each such bias, however, decreases from the BS to the SI, to the SV, and to the SVSI model. For instance, the BS percentage pricing errors will on the average be 2.29 points higher when the yield spread $SLOPE(t)$ increases by one point, whereas the SV and the SVSI percentage errors will only be, respectively, 0.32 and 0.34 points higher in response. Thus, a higher yield spread on the term structure means higher pricing errors, regardless of the option model used. This points out that *a possible direction to further improve pricing performance is to include the yield spread as a second factor in the term structure model of interest rates*. Other noticeable patterns include the following. The BS pricing errors are decreasing, while the SI, the SV, and the SVSI pricing errors are increasing, in both the option's time to expiration and the underlying stock's volatility on the previous day. The deeper in the money the call or the wider its bid–ask spread, the lower the SI's, the SV's, and the SVSI model's mispricing. But, for the BS model, its mispricing increases with moneyness and decreases with bid–ask spread.

Even though all four models' pricing errors are significantly related to each independent variable, the collective explanatory power of these variables is not so impressive. The adjusted R^2 is 29 % for the BS formula's pricing errors, 22 % for the SI's, 12 % for the SV's, and 7 % for the SVSI model's. Therefore, while both the BS and the SI model have significant overall biases related to contract terms and market conditions (indicating systematic model misspecifications), the remaining pricing errors under the SV and the SVSI are not as significantly associated with these variables. About 93 % of the SVSI model's pricing errors cannot be explained by these variables!

As reported in Table 98.6, delta-neutral hedging errors by the BS and the SI model tend to increase with the moneyness and the bid–ask spread of the target call, but decrease with the non-contract-specific yield spread and lagged stock volatility variables. Therefore, the two models are misspecified for hedging purposes, and they lead to systematic hedging biases. But, overall, these variables can explain only 1 % of the hedging errors by the two models. And, even more impressively, none of the included independent variables can explain any of the remaining hedging errors by the SV and the SVSI model, as their R^2 values are both zero.

Finally, when the dollar pricing errors are used to replace the percentage pricing errors or when the percentage hedging errors are employed to replace the dollar hedging errors in the above regressions, the sign of each resulting coefficient estimate and the magnitude of each R^2 value in Table 98.6 remain unchanged. Thus, the conclusions drawn from Table 98.6 are independent of the choice of the pricing or hedging error measure. Results from these exercises are not reported here but available upon request.

98.4.4 Robustness of Empirical Results

Using the entire sample period data, we have concluded that the evidence, based on both static performance and dynamic performance measures, is in favor of both the SVSI and the SV model. However, it is important to demonstrate that this conclusion still holds when alternative test designs and different sample periods are used. Below we briefly report results from two controlled experiments.

According to Rubinstein (1985), the volatility smile pattern and the nature of pricing biases are time period dependent. To see whether our conclusion may be reversed, we separately examined the pricing and hedging performance of the models in three subperiods: 06:1988–05:1989, 06:1989–05:1990, and 06:1990–05:1991. Each subperiod contains about 10,000 call option observations. As the results are similar for each subperiod, we provide the percentage pricing errors in Panel A and the absolute delta-neutral hedging errors in Panel B of Table 98.7, for the subperiod 06:1990–05:1991. It is seen that these results are qualitatively the same as those in Tables 98.3 and 98.4.

We examined the pricing and hedging error measures of each model when the structural parameters were not updated daily. Rather, retain the structural parameter values estimated from the options of the first day of each month, and then, for the remainder of the month, use them as input to compute the corresponding model-based price for each traded option, except that the implied spot volatility

Table 98.7 Robustness analysis

Panel A: percentage pricing errors, 06:1990–05:1991							
Moneyness	Model	Term to expiration (days)					
S/K		<30	30–60	60–90	90–120	120–180	≥180
<0.93	BS	−76.51	−96.75	−74.73	−78.71	−61.36	−46.83
	SI	−26.88	−62.89	−38.89	−34.40	−18.68	−5.14
	SV	−25.05	−35.80	−12.59	−1.25	1.64	8.51
	SVSI	−20.56	−31.82	−8.15	−3.00	0.58	3.57
0.93–0.95	BS	−54.99	−59.06	−31.97	−24.29	−19.28	−12.36
	SI	−46.25	−46.77	−28.99	−10.62	−9.55	−5.37
	SV	−26.57	−25.32	−8.57	−1.60	−0.33	1.50
	SVSI	−28.25	−21.71	−6.26	−2.18	−0.04	0.79
0.95–0.97	BS	−34.72	−29.97	−16.71	−12.74	−10.27	−7.33
	SI	−31.85	−24.18	−16.79	−4.92	−5.35	−4.49
	SV	−20.09	−13.89	−5.54	−1.68	−0.56	0.61
	SVSI	−15.83	−13.08	−4.20	−3.29	−0.91	0.18
0.97–0.99	BS	−15.93	−10.36	−5.84	−3.22	−3.46	−2.14
	SI	−15.27	−7.45	−7.22	2.36	−2.05	−1.24
	SV	−13.09	−7.04	−3.64	0.77	−0.47	0.15
	SVSI	−12.38	−7.49	−3.37	−0.90	−0.83	−0.33
0.99–1.01	BS	−3.92	−1.23	0.62	1.54	0.65	1.99
	SI	−3.24	−0.09	−0.87	5.22	0.38	−0.09
	SV	−6.69	−3.43	−1.23	1.22	−0.10	0.14
	SVSI	−7.46	−4.17	−1.49	−0.33	−0.48	−0.27
1.01–1.03	BS	2.48	3.36	4.17	4.05	3.28	2.93
	SI	2.73	3.75	2.78	5.57	1.82	−0.60
	SV	−0.92	−0.56	0.24	1.44	0.19	−0.24
	SVSI	−1.41	−1.02	−0.09	−0.59	−0.16	−0.53
1.03–1.05	BS	3.93	4.86	5.35	5.57	4.62	3.41
	SI	3.95	4.92	4.08	6.62	2.61	−0.08
	SV	1.21	0.74	0.82	1.87	0.39	−0.17
	SVSI	0.85	0.19	0.48	0.79	0.08	−0.40
1.05–1.07	BS	3.69	5.07	5.84	6.36	5.12	4.93
	SI	3.82	5.08	4.67	6.55	2.87	1.29
	SV	1.83	1.52	1.49	2.07	0.51	−0.54
	SVSI	1.68	1.18	1.17	1.38	0.32	−0.57
>1.07	BS	1.99	2.98	3.58	4.35	4.00	3.59
	SI	2.44	2.90	2.77	4.08	1.93	−1.01
	SV	1.43	1.40	1.21	1.47	0.45	−0.69
	SVSI	1.38	1.30	1.18	1.18	0.46	−0.40

Panel B: absolute hedging errors (1 and 5 days), 06:1990–05:1991							
		1-day revision					
Moneyness	Model	Term to expiration (days)			5-day revision		
S/K		<60	60–180	>180	<60	60–180	>180
<0.93	BS	NA	0.42	0.48	NA	1.13	0.99

(*continued*)

Table 98.7 (continued)

Panel B: absolute hedging errors (1 and 5 days), 06:1990–05:1991							
		1-day revision					
Moneyness	Model	Term to expiration (days)			5-day revision		
S/K		<60	60–180	>180	<60	60–180	>180
	SI		0.46	0.45		0.77	0.82
	SV		0.17	0.22		0.18	0.30
	SVSI		0.17	0.22		0.18	0.30
0.93–0.95	BS	NA	0.40	0.50	NA	0.97	0.95
	SI		0.45	0.47		0.73	0.75
	SV		0.13	0.25		0.15	0.30
	SVSI		0.13	0.25		0.15	0.30
0.95–0.97	BS	NA	0.37	0.44	NA	0.96	0.85
	SI		0.45	0.44		0.77	0.86
	SV		0.16	0.22		0.16	0.29
	SVSI		0.16	0.22		0.16	0.29
0.97–0.99	BS	0.39	0.42	0.47	0.72	0.95	0.97
	SI	0.33	0.45	0.41	0.66	0.74	0.76
	SV	0.14	0.17	0.17	0.16	0.17	0.24
	SVSI	0.14	0.17	0.16	0.15	0.17	0.23
0.99–1.01	BS	0.41	0.43	0.50	0.99	0.91	0.89
	SI	0.40	0.48	0.50	0.79	0.71	0.78
	SV	0.16	0.16	0.17	0.20	0.17	0.28
	SVSI	0.16	0.16	0.17	0.19	0.17	0.26
1.01–1.03	BS	0.40	0.46	0.47	0.99	0.89	0.83
	SI	0.45	0.44	0.45	0.74	0.71	0.73
	SV	0.17	0.17	0.18	0.19	0.20	0.25
	SVSI	0.17	0.17	0.17	0.19	0.20	0.25
1.03–1.05	BS	0.45	0.43	0.50	0.88	0.85	0.97
	SI	0.46	0.44	0.48	0.71	0.72	0.68
	SV	0.17	0.14	0.17	0.18	0.16	0.27
	SVSI	0.17	0.14	0.17	0.18	0.16	0.27
1.05–1.07	BS	0.46	0.47	0.51	0.73	0.78	0.77
	SI	0.47	0.45	0.50	0.61	0.67	0.68
	SV	0.18	0.14	0.22	0.19	0.16	0.24
	SVSI	0.17	0.14	0.21	0.19	0.16	0.22
>1.07	BS	0.41	0.45	0.53	0.62	0.70	0.81
	SI	0.38	0.46	0.50	0.48	0.64	0.75
	SV	0.17	0.15	0.22	0.18	0.19	0.32
	SVSI	0.16	0.15	0.21	0.18	0.18	0.31

The reported percentage pricing error is the sample average of the market price minus the model price divided by the market price. The sample period is 06:1990–05:1991 for a total of 11,979 call options

The average absolute hedging error for each model is reported based on the subsample period 06:1990–05:1991 (with a total of 6,440 observations) is updated each day based on the previous day's option prices. The obtained absolute pricing errors for the subperiod 06:1990–05:1991 indicate that the performance ranking of the four models remains the same as before.

In addition, when we used only ATM (or only ITM or only OTM) option prices to back out each model's parameter values, the resulting pricing and hedging errors did not change the performance ranking of the models either. This means that even if one would estimate and use a matrix of implied volatilities (across moneynesses and maturities) to accordingly price and hedge options in different moneyness–maturity categories, it would still not change the fact that the SV and the SVSI models are better specified than the other two for pricing and hedging. Given that the implied-volatility matrix method has gained some popularity among practitioners, our results should be appealing. On the one hand, they suggest that with the SV and the SVSI models, there is far less a need to engage in moneyness- and maturity-related fitting. On the other hand, if one is still interested in the matrix method, the SV and the SVSI models should be better model choices.

Early in the project we used only option transaction price data for the pricing and hedging estimations. But, that meant a far smaller data set, especially for the hedging estimations. Nonetheless, the results obtained from the transaction prices were similar to these presented and discussed in this chapter.

98.5 Concluding Remarks

We have developed and analyzed a simple option pricing model that admits both stochastic volatility and stochastic interest rates. It is shown that this closed-form pricing formula is practically implementable, leads to useful analytical hedge ratios, and contains many known option formulas as special cases. This last feature has made it relatively straightforward to conduct a comparative empirical study of the four classes of option pricing models.

According to the pricing and hedging performance measures, the SVSI and the SV models both perform much better than the SI and the BS models, as the former typically reduce the pricing and hedging errors of the latter by more than a half. These error reductions are also economically significant. Furthermore, the hedging errors by the SV and the SVSI models are relatively insensitive to the frequency of portfolio revision, whereas those of the SI and the BS models are sensitive. Given that both the SV and the SVSI models can be easily implemented on a personal computer, they should thus be better alternatives to the widely applied BS formula. A regression-based analysis of the pricing and hedging errors indicates that while the BS and the SI models show significant pricing biases related to moneyness, time to expiration, bid–ask spread, lagged stock volatility, and interest rate term spread, pricing errors by the SV and the SVSI models are not as systematically related to either contract-specific or market-dependent variables. Overall, the results lend empirical support to the claim that incorporating stochastic interest rates and, especially, stochastic volatility can both improve option pricing and hedging performance substantially and resolve some known empirical biases associated with the BS model.

The empirical issues and questions addressed in this chapter can also be reexamined using data from individual stock options, American-style index options, options on futures, currency and commodity options, and so on. Eventually, the acceptability of option pricing models with the added features will be judged not only by its easy implementability or even its impressive pricing and hedging performance as demonstrated in this chapter using European-style index calls but also by its success or failure in pricing and hedging other types of options. These extensions are left for future research.

Appendix 1

Proof of the Option Pricing Formula in Eq. 98.10. The valuation PDE in Eq. 98.9 can be rewritten as

$$\begin{aligned}&\frac{1}{2}\frac{\partial^2 C}{\partial L^2}+\left(R-\frac{1}{2}V\right)\frac{\partial C}{\partial L}+\rho\sigma_v V\frac{\partial^2 C}{\partial L\partial V}+\frac{1}{2}\sigma_v^2 V\frac{\partial^2 C}{\partial V^2}\\&+[\theta_v-\kappa_v V]\frac{\partial C}{\partial V}+\frac{1}{2}\sigma_R^2 R\frac{\partial^2 C}{\partial R^2}+[\theta_R-\kappa_R R]\frac{\partial C}{\partial R}-\frac{\partial C}{\partial \tau}-RC=0,\end{aligned} \tag{98.30}$$

where we have applied the transformation $L(t) \equiv \ln[S(t)]$. Inserting the conjectured solution in Eq. 98.10 into Eq. 98.30 produces the PDEs for the risk-neutralized probabilities, Π_j for $j = 1, 2$:

$$\begin{aligned}&\frac{1}{2}\frac{\partial^2 \prod_1}{\partial L^2}+\left(R+\frac{1}{2}V\right)\frac{\partial \prod_1}{\partial L}+\rho\sigma_v V\frac{\partial^2 \prod_1}{\partial L\partial V}+\frac{1}{2}\sigma_v^2 V\frac{\partial^2 \prod_1}{\partial V^2}\\&+[\theta_v-(\kappa_v-\rho\sigma_v)V]\frac{\partial \prod_1}{\partial V}+\frac{1}{2}\sigma_R^2 R\frac{\partial^2 \prod_1}{\partial R^2}+[\theta_R-\kappa_R R]\frac{\partial \prod_1}{\partial R}-\frac{\partial \prod_1}{\partial \tau}=0,\end{aligned} \tag{98.31}$$

and

$$\begin{aligned}&\frac{1}{2}\frac{\partial^2 \prod_2}{\partial L^2}+\left(R-\frac{1}{2}V\right)\frac{\partial \prod_2}{\partial L}+\rho\sigma_v V\frac{\partial^2 \prod_2}{\partial L\partial V}+\frac{1}{2}\sigma_v^2 V\frac{\partial^2 \prod_2}{\partial V^2}+[\theta_v-\kappa_v V]\frac{\partial \prod_2}{\partial V}\\&+\frac{1}{2}\sigma_R^2 R\frac{\partial^2 \prod_2}{\partial R^2}+\left[\theta_R-\left(\kappa_R-\frac{\sigma_R^2}{B(t,\tau)}\frac{\partial B(t,\tau)}{\partial R}\right)R\right]\frac{\partial \prod_2}{\partial R}-\frac{\partial \prod_2}{\partial \tau}=0.\end{aligned} \tag{98.32}$$

Observe that Eqs. 98.31 and 98.32 are the Fokker–Planck forward equations for probability functions. This implies that Π_1 and Π_2 must indeed be valid probability functions, with values bounded between 0 and 1. These PDEs must be separately solved subject to the terminal condition:

$$\prod_j (t+\tau, 0) = 1_{L(t+\tau)\geq K}\; j = 1, 2. \tag{98.33}$$

The corresponding characteristic functions for Π_1 and Π_2 will also satisfy similar PDEs:

$$\begin{aligned}&\frac{1}{2}\frac{\partial^2 f_1}{\partial L^2} + \left(R + \frac{1}{2}V\right)\frac{\partial f_1}{\partial L} + \rho\sigma_v V\frac{\partial^2 f_1}{\partial L \partial V} + \frac{1}{2}\sigma_v^2 V\frac{\partial^2 f_1}{\partial V^2}\\ &+[\theta_v - (\kappa_v - \rho\sigma_v)V]\frac{\partial f_1}{\partial V} + \frac{1}{2}\sigma_R^2 R\frac{\partial^2 f_1}{\partial R^2} + [\theta_R - \kappa_R R]\frac{\partial f_1}{\partial R} - \frac{\partial f_1}{\partial \tau} = 0,\end{aligned} \tag{98.34}$$

and

$$\begin{aligned}&\frac{1}{2}\frac{\partial^2 f_2}{\partial L^2} + \left(R - \frac{1}{2}V\right)\frac{\partial f_2}{\partial L} + \rho\sigma_v V\frac{\partial^2 f_2}{\partial L \partial V} + \frac{1}{2}\sigma_v^2 V\frac{\partial^2 f_2}{\partial V^2} + [\theta_v - \kappa_v V]\frac{\partial f_2}{\partial V}\\ &+\frac{1}{2}\sigma_R^2 R\frac{\partial^2 f_2}{\partial R^2} + \left[\theta_R - \left(\kappa_R - \frac{\sigma_R^2}{B(t,\tau)}\frac{\partial B(t,\tau)\partial f_2}{\partial R}\right)R\right]\frac{\partial f_2}{\partial R} - \frac{\partial f_2}{\partial \tau} = 0,\end{aligned} \tag{98.35}$$

with the boundary condition:

$$f_j(t+\tau, 0; \phi) = e^{i\phi L(t+\tau)}\; j = 1, 2. \tag{98.36}$$

Conjecture that the solution to the PDEs (98.34) and (98.35) is, respectively, given by

$$\begin{aligned}f_1(t+\tau, S(t), V(t), R(t); \phi) = \exp\{&u_r(\tau) + u_v(\tau) + x_r(\tau) + R(t)\\ &+x_v(\tau)V(t) + i\phi\ln[S(t)]\}\end{aligned} \tag{98.37}$$

$$\begin{aligned}f_2(t, \tau, S(t), V(t), R(t); \phi) = \exp\{&z_r(\tau) + z_v(\tau) + y_r(\tau)\,R(t)\\ &+y_v(\tau)V(t) + i\phi\ln[S(t)] - \ln[B(t,\tau)]\}\end{aligned} \tag{98.38}$$

with $u_r(0) = u_v(0) = x_r(0) = x_v(0) = 0$ and $z_r(0) = z_v(0) = y_r(0) = y_v(0) = 0$. Solving the resulting system of differential equations and noting that $B(t + \tau, 0) = 1$ will respectively produce the desired characteristic functions in Eqs. 98.12 and 98.13.

Both the constant interest rate–stochastic volatility and constant volatility–stochastic interest rate option pricing models are nested in Eq. 98.10. In the constant interest rate–stochastic volatility model, for instance, the partial derivatives with respect to R vanish in Eq. 98.30. The general solution in Eqs. 98.37, 98.38, and 98.40 will still apply except that now $R(t) = R$ (a constant), $B(t, \tau) = e^{-R\tau}$, $x_r(\tau) = i\phi\tau$, $y_r(\tau) = (i\phi - 1)\tau$, and $u_r(\tau) = z_r(\tau) = 0$. The final characteristic functions $\hat{f}_j$ for the constant interest rate–stochastic volatility option model are respectively given by

$$\hat{f}_1 = \exp\left\{-i\phi\ln[B(t,\tau)] - \frac{\theta_v}{\sigma_v^2}\left[2\ln\left(1 - \frac{[\xi_v - \kappa_v + (1+i\phi)\rho\sigma_v]\left(1-e^{-\xi_v\tau}\right)}{2\xi_v}\right)\right]\right.$$
$$-\frac{\theta_v}{\sigma_v^2}[\xi_v - \kappa_v + (1+i\phi)\rho\sigma_v]\tau + i\phi\ln[S(t)]$$
$$\left.+\frac{i\phi(i\phi+1)\left(1-e^{-\xi_v\tau}\right)}{2\xi_v - [\xi_v - \kappa_v + (1+i\phi)\rho\sigma_v](1-e^{-\xi_v\tau})}V(t)\right\}, \quad (98.39)$$

and

$$\hat{f}_2 = \exp\left\{-i\phi\ln[B(t,\tau)] - \frac{\theta_v}{\sigma_v^2}\left[2\ln\left(1 - \frac{\left[\xi_v^* - \kappa_v + i\phi\rho\sigma_v\right]\left(1-e^{-\xi_v^*\tau}\right)}{2\xi_v^*}\right)\right]\right.$$
$$\left.-\frac{\theta_v}{\sigma_v^2}\left[\xi_v^* - \kappa_v + i\phi\rho\sigma_v\right]\tau + i\phi\ln[S(t)] + \frac{i\phi(i\phi+1)\left(1-e^{-\xi_v^*\tau}\right)}{2\xi_v^* - \left[\xi_v^* - \kappa_v + i\phi\rho\sigma_v\right]\left(1-e^{-\xi_v^*\tau}\right)}V(t)\right\}, \quad (98.40)$$

Similarly, the constant volatility–stochastic interest rate option model obtains with $V(t) = V$ (a constant), $x_v(\tau) = \frac{1}{2}i\phi(1+i\phi)\tau$, $y_v(\tau) = \frac{1}{2}i\phi(i\phi-1)\tau$, and $u_v(\tau) = z_v(\tau) = 0$. The final characteristic functions $\widetilde{f}_j$ for the stochastic interest rate–constant volatility model are

$$\widetilde{f}_1 = \exp\left\{\frac{1}{2}i\phi(1+i\phi)V\tau + i\phi\ln[S(t)]\right.$$
$$-\frac{\theta_R}{\sigma_R^2}\left[2\ln\left(1 - \frac{[\xi_R - \kappa_R]\left(1-e^{-\xi_R\tau}\right)}{2\xi_R}\right) + [\xi_R - \kappa_R]\tau\right]$$
$$\left.+\frac{2i\phi\left(1-e^{-\xi_R\tau}\right)}{2\xi_R - [\xi_R - \kappa_R](1-e^{-\xi_R\tau})}R(t)\right\}, \quad (98.41)$$

and

$$\widetilde{f}_2 = \exp\left\{\frac{1}{2}i\phi(i\phi-1)V\tau + i\phi\ln[S(t)] - \ln[B(t,\tau)]\right.$$
$$-\frac{\theta_R}{\sigma_R^2}\left[2\ln\left(1 - \frac{\left[\xi_R^* - \kappa_R\right]\left(1-e^{-\xi_R^*\tau}\right)}{2\xi_R^*}\right) + \left[\xi_R^* - \kappa_R\right]\tau\right], \quad (98.42)$$
$$\left.+\frac{2(i\phi-1)\left(1-e^{-\xi_R^*\tau}\right)}{2\xi_R^* - \left[\xi_R^* - \kappa_R\right]\left(1-e^{-\xi_R^*\tau}\right)}R(t)\right\}$$

Expressions for the Gamma Measures. The various second-order partial derivatives of the call price in Eq. 98.10, which are commonly referred to as Gamma measures, are given below:

$$\Gamma_S \equiv \frac{\partial^2 C(t,\tau)}{\partial S^2} = \frac{\partial \prod_1}{\partial S} = \frac{1}{\pi}\int_0^\infty Re\left[(i\phi)^{-1} e^{-i\phi \ln[K]} f_1 \frac{i\phi}{S}\right] d\phi. > 0. \tag{98.43}$$

$$\Gamma_V \equiv \frac{\partial^2 C(t,\tau)}{\partial V^2} = S(t)\frac{\partial^2 \prod_1}{\partial V^2} - KB(t,\tau)\frac{\partial^2 \prod_2}{\partial V^2} \tag{98.44}$$

$$\Gamma_R \equiv \frac{\partial^2 C(t,\tau)}{\partial R^2} = S(t)\frac{\partial^2 \prod_1}{\partial R^2} - KB(t,\tau)\left\{\frac{\partial^2 \prod_2}{\partial R^2} 2\varrho(\tau)\frac{\partial \prod_2}{\partial R} + \varrho^2(\tau)\prod\nolimits_2\right\}. \tag{98.45}$$

$$\Gamma_{S,V} \equiv \frac{\partial^2 C(t,\tau)}{\partial S \partial V} = \frac{\partial \prod_1}{\partial V} = \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[(i\phi)^{-1} e^{-i\phi \ln[K]} \frac{\partial^2 f_1}{\partial V}\right] d\phi. \tag{98.46}$$

where for $g = V, R$ and $j = 1, 2$

$$\frac{\partial^2 \prod_j}{\partial g^2} = \frac{1}{\pi}\int_0^\infty \mathrm{Re}\left[(i\phi)^{-1} e^{-i\phi \ln[K]} \frac{\partial^2 f_j}{\partial g^2}\right] d\phi. \tag{98.47}$$

Acknowledgments We would like to thank Sanjiv Das, Ranjan D'Mello, Helyette Geman, Eric Ghysels, Frank Hatheway, Steward Hodges, Ravi Jagannathan, Andrew Karolyi, Bill Kracaw, C. F. Lee, Dilip Madan, Louis Scott, René Stulz, Stephen Taylor, Siegfried Trautmann, Alex Triantis, and Alan White for their helpful suggestions. Any remaining errors are our responsibility alone.

References

Amin, K., & Jarrow, R. (1992). Pricing options on risky assets in a stochastic interest rate economy. *Mathematical Finance, 2*, 217–237.

Amin, K., & Ng, V. (1993). Option valuation with systematic stochastic volatility. *Journal of Finance, 48*, 881–910.

Andersen, T., & Lund, J. (1997). Estimating continuous time stochastic volatility models of the short term interest rate. *Journal of Econometrics, 77*, 343–377.

Bailey, W., & Stulz, R. (1989). The pricing of stock index options in a general equilibrium model. *Journal of Financial and Quantitative Analysis, 24*, 1–12.

Bakshi, G., Cao, C., & Chen, Z. (1997). Empirical performance of alternative option pricing models. *Journal of Finance, 52*, 2003–2049.

Bakshi, G., Cao, C., & Chen, Z. (2000a). Do call prices and the underlying stock always move in the same direction? *Review of Financial Studies, 13*, 549–584.

Bakshi, G., Cao, C., & Chen, Z. (2000b). Pricing and hedging long-term options. *Journal of Econometrics, 94*, 277–318.

Bakshi, G., & Chen, Z. (1997). An alternative valuation model for contingent claims. *Journal of Financial Economics, 44*, 123–165.

Barone-Adesi, G., & Whaley, R. (1987). Efficient analytic approximation of American option values. *Journal of Finance, 42*, 301–320.
Bates, D. (1996a). Testing option pricing models. In G. S. Maddala & C. R. Rao (Eds.), *Statistical methods in finance* (Handbook of statistics, Vol. 14, pp. 567–611). Amsterdam: Elsevier.
Bates, D. (1996b). Jumps and stochastic volatility: Exchange rate processes implicit in Deutschemark options. *Review of Financial Studies, 9*(1), 69–108.
Bates, D. (2000). Post-87 crash fears in S&P 500 futures options. *Journal of Econometrics, 94*, 181–238.
Black, F. (1975). Fact and fantasy in the use of options. *Financial Analyst Journal, 31*, 899–908.
Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy, 81*, 637–659.
Cao, C., & Huang, J. (2008). Determinants of S&P 500 index option returns. *Review of Derivatives Research, 10*, 1–38.
Chan, K., Karolyi, A., Longstaff, F., & Sanders, A. (1992). An empirical comparison of alternative models of the short- term interest rate. *Journal of Finance, 47*, 1209–1227.
Cox, J., Ingersoll, J., & Ross, S. (1985). A theory of the term structure of interest rates. *Econometrica, 53*, 385–408.
Cox, J., & Ross, S. (1976). The valuation of options for alternative stochastic processes. *Journal of Financial Economics, 3*, 145–166.
Dumas, B., Fleming, J., & Whaley, R. (1998). Implied volatility functions: Empirical tests. *Journal of Finance, 53*(6), 2059–2106.
Figlewski, S. (1989). Option arbitrage in imperfect markets. *Journal of Finance, 44*, 1289–1311.
Galai, D. (1983a). The components of the return from hedging options against stocks. *Journal of Business, 56*, 45–54.
Galai, D. (1983b). A survey of empirical tests of option pricing models. In M. Brenner (Ed.), *Option pricing* (pp. 45–80). Lexington: Heath.
George, T., & Longstaff, F. (1993). Bid-ask spreads and trading activity in the S&P 100 index options market. *Journal of Financial and Quantitative Analysis, 28*, 381–397.
Hansen, L. (1982). Large sample properties of generalized method of moments estimators. *Econometrica, 50*, 1029–1054.
Harrison, M., & Kreps, D. (1979). Martingales and arbitrage in multiperiod securities markets. *Journal of Economic Theory, 20*, 381–408.
Harvey, C., & Whaley, R. (1992a). Market volatility and the efficiency of the S&P 100 index option market. *Journal of Financial Economics, 31*, 43–73.
Harvey, C., & Whaley, R. (1992b). Dividends and S&P 100 index option valuation. *Journal of Futures Markets, 12*, 123–137.
Heston, S. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies, 6*, 327–343.
Hull, J., & White, A. (1987a). The pricing of options with stochastic volatilities. *Journal of Finance, 42*, 281–300.
Hull, J., & White, A. (1987b). Hedging the risks from writing foreign currency options. *Journal of International Money and Finance, 6*, 131–152.
Kim, I.-J. (1990). The analytical valuation of American options. *Review of Financial Studies, 3*(4), 547–572.
Longstaff, F. (1995). Option pricing and the martingale restriction. *Review of Financial Studies, 8*(4), 1091–1124.
Madan, D., Carr, P., & Chang, E. (1998). The variance gamma process and option pricing. *European Finance Review, 2*, 79–105.
McBeth, J., & Merville, L. (1979). An empirical examination of the Black–Scholes call option pricing model. *Journal of Finance, 34*, 1173–1186.
Melino, A., & Turnbull, S. (1990). Pricing foreign currency options with stochastic volatility. *Journal of Econometrics, 45*, 239–265.
Melino, A., & Turnbull, S. (1995). Misspecification and the pricing and hedging of long-term foreign currency options. *Journal of International Money and Finance, 45*, 239–265.

Merton, R. (1973). Theory of rational option pricing. *Bell Journal of Economics, 4*, 141–183.
Nandi, S. (1996). *Pricing and hedging index options under stochastic volatility* (Working Paper). Federal Reserve Bank of Atlanta.
Ross, S. (1995). *Hedging long-run commitments: Exercises in incomplete market pricing* (Working Paper). Yale School of Management.
Rubinstein, M. (1985). Nonparametric tests of alternative option pricing models using all reported trades and quotes on the 30 most active CBOE options classes from August 23, 1976 through August 31, 1978. *Journal of Finance*, 455–480.
Rubinstein, M. (1994). Implied binomial trees. *Journal of Finance, 49*, 771–818.
Scott, L. (1987). Option pricing when the variance changes randomly: Theory, estimators, and applications. *Journal of Financial and Quantitative Analysis, 22*, 419–438.
Scott, L. (1997). Pricing stock options in a jump-diffusion model with stochastic volatility and interest rates: Application of Fourier inversion methods. *Mathematical Finance, 7*, 413–426.
Stein, E., & Stein, J. (1991). Stock price distributions with stochastic volatility. *Review of Financial Studies, 4*, 727–752.
Whaley, R. (1982). Valuation of American call options on dividend paying stocks. *Journal of Financial Economics, 10*, 29–58.
Wiggins, J. (1987). Option values under stochastic volatilities. *Journal of Financial Economics, 19*, 351–372.

The Le Châtelier Principle of the Capital Market Equilibrium

99

Chin W. Yang, Ken Hung, and Matthew D. Brigida

Contents

Abstract

This chapter purports to provide a theoretical underpinning for the problem of the Investment Company Act. The theory of the Le Chatelier principle is well known in thermodynamics. The system tends to adjust itself to a new equilibrium as far as possible. In capital market equilibrium, added constraints on portfolio investment in each stock can lead to inefficiency manifested in the right-shifting efficiency frontier. According to the empirical study, the potential loss can amount to millions of dollars coupled with a higher risk-free rate and greater transaction and information costs.

C.W. Yang (✉)
Clarion University of Pennsylvania, Clarion, PA, USA

National Chung Cheng University, Chia-yi, Taiwan
e-mail: yang@clarion.edu; yang@mail.clarion.edu; ecdycw@ccu.edu.tw

K. Hung
Division of International Banking & Finance Studies, Texas A&M International University, Laredo, TX, USA
e-mail: ken.hung@tamiu.edu

M.D. Brigida
Department of Finance, Clarion University of Pennsylvania, Clarion, PA, USA
e-mail: mbrigida@clarion.edu

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1_47,

Keywords
Markowitz model • Efficient frontiers • With constraints • Without constraints • Le Chatelier principle • Thermodynamics • Capital market equilibrium • Diversified mutual funds • Quadratic programming • Investment Company Act

99.1 Introduction

In the wake of a growing trend of deregulation in various industries (e.g., utility, banking, and airline), it has become more and more important to study the responsiveness of the market to the exogenous perturbations as the system is gradually constrained. According to the law of thermodynamics, the system tends to adjust itself to a new equilibrium by counteracting the change as far as possible. This law, the Le Chatelier principle, was applied to economics by Samuelson (1949, 1960, 1972), Silberberg (1971, 1974, 1978), and to a class of spatial equilibrium models: linear programming, fixed demand, quadratic programming, full-fledged spatial equilibrium model by Labys and Yang (1996). Recently, it has been applied to optimal taxation by Diamond and Mirrlees (2002).

According to subchapter M of the Investment Company Act of 1940, a diversified mutual fund cannot have more than 5 % of total assets invested in any single company and the acquisition of securities does not exceed 10 % of the acquired company's value. By meeting this diversification threshold, funds are considered "pass-through" entities enabling capital gains and income taxes to accrue to the fund's investors. This diversification rule, on the one hand, reduces the portfolio risk according to the fundamental result of investment theory. On the other hand, more and more researchers begin to raise questions as to the potential inefficiency arising from the Investment Company Act (see Elton and Gruber 1991; Roe 1991; Francis 1993; Kohn 1994). Further, Almazan et al. (2004) document inefficiencies from a broad set of mutual fund investment constraints. With the exception of the work by Cohen and Pogue (1967), Frost and Savarino (1988), and Loviscek and Yang (1997), there is very little evidence to refute or favor this conjecture.

Empirical findings (e.g., Loviscek and Yang 1997) suggest that over 300 growth mutual funds evaluated by Value Line show that the average weight for the company given the greatest share of a fund's assets was 4.29 %. However, the Le Chatelier principle in terms of the Investment Company Act has not been scrutinized in the literature of finance. The objective of this chapter is to investigate the Le Chatelier principle applied to the capital market equilibrium in the framework of the Markowitz portfolio selection model.

99.2 The Le Chatelier Principle of the Markowitz Model

In a portfolio of n securities, Markowitz (1952, 1956, 1959, 1990, 1991) formulated the portfolio selection model in the form of a quadratic programming as shown below:

$$\min_{x_i x_j} \upsilon = \sum\nolimits_{i \in I} \sum\nolimits_{j \in J} x_i x_j \sigma_{ij} \tag{99.1}$$

subject to

$$\sum\nolimits_{i \in I} r_i x_i \geq k \tag{99.2}$$

$$\sum\nolimits_{i \in I} x_i = 1 \tag{99.3}$$

$$x_i \geq 0 \ \ \forall \ \ i \in I \tag{99.4}$$

where
x_i = proportion of investment in security i
σii = variance of rate of return of security i
σij = covariance of rate of return of security i and j
r_i = expected rate of return of security i
k = minimum rate of return of the portfolio
I and J are sets of positive integers
The resulting Lagrange function is therefore

$$L = \upsilon + \lambda\left(k - \sum r_i x_{ij}\right) + \gamma\left(1 - \sum x_i\right) \tag{99.5}$$

The solution to the Markowitz is well known (1959). The Lagrange multiplier of constraint Eq. 99.2 assumes the usual economic interpretation: change in total risk in response to an infinitesimally small change in k while all other decision variables adjust to their new equilibrium levels, i.e., $\lambda = dv/dk$. Hence, the Lagrange multiplier is of utmost importance in determining the shape of the efficiency frontier curve in the capital market. Note that values of x_{is} are unbounded between 0 and 1 in the Markowitz model. However, in reality, the proportion of investment on each security many times cannot exceed a certain percentage to ensure adequate diversification. As the maximum investment proportion on each security decreases from 99 % to 1 %, the solution to the portfolio selection model becomes more constrained, i.e., the values of optimum xs are bounded within a narrower range as the constraint is tightened. Such impact on the objective function v is straightforward: as the system is gradually constrained, the limited freedom of optimum xs gives rise to a higher and higher risk level as k is increased. For example, if parameter k is increased gradually, the Le Chatelier principle implies that in the original Markowitz minimization system, isorisk contour has the smallest curvature to reflect the most efficient adjustment mechanism:

$$abs\left(\frac{\partial^2 \upsilon}{\partial k^2}\right) \leq abs\left(\frac{\partial^2 \upsilon^*}{\partial k^2}\right) \leq abs\left(\frac{\partial^2 \upsilon^{**}}{\partial k^2}\right) \tag{99.6}$$

where v^* and v^{**} are the objective function (total portfolio risk) corresponding to the additional constrains of $x_i < s^*$ and $x_i < s^{**}$ for all i and $s^* > s^{**}$ represent different investment proportions allowed under V^* and V^{**} and abs denotes absolute value. Via the envelope theorem (Dixit 1990) we have

$$\frac{d\{L(x_i(k),k) = \upsilon(x_i(k))\}}{dk} = \frac{\partial\{L(x_i(k),k) = \upsilon(x_i(k))\}}{\partial k} = \lambda|x_i = x_i(k) \tag{99.7}$$

Hence, Eq. 99.6 can be rewritten as

$$abs\left(\frac{\partial \lambda}{\partial k}\right) \leq abs\left(\frac{\partial \lambda^*}{\partial k}\right) \leq abs\left(\frac{\partial \lambda^{**}}{\partial k}\right) \tag{99.8}$$

Equation 99.8 states that the Lagrange multiplier of the original Markowitz portfolio selection model is less sensitive to an infinitesimally small change in k than that of the model when the constraints are gradually tightened. Note that the Lagrange multiplier λ is the reciprocal of the slope of the efficiency frontier curve frequently drawn in investment textbooks. Hence, the original Markowitz model has the steepest slope for a given set of x_i s. However, the efficiency frontier curve of the Markowitz minimization system has a vertical segment corresponding to a range of low ks and a constant v. Only within this range do the values of optimum xs remain equal under various degrees of constraints. Within this range constraint Eq. 99.2 is not active; hence the Lagrange multiplier is zero. As a result, equality relation holds for Eq. 99.8. Outside this range, the slopes of the efficiency frontier curve are different owing to the result of Eq. 99.8.

99.3 Simulation Results

To verify the result implied by the Le Chatelier principle, we employ a five-stock portfolio with $x_i < 50\ \%$ and $x_i < 40\ \%$. The numerical solutions are reported in Table 99.1. An examination of Table 99.1 indicates that the efficiency frontier curve is vertical and all optimum xs are identical between $0.001 < k < 0.075$. After that, the solutions of xs begin to change for the three models. Note that the maximum possible value for x_4 remains 0.4 throughout the simulation for $k > 0.075$ for the model with the tightest constraint xi < 0.4. In the case of $x_i < 0.5$, a relatively loosely constrained Markowitz system, all the optimum values of decision variables remain the same as the original Markowitz model between $0.01 < k < 0.1$. Beyond that range, the maximum value of x_4 is limited to 0.5. As can be seen from Table 99.1, the total risk v responds less volatile to the change in k in the original unconstrained Markowitz system than that in the constrained systems. In other words, the original Markowitz minimization system guarantees

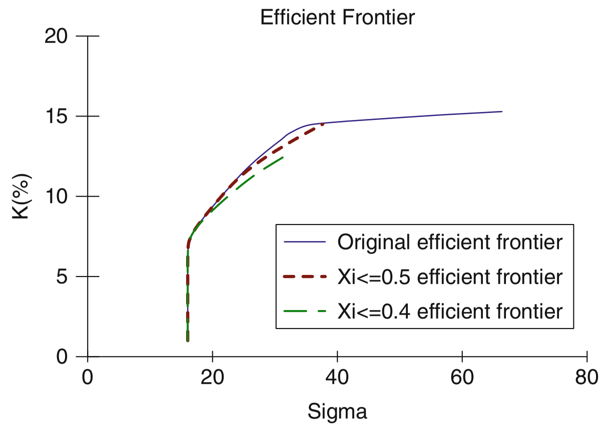
Table 99.1 Simulation Results

Least constrained solution (original Markowitz model)							Solution with $x_i \leq 0.5$						Solution with $x_i \leq 0.4$					
k (%)	$v(10^{-5})$	x_1 %	x_2 %	x_3 %	x_4 %	x_5 %	$v(10^{-5})$	x_1 %	x_2 %	x_3 %	x_4 %	x_5 %	$v(10^{-5})$	x_1 %	x_2 %	x_3 %	x_4 %	x_5 %
1	257.2	39.19	0	31.87	28.94	0	257.2	39.19	0	31.87	28.94	0	257.2	39.19	0	31.87	28.94	0
2	257.2	39.19	0	31.87	28.94	0	257.2	39.19	0	31.87	28.94	0	257.2	39.19	0	31.87	28.94	0
3	257.2	39.19	0	31.87	28.94	0	257.2	39.19	0	31.87	28.94	0	257.2	39.19	0	31.87	28.94	0
4	257.2	39.19	0	31.87	28.94	0	257.2	39.19	0	31.87	28.94	0	257.2	39.19	0	31.87	28.94	0
5	257.2	39.19	0	31.87	28.94	0	257.2	39.19	0	31.87	28.94	0	257.2	39.19	0	31.87	28.94	0
6	257.2	39.19	0	31.87	28.94	0	257.2	39.19	0	31.87	28.94	0	257.2	39.19	0	31.87	28.94	0
7	260.8	35.02	0	32.6	32.38	0	260.8	35.02	0	32.6	32.38	0	260.8	35.02	0	32.6	32.38	0
7.5	274.8	30.54	0	32.77	36.69	0	274.8	30.54	0	32.77	36.69	0	274.8	30.54	0	32.77	36.69	0
8	299.3	25.82	0	33.27	40.91	0	299.3	25.82	0	33.27	40.91	0	300.5	24.91	0	34.55	40	5.39
8.5	333.1	21.65	0	33.26	43.63	1.45	333.1	21.65	0	33.26	43.63	1.45	340.2	20.42	0	35.34	40	4.24
9	371.2	17.82	0	32.92	45.73	3.53	371.2	17.82	0	32.92	45.73	3.53	387.7	15.93	0	36.13	40	7.94
9.5	413.2	14.05	0	32.53	47.64	5.79	413.2	14.05	0	32.53	47.64	5.79	443	11.44	0	36.92	40	11.64
10	459	9.68	0.58	32.17	49.59	7.98	459	9.68	0.58	32.17	49.59	7.98	506.2	6.95	0	37.71	40	15.34
10.5	508.3	4.83	1.96	31.44	51.56	10.2	509.5	4.25	2.1	32.23	50	11.42	576.7	1.23	1.93	37.7	40	19.15
11	560.9	0	3.53	30.46	53.55	12.46	567.5	0	2.66	32.03	50	15.31	656.5	0	0.21	36.45	40	23.34
11.5	619.9	0	1.34	27.91	55.8	14.95	637.4	0	0	30.39	50	19.62	751.7	0	0	31.79	40	28.22
12	687.5	0	0	24.31	58.11	17.58	724.5	0	0	25.39	50	24.62	866.3	0	0	26.79	40	33.22
12.5	765.4	0	0	19.02	60.68	20.3	826.7	0	0	20.52	50	29.48	995.2	0	0	21.91	40	38.09
13	854.3	0	0	13.73	63.2	23.07	949.7	0	0	15.53	50	34.48						
13.5	954	0	0	8.45	65.72	25.83	1,086.8	0	0	10.65	50	39.45						

(*continued*)

Table 99.1 (continued)

Least constrained solution (original Markowitz model)							Solution with $x_i \leq 0.5$						Solution with $x_i \leq 0.4$					
k (%)	$v(10^{-5})$	x_1 %	x_2 %	x_3 %	x_4 %	x_5 %	$v(10^{-5})$	x_1 %	x_2 %	x_3 %	x_4 %	x_5 %	v(10^{-5})	x_1 %	x_2 %	x_3 %	x_4 %	x_5 %
14	1,064.6	0	0	3.16	68.25	28.59	1,243.3	0	0	5.73	50	44.28						
14.5	1,309.1	0	0	0	55.63	44.37	1,417.7	0	0	0.79	50	49.21						
15	2,847.3	0	0	0	20	80												
15.29	4,402	0	0	0	0	100												



a smallest possible total risk due to the result of the Le Chatelier principle: a thermodynamic system (risk-return space) can most effectively adjust itself to the parametric change (temperature or minimum rate of return of a portfolio or k) if it is least constrained.

99.4 Policy Implications of the Le Chatelier Principle

As shown in the previous section, the efficiency frontier curve branches off to the right first for the most binding constraint of $x_i < s^{**}$. Consequently, the tangency point between the efficiency frontier curve and a risk-free rate on the vertical axis must occur at a higher risk-free rate. As the value of maximum investment proportion for each stock s decreases, i.e., the constraint becomes more binding, there is a tendency for the risk-free rate to be higher in order to sustain an equilibrium (tangency) state. Second, one can assume the existence of a family of isowelfare functions (or indifference curves) in the v-k space. The direct impact of the Le Chatelier principle on the capital market equilibrium is a lower level of welfare measure due to the right branching-off of the efficiency frontier curve. In sum, as the constraint on the maximum investment proportion is tighter, the risk-free rate will be higher and investors in the capital market will in general experience a lower welfare level. In particular, the 5 % rule carries a substantial cost in terms of shifting of the efficiency frontier to the right. The study by Loviscek and Yang (1997) based on a 36-security portfolio indicates the loss is about 1–2 percentage points and the portfolio risk is 20–60 % higher. Given the astronomical size of a mutual fund, 1–2 percentage point translates into millions of dollars potential loss in daily return. Furthermore, over diversification would incur greater transaction and information cost, which speaks against the Investment Company Rule.

99.5 Concluding Remarks

In this note, we apply the Le Chatelier principle in the thermodynamics to the Markowitz portfolio selection model. The analogy is clear: as a thermodynamic system (or the capital market in the v-k space) undergoes some parametric changes (temperature or minimum portfolio rate of change k), the system will adjust most effectively if it is least constrained. The simulation shows that as the constraint becomes more and more tightened, the optimum investment proportions are less and less sensitive. Via the envelop theorem, it is shown that investors will be experiencing a higher risk-free rate and a lower welfare level in the capital market, if a majority of investors in the capital market experience the same constraint, i.e., maximum investment proportion on each security. Moreover, the potential loss in daily returns can easily be in millions on top of much greater transaction and information costs.

References

Almazan, A., Brown, K. C., Carlson, M., & Chapman, D. A. (2004). Why constrain your mutual fund manager? *Journal of Financial Economics, 73*, 289–321.

Cohen, K. J., & Pogue, J. A. (1967). An empirical evaluation of alternative portfolio selection models. *Journal of Business, 40*, 166–189.

Diamond, P., & Mirrlees, J. A.(2002). *Optimal taxation and the Le Chatelier principle*. MIT Department of Economics. Working paper series, 24 Sept 2002.

Dixit, A. K. (1990). *Optimization in economic theory* (2nd ed.). Oxford: Oxford University Press.

Elton, E. J., & Gruber, M. J. (1991). *Modern portfolio theory and investment analysis*. New York: Wiley.

Francis, J. C. (1993). *Management of investment*. New York: McGraw-Hill.

Frost, P. A., & Savarino, J. E. (1988). For better performance, constrain portfolio weights. *Journal of Portfolio Management, 14*, 29–34.

Kohn, M. G. (1994). *Financial institutions and markets*. New York: McGraw-Hill.

Labys, W. C., & Chin-Wei Yang. (1996). Le Chatelier principle and the flow sensitivity of spatial commodity models. In J. C. J. M. van de Bergh, P. Nijkamp, & P. Rietveld (Eds.), *Recent advances in spatial equilibrium modeling: Methodology and applications* (pp. 96–110). Berlin: Springer.

Loviscek, A. L., & Chin-Wei Yang. (1997). Assessing the impact of the Investment Company Act's diversification rule: A portfolio approach. *Midwest Review of Finance and Insurance, 11*(1), 57–66.

Markowitz, H. M. (1952). Portfolio selection. *The Journal of Finance, VII*(1), 77–91.

Markowitz, H. M. (1956). The optimization of a quadratic function subject to linear constraints. *Naval Research Logistics Quarterly, 3*, 111–133.

Markowitz, H. M. (1959). *Portfolio selection*. New York: Wiley.

Markowitz, H. M. (1990). *Mean-variance analysis in portfolio choice and capital markets*. Oxford: Basil Blackwell.

Markowitz, H. M. (1991). Foundation of portfolio theory. *Journal of Finance, XLVI*(2), 469–477.

Roe, M. J. (1991). Political element in the creation of a mutual fund industry. *University of Pennsylvania Law Review, 1469*, 1–87.

Samuelson, P. A. (1949). *The Le Chatelier principle in linear programming*. Rand Corporation monograph.

Samuelson, P. A. (1960). An extension of the Le Chatelier principle. *Econometrica, 28*, 368–379.

Samuelson, P. A. (1972). Maximum principles in analytical economics. *American Economic Review, 62*(3), 249–262.

Silberberg, E. (1971). The Le Chatelier principle as a corollary to a generalized envelope theorem. *Journal of Economic Theory, 3*, 146–155.

Silberberg, E. (1974). A revision of comparative statics methodology in economics, or how to do economics on the back of an envelope. *Journal of Economic Theory, 7*, 159–172.

Silberberg, E. (1978). *The structure of economics: A mathematical analysis*. New York: McGraw-Hill. Value line, mutual fund survey (New York: Value Line, 1994).

Author Index

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1,

D

E

F

I

J

K

N

O

Q

S

T

X

Y

Subject Index

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*, DOI 10.1007/978-1-4614-7750-1,

F

H

N

Y

Z

Reference Index

Aas, K., Czado, C., Frigessi, A., & Bakken, H. (2007). Pair-copula constructions of multiple dependence. *Insurance: Mathematics and Economics, 2*, 1–25.

Abarbanell, J. S., & Bushee, B. J. (1997). Fundamental Analysis, Future Earnings, and Stock Prices. *Journal of Accounting Research, 35*, 1–24.

Abdel-Aader, M. G., & Qing, K. Y. (2007). Risk-adjusted performance, selectivity, timing ability and performance persistence of Hong Kong mutual funds. *Journal of Asia-Pacific Business, 8*(2), 25–58.

Abdel-khalik, A. R., & Lusk, E. J. (1974). Transfer pricing - a synthesis. *The Accounting Review, 49*, 8–23.

Abel, A. B. (1988). Stock prices under time-varying dividend risk: an exact solution in an infinite-horizon general equilibrium model. *Journal of Monetary Economics, 22*, 373–393.

Abhyankar, A., Basu, D., & Streme, A. (2007). Portfolio efficiency and discount factor bounds with conditioning information: An empirical study. *Journal of Banking and Finance, 31*, 419–437.

Aboody, D., & Kaznik, R. (2000). CEO stock option awards and the timing of corporate voluntary disclosures. *Journal of Accounting Economics, 29*(1), 73–100.

Abraham, J. M., & Hendershott, P. H. (1996). Bubbles. *Journal of Housing Research, 7*(2), 191–208.

Abramson, D., et al. (2006). Deploying Scientific Applications to the PRAGMA Grid Testbed: Strategies and Lessons. Sixth IEEE International Symposium on Cluster Computing and the Grid, 241–248.

Acharya, S. (1993). Value of latent information: alternative event study methods. *Journal of Finance, 48*, 363–385.

Acharya, V. V., Pedersen, L. H., Philippon, T, & Richardson, M. (2010). Measuring systemic risk, Working Paper 10-02, Federal Reserve Bank of Cleveland.

Ackert, L., Charupat, N., Church, B., & Deaves, R. (2002). Bubbles in Experimental Asset Markets: Irrational Exuberance No More, working paper 2002-24.

Ackert, L., Charupat, N., Deaves, R., & Kluger, B. (2006). The Origins of Bubbles in Laboratory Asset Markets, Federal Reserve Bank of Atlanta, working paper 2006-6.

Adams, J., & Montesi, C. J. (1995). *Major issues related to hedge accounting*. Newark, CT: Financial Accounting Standard Board.

Adams, Z., Füss, R, & Gropp, R. (2010). Modeling spillover effects among financial institutions: A State-Dependent Sensitivity Value-at-Risk (SDSVaR) approach, Research Paper 10-12, European Business School.

Adrian, T., & Brunnermeier, M. K. (2011). *CoVaR, Staff Reports 348*. New York: Federal Reserve Bank of New York.

Afonso, A. (2003). Understanding the Determinants of Sovereign Debt Ratings: Evidence for the Two Leading Agencies. *Journal of Economics and Finance, 27*(1), 56–74.

Afonso A., Gomes P., Rother P. (2007). What "Hides" behind Sovereign Debt Ratings? European Central Bank Working Paper Series 711.

C.-F. Lee, J. Lee (eds.), *Handbook of Financial Econometrics and Statistics*,
DOI 10.1007/978-1-4614-7750-1,

Afza, T., & Rauf, A. (2009). Performance Evaluation of Pakistani mutual funds. *Pakistan Economic and Social Review, 47*(2), 199–214.

Agarwal, V., & Naik, N. Y. (2004). Risk and portfolios decisions involving hedge funds. *Review of Financial Studies, 17*, 63–98.

Agca, S., & Mozumdar, A. (2007). Investment-Cash Flow Sensitivity: Myth or Reality? Working paper, George Washington University.

Aggarwal, R., & Samwick, A. (1999). The other side of the trade-off: The impact of risk on executive compensation. *Journal of Political Economy, 107*, 65–105.

Aggarwal, R., & Goodell, J. W. (2009). Markets and institutions in financial intermediation: national characteristics as determinants. *Journal of Banking and Finance, 33*, 1770–1780.

Aggarwal, R., & Goodell, J. W. (2011a). International variations in expected equity premia: Role of financial architecture and governance. *Journal of Banking and Finance, 35*, 3090–3100.

Aggarwal, R., & Goodell, J. W. (2011b). Variations in emerging market equity premia: Impact of financial architecture. *Journal of the Asia Pacific Economy, 16*, 313–330.

Aggarwal, R., & Zhao, S. (2009). The diversification discount puzzle: Empirical evidence for a transactions cost resolution. *Financial Review, 44*, 113–135.

Aggarwal, R. K., & Samwick, A. A. (2003). Performance incentives within firms: The effect of managerial responsibility. *Journal of Finance, 58*, 1613–1649.

Agrawal, A., & Mandelker, G. N. (1987). Managerial incentives and corporate investment and financing decisions. *Journal of Finance, 42*, 823–837.

Agrawal, A, & Chen, M. (2012). Analyst conflicts and research quality. *Quarterly Journal of Finance*, forthcoming.

Agrawal, A., & Jayaraman, N. (1994). The dividend policies of all-equity firms: A direct test of the free cash flow theory. *Managerial & Decision Economics, 15*, 139–148.

Agrawal, A., Jaffe, J., & Mandelker, G. (1992). The post-merge performance of acquiring firms: A re-examination of an anomaly. *Journal of Finance, 47*, 1605–1621.

Agresti, A. (2007). *An Introduction to Categorical Data Analysis* (2nd ed.). Hoboken, New Jersey: Wiley Interscience.

Ahmad, K., Gillam, L., & Cheng, D. (2005). Textual and Quantitative Analysis: Towards a new, e-mediated Social Science. Proceedings of the 1st International Conference on e-Social Science, Manchester, 22–24.

Ahn, D.H., Gao, B. (1999). A parametric nonlinear model of term structure dynamics. *The Review of Financial Studies* 12 4, pp. 721.762.

Ahn, D. H., Conrad, J., & Dittmar, R. F. (2003a). Risk adjustment and trading strategies. *Review of Financial Studies, 16*, 459–485.

Ahn, D.-H., Dittmar, R. F., Ronald Gallant, A., & Gao, B. (2003b). Purebred or Hybrid?: Reproducing the Volatility in Term Structure Dynamics. *Journal of Econometrics, 116*, 147–180.

Ahn, D.-H., Dittmar, R. F., & Ronald Gallant, A. (2002). Quadratic Term Structure Models: Theory and Evidence. *Review of Financial Studies, 15*, 243–288.

Aitchison, J., & Brown, J. A. C. (1973). *Lognormal Distribution, with Special Reference to Its Uses in Economics*. Cambridge: Cambridge University Press.

Aït-Sahalia, Y. (1996). Testing Continuous Time Models of the Spot Interest Rate. *Review of Financial Studies, 9*, 385–426.

Aït-Sahalia, Y. (2002). Maximum Likelihood Estimation of Discretely Sampled Diffusions: A Closed Form Approximation Approach. *Econometrica, 70*, 223–262.

Aït-Sahalia, Y. (2007). Estimating Continuous-Time Models Using Discretely Sampled Data. Econometric Society World Congress Invited Lecture, in Advances in Economics and Econometrics, Theory and Applications, NinthWorld Congress, edited by Richard Blundell, Persson Torsten and Whitney K. Newey, Econometric Society Monographs, Cambridge University Press.

Ait-Sahalia, Y., & Lo, A. W. (2000). Nonparametric risk management and implied risk aversion. *Journal of Econometrics, 94*, 9–51.

Aït-Sahalia, Y., Fan, J., & Peng, H. (2009). Nonparametric Transition-Based Tests for Diffusions. *Journal of the American Statistical Association, 104*, 1102–1116.
Aït-Sahalia, Y., & Duarte, J. (2003). Nonparametric option pricing under shape restrictions. *Journal of Econometrics, 116*, 9–47.
Aït-Sahalia, Y., & Andrew, W. L. (1998). Nonparametric Estimation of State-Price Densities Implicit in Financial Asset Prices. *Journal of Finance, American Finance Association, 53*(2), 499–547.
Aivazian, V., Booth, L., & Cleary, S. (2003). Do emerging market firms follow different dividend policies from U.S. firms? *Journal of Financial Research, 26*, 371–387.
Akaike, H. (1973). Information theory and an extension of the maximum likelihood principle. In Second international symposium on information theory, vol 1, B.N. Petrov and F. Csaki (Eds), 267–281.
Akay, O., Griffiths, M. D., & Winters, D. B. (2010). On the robustness of range-based volatility estimators. *Journal of Financial Research, 33*, 179–199.
Akgiray, V. (1989). Conditional Heteroskedasticity in Time Series of Stock Returns: Evidence and Forecasts. *Journal of Business, 62*, 55–80.
Akhigbe, A., Borde, S. F., & Madura, J. (1993). Dividend policy and signaling by insurance companies. *Journal of Risk & Insurance, 60*, 413–428.
Akyurek, C., & Kutan, M. A. (2008). Inflation targeting, policy rates and exchange rate volatility: evidence from Turkey. *Comparative Economic Studies, 50*, 460–493.
Albert, J. H., & Chib, S. (1993). Bayes inference via Gibbs sampling of autoregressive time series subject to Markov mean and variance shifts. *Journal of Business and Economic Statistics, 11*(1), 1–15.
Albrecher, H., Mayer, P., Schoutens, W., & Tistaert, J. (2007). The little Heston trap, Wilmott Magazine, January, 83–92.
Albuquerque, A. (2005). Who are your peers? A study of relative performance evaluation. Working Paper, Simon School of Business, University of Rochester.
Alexander, T Obeid. (2004). A modified approach for Risk-adjusted performance attribution. *Financial markets and portfolio management*, 18(3) 285–305.
Alexe, G., Alexe, S., Bonates, T. O., & Kogan, A. (2007). Logical Analysis of Data – the Vision of Peter L. Hammer. *Annals of Mathematics and Artificial Intelligence, 49*, 265–312.
Alexe, S. (2002). Datascope - A New Tool for Logical Analysis of Data. DIMACS Mixer Series. DIMACS, Rutgers University.
Alizadeh, A. H., & Nomikos, N. (2004). A Markov regime switching approach for hedging stock indices. *Journal of Futures Markets, 24*, 649–674.
Alizadeh, A. H., Nomikos, N., & Pouliasis, P. K. (2008). A Markov regime switching approach for hedging energy commodities. *Journal of Banking and Finance, 32*, 1970–1983.
Alizadeh, S., Brandt, M. W., & Diebold, F. X. (2002). Range-based estimation of stochastic volatility models. *Journal of Finance, 57*, 1047–1092.
Allen, D. (1974). The relationship between variable selection and data agumentation and a method for prediction. *Technometrics, 16*, 125–127.
Allen, F., & Gale, D. (2000). Bubbles and Crises. *Economic Journal, 110*, 236–255.
Allen, F., & Gorton, G. (1988). Rational finite bubbles, Rodney L. White Center for Financial Research working paper no. 41–88.
Allison, P. D. (1984). *Event history analysis*. Beverly Hills, California: Sage Publications.
Allison, P. D. (Ed.). (1995). *Survival analysis using SAS: A practical guide*. Cary, North Carolina: SAS Institute Inc.
Almeida, H., & Campello, M. (2007). Financial Constraints, Asset Tangibility and Corporate Investment. *Review of Financial Studies, 20*, 1429–1460.
Almond, J., & Snelling, D. (1999). UNICORE: uniform access to supercomputing as an element of electronic commerce. *Future Generation Computer Systems, 15*(5), 539–548.
Aloui, C., & Mabrouk, S. (2010). Value-at-risk estimations of energy commodities via long-memory, asymmetry and fat-tailed GARCH models. *Energy Policy, 38*, 2326–2339.

Altissimo, F., & Mele, A. (2009). Simulated Nonparametric Estimation of Dynamic Models with Application in Finance. *Review of Economic Studies, 76*, 413–450.
Altman, E. I., & Rijken, H. A. (2004). How Rating Agencies Achieve Rating Stability. *Journal of Banking and Finance, 28*, 2679–2714.
Altman, E. I., & Saunders, A. (1997). Credit Risk Measurement: Developments over the Last 20 Years. *Journal of Banking and Finance, 21*(11–12), 1721–1742.
Altman, E. (1992). Revisiting the high yield bond market. *Financial Management, 21*(2), 78–92.
Altman, E. (1998). The importance and subtlety of credit rating migration. *Journal of Banking and Finance, 22*, 1231–1247.
Altman, E., & Kao, D. (1991). *Corporate bond rating drift: An examination of credit quality rating changes over time, The Research Foundation of the Institute of Chartered Financial Analysts*. Virginia: Charlottesville.
Altman, E., & Kao, D. L. (1992). Rating drift of high yield bonds. *Journal of Fixed Income, 2*, 15–20.
Altman, E., & Kao, D. (1992b). The implications of corporate bond rating drifts. *Financial Analysts Journal* May-June, 64–75.
Altman, E. I. (1968). Financial ratios discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance, 23*, 589–609.
Altman, E. I., & Sounders, A. (1997). Credit risk measurement: Developments over the last 20 years. *Journal of Banking and Finance, 21*, 1721–1742.
Altman, E. I., & Katz, S. (1976). Statistical bond rating classification using financial and accounting data. Proceedings of the Conference on Topical Research in Accounting, New York.
Altman, E. I., Hadelman, R. G., & Narayanan, P. (1977). Zeta analysis, a new model to identify bankruptcy risk of corporations. *Journal of Banking and Finance, 10*, 29–51.
Altman, E. I. (1983). *Predicting corporate bankruptcy: the Z-Score model. In: Corporate Financial Distress: A Complete Guide to Predicting, Avoiding and Dealing with Bankruptcy*. New York: Wiley.
Altman, E. I., & Eisenbeis, R. A. (1968). Financial applications of discriminant analysis: A clarification. *Journal of Financial and Quantitative Analysis, 13*, 185–95.
Altman, E. I., Haldeman, R., & Narayanan, P. (1977). ZETA Analysis, a new model for bankruptcy classification. *Journal of Banking and Finance, 1*, 29–54.
Altonji, J., & Segal, L. (1996). Small-Sample Bias in GMM Estimation of Covariance Structures. *Journal of Business & Economic Statistics, 14*, 353–366.
Amad, K., et al. (2004). Financial Information Grid –an ESRC e-Social Science Pilot. Proceedings of the Third UK e-Science Programme All-Hands Meeting, Nottingham, UK.
Amar Ranu, & Depali Ran. (2010). The Best Performing Equity Funds, The Analyst, Jan.
Amato, J., & Furfine, C. (2004). Are credit ratings procyclical? *Journal of Banking and Finance, 28*, 2641–2677.
Amemiya, T. (1974). A Note on a Fair and Jaffee Model. *Econometrica, 42*, 759–762.
Amemiya, T. (1980). Selection of Regressors. *International Economic Review, 21*, 331–54.
Amihud, Y. (2002). Illiquidity and stock returns: Cross-section and time-series effects. *Journal of Financial Markets, 5*, 31–56.
Amihud, Y., & Mendelson, H. (1986). Asset pricing and the bid-ask spread. *Journal of Financial Economics, 17*, 223–249.
Amihud, Y., & Mendelson, H. (1991). Liquidity, maturity, and the yields on US Treasury securities. *The Journal of Finance, 46*, 1411–1425.
Amihud, Y., & Mendelson, H. (1987). Trading Mechanism and Stock Returns: An Empirical Investigation. *Journal of Finance, 42*, 533–553.
Amin, K., & Jarrow, R. (1992). Pricing Options on Risky Assets in a Stochastic Interest Rate Economy, *Mathematical Finance, 2*, 217–237.
Amin, K., & Ng, V. (1993). Option Valuation With Systematic Stochastic Volatility. *Journal of Finance, 48*, 881–910.

Amit Singh Sisodiya. (2004). Mutual Fund Industry in India- Coming of Age, Chartered Financial Analyst, July, page 17–22.
Anagnostopoulou, S. C. (2008). R&D expenses and firm valuation: a literature review. *International Journal of Accounting and Information Management, 16*, 5–24.
Ananth, C. V., & Kleinbaum, D. G. (1997). Regression models for ordinal responses: a review of methods and applications. *International journal of epidemiology, 26*, 1323–1333.
Anas, A., & Eum, S. J. (1984). Hedonic Analysis of a Housing Market in Disequilibrium. *Journal of Urban Economics, 15*, 87–106.
Andersen, J. V., & Sornette, D. (2001). Have Your Cake and Eat It Too: Increasing Returns while Lowering Large Risks! *Journal of Risk Finance, 2*, 70–82.
Andersen, L., & Sidenius, J. (2004). Extensions to the Gaussian Copula: Random Recovery and Random Factor Loadings. *Journal of Credit Risk, 1*(1), 29–70.
Andersen, L., & Brotherton-Ratcliffe, R. (2001). Extended LIBOR Market Models with Stochastic Volatility, working paper, Gen Re Securities.
Andersen, L., Sidenius, J., & Basu, S. (2003). All Your Hedges in One Basket. Risk November, 67–72.
Andersen, P. K. (1992). Repeated assessment of risk factors in survival analysis. *Statistical Methods in Medical Research, 1*, 297–315.
Andersen, T. G. (1996). Return volatility and trading volume: An information flow interpretation of stochastic volatility. *Journal of Finance, 51*, 169–204.
Andersen, T. G., Bollerslev, T., & Diebold, F. X. (2004). Parametric and nonparametric volatility measurement. In Y. Aït-Sahalia & L. P. Hansen (Eds.), *Handbook of Financial Econometrics*. Amsterdam: North Holland.
Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association, 96*, 42–55.
Andersen, T. G., Bollerslev, T., Diebold, F. X., & Ebens, H. (2001). The distribution of realized stock return volatility. *Journal of Financial Economics, 61*, 43–76.
Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2003). Modeling and forecasting realized volatility. *Econometrica, 71*, 579–626.
Andersen, T., & Lund, J. (1997). Estimating Continuous Time Stochastic Volatility Models of the Short Term Interest Rate. *Journal of Econometrics, 77*, 343–377.
Andersen, T., Bollerslev, T., Diebold, F., & Vega, C. (2007). Real time price discovery in global stock, bond and foreign exchange markets. *Journal of International Economics, 73*, 251–277.
Andersen, T., Bollerslev, T., Christoffersen, P., & Diebold, F. (2006). Volatility and correlation forecasting. In G. Elliott, C. W. J. Granger, & A. Timmermann (Eds.), *Handbook of Economic Forecasting* (pp. 778–878). Amsterdam: North-Holland.
Andersen, T., Bollerslev, T., Christoffersen, P., & Diebold, F. (2007). Practical volatility and correlation modeling for financial market risk management. In M. Carey & R. Stulz (Eds.), *The Risks of Financial Institutions* (pp. 513–548). Chicago: University of Chicago Press for NBER.
Andersen, T. G., & Lund, J. (1997). Estimating Continuous Time Stochastic Volatility Models of the Short Term Interest Rate. *Journal of Econometrics, 77*, 343–378.
Andersen, T. G., & Bollerslev, T. (1997). Heterogeneous information arrivals and return volatility dynamics: uncovering the long-run in high frequency returns. *Journal of Finance, 52*, 975–1005.
Andersen, T. G., Bollerslev, T., Diebold, F. X., & Ebens, H. (2001). The distributions of realized stock return volatility. *Journal of Financial Economics, 61*, 43–76.
Andersen, T.G., Benzoni, L., & Lund, J. (2004). Stochastic Volatility, Mean Drift, and Jumps in the Short-Term Interest Rate. Working Paper, Northwestern University.
Andersen, T. G., Bollerslev, T., Diebold, F. X., & Labys, P. (2001). The distribution of realized exchange rate volatility. *Journal of the American Statistical Association, 96*(453), 42–55.
Andersen, T. G., & Luca Benzoni. (2006). Do bonds span volatility risk in the U.S. Treasury market? A specification test of Affine term structure models, Working paper, Northwestern University.

Anderson, D. R., Sweeney, D. J., & Williams, T. A. (2011a). *Statistics for Business and Economics* (11th ed.). Mason, OH: South-Western, Cengage Learning.

Anderson, T. W., Kunitomo, N., & Matsushita, Y. (2011b). On Finite Sample Properties of Alternative Estimators of Coefficients in a Structural Equation with Many Instruments. *Journal of Econometrics., 165*, 58–69.

Anderson, G. J. (1983). The internal financing decisions of the industrial and commercial sector: A reappraisal of the Lintner model of dividend disbursements. *Economica, 50*, 235–248.

Anderson, H. M., & Vahid, F. (2007). Forecasting the volatility of Australian stock returns. *Journal of Business and Economic Statistics, 25*, 76–90.

Anderson, T. W. (1994). *The Statistical Analysis of Time Series*. New york: Wiley-Interscience.

Anderson, T. W. (2003a). *An Introduction to Multivariate Statistical Analysis*. New york: Wiley-Interscience.

Anderson, T. W. (2003b). *An Introduction to Multivariate Statistical Analysis* (3rd ed.). New York: John Wiley and Sons.

Anderson, T. W., & Hsiao, C. (1981a). Estimation of dynamic models with error components. *Journal of the American Statistical Association, 76*, 589–606.

Anderson, T. W., & Hsiao, C. (1981b). Estimation of dynamic models with error components. *Journal of the American Statistical Association, 76*, 598–606.

Anderson, T. W., & Rubin, H. (1949). Estimation of the parameters of a single equation in a complete system of stochastic equations. *Annuals of Mathematical Statistics, 20*, 46–63.

Anderson, T. W., Kunitomo, N., & Matsushita, Y. (2010). On the Asymptotic Optimality of the LIML Estimator with Possibly Many Instruments. *Journal of Econometrics., 157*, 191–204.

Anderson, T. W., & Darling, D. A. (1952). Asymptotic Theory of Certain 'Goodness of Fit' Criteria Based on Stochastic Processes. *The Annals of Mathematical Statistics, 23*, 193–212.

Anderson, T. W., & Darling, D. A. (1954). A Test of Goodness of Fit. *Journal of the American Statistical Association, 49*, 765–769.

Andews, D. W. K., & Kim, J. (2003). End-of-sample cointegration breakdown tests. Cowles Foundation Discussion Paper No. 1404.

Andrews, D. W. K. (1993). An Introduction to Econometric Applications of Empirical Process Theory for Dependent Random Variables. *Econometric Reviews, 12*, 183–216.

Andrews, D. W. K. (1997). A Conditional Kolmogorov Test. *Econometrica, 65*, 1097–1128.

Andrews, D. W. K., Lee, I., & Ploberger, W. (1996). Optimal changepoint tests for Normal linear regression. *Journal of Econometrics, 70*, 9–36.

Andrews, D. W. K., Lieberman, O., & Marmer, V. (2006). Higher order improvements of the parametric bootstrap for long-memory gaussian processes. *Journal of Econometrics, 133*, 673–702.

Andrews, D. W. K. (1993). Tests for parameter instability and structural change with unknown change point. *Econometrica, 61*, 821–856.

Ang, A., & Bekaert, G. (2002a). International asset allocation with regime shifts. *Review of Financial Studies, 15*, 1137–1187.

Ang, A., & Bekaert, G. (2002b). International Asset Allocation with Regime Shifts. *Review of Financial Studies, 15*(4), 1137–1187.

Ang, A., & Chen, J. (2002). Asymmetric correlations of equity portfolios. *Journal of Financial Economics, 63*, 443–494.

Ang, A., Chen, J., & Xing, Y. (2006a). Downside risk. *The Review of Financial Studies, 19*, 1191–1239.

Ang, A., Hodrick, R., Xing, Y., & Zhang, X. (2006b). The cross-section of volatility and expected returns. *Journal of Finance, 61*, 259–299.

Ang, J. S., & Zhang, S. (2004). An evaluation of testing procedures for long horizon event studies. *Review of Quantitative Finance and Accounting, 23*, 251–274.

Ang, J. S., Davitopolous, D., & Schwarz, T. (2009a). The creation and control of speculative bubbles in a laboratory setting. In A. Lee & C. F. Lee (Eds.), *Handbook of Quantitative Finance and Risk Management* (p. 137). New York: Springer.

Ang, J. S., Jou, G. D., & Lai, T. Y. (2009). Alternative Formulas to Compute Implied Standard Deviation. *Review of Pacific Basin Financial Markets and Policies, 12*, 159–176 (June).

Ang, J. S., Jou. G. D., & Lai, T. Y., (2012). A Comparison of Formulas to Compute Implied Standard Deviation. Encyclopedia of Finance, 2nd. ed. by Lee, C. F., A. C. Lee and J. C. Lee, Springer, forthcoming.

Ang, J., & Schwarz, T. (1985). Risk aversion and information structure: An experimental study of price variability in the securities market. *Journal of Finance, 40*, 924–844.

Ang, R., (2011). Sarbanes-Oxley effectiveness on the earnings management. Available at SSRN: http://ssrn.com/abstract=1791766.

Angel, J. J., Christophe, S. E., & Ferri, M. G. (2003). A close look at short selling on Nasdaq. *Financial Analysts Journal, 59*, 66–74.

Angelidis, T., Benos, A., & Degiannakis, S. (2004). The use of GARCH models in VaR estimation. *Statistical Methodology, 1*, 105–128.

Angrist, J. D., & Lavy, V. (1999). Using Maimonides' rule to estimate the effect of class size on scholastic achievement. *Quarterly Journal of Economics, 14*(2), 533–575.

Angrist, J. D., Imbens, G. W., & Krueger, A. B. (1999). Jackknife instrumental variables estimation. *Journal of Applied Econometrics, 14*, 57–67.

Ansoff, H. Igor., Brandenburg, R. G., Portner, F. E., & Raymond Radosevich. (1971). Acquisition Behavior of U.S. Manufacturing Firms. Vanderbilt University Press, 1946–1965.

Antoniadis, A. (1997). Wavelets in statistics: A review (with discussion). *Journal of the Italian Statistical Association, 6*, 97–144.

Antoniadis, A., & Fan, J. (2001). Regularization of wavelets approximations (with discussion). *Journal of the American Statistical Association, 96*, 939–967.

Antoniou, A., Koutmos, G., & Pericli, A. (2005). Index Futures and Positive Feedback Trading: Evidence from Major Stock Exchanges. *Journal of Empirical Finance, 12*, 219–238.

Aoki, M. (1967). *Optimization of Stochastic Systems*. New York, NY: Academic Press.

Arak, M., & Goodman, L. (1987). Treasury bond futures: Valuing the delivery option. *Journal of Futures Markets, 7*(3), 269–286.

Ardittti, Fred D., (1980). A Survey of Valuation and the Cost of Capital. Research in Finance 2, 1–56, JAI Press Inc.

Areal, N. M. P. C., & Taylor, S. J. (2002). The realized volatility of FTSE-100 futures prices. *Journal of Futures Markets, 22*, 627–648.

Arellano, M., & Bover, O. (1995). Another look at the instrumental variable estimation of error-components models. *Journal of Econometrics, 68*(1), 29–51.

Arellano, M., & Bond, S. (1991). Some Tests of Specification for Panel Data: Monte Carlo Evidence and an Application to Employment Equations. *Review of Economic Studies, 58*, 277–297.

Arkes, H., Dawson, N., Speroff, T., Harrell, F., Alzola, C., Phillips, R., Desbiens, N., Oye, R., Knaus, W., & A. Connors the Support Investigators. (1995). The covariance decomposition of the probability score and its use in evaluating prognostic estimates. *Medical Decision Making, 15*(2), 120–131.

Armstrong, J. S. (1989). Combining forecasts: the end of the beginning or the beginning of the end? *International Journal of Forecasting, 5*(4), 585–592.

Arnold, L., Redmand, N. S. G., & Manakyan, H. (2000). The Performance of Global and International Mutual Funds. *Journal of Financial and strategic Decisions, 13*(1), 75–85. Spring.

Arnold, B. C. (1980). *Majorization and the Lorenz Order: A Brief Introduction, Lecture Notes in Statistics* (Vol. 43). Berlin: Springer-Verlag.

Arrington, E. E., Hillison, W., & Jensen, R. E. (1984). An application of analytical hierarchy process to model expert judgment on analytic review procedures. *Journal of Accounting Research, 22*, 298–312.

Artzner, P., Delbaen, F., Eber, J.-M., & Heath, D. (1999). Coherent Measures of Risk. *Mathematical Finance, 9*, 203–228.

Aruoba, B. S., Diebold, F. X., & Rudebusch, G. D. (2006). The macroeconomy and yield curve: a dynamic latent factor approach. *Journal of Econometrics, 127*, 309–338.

Asai, M., McAleer, M., & Yu, J. (2006). Multivariate stochastic volatility. *Econometric Reviews, 25*(2–3), 145–175.

Asai, M. (2008). Autoregressive stochastic volatility models with heavy-tailed distributions: A comparison with multifactors volatility models. *Journal of Empirical Finance, 15*, 322–345.

Asai, M., & Unite, A. (2010). General asymmetric stochastic volatility models using range data: estimation and empirical evidence from emerging equity markets. *Applied Financial Economics, 20*, 1041–1049.

Asai, M., & Brugal, I. (2012). Forecasting volatility using range data: analysis for emerging equity markets in Latin America. *Applied Financial Economics, 22*, 461–470.

Asai, M., & McAleer, M. (2009). The structure of dynamic correlations in multivariate stochastic volatility models. *Journal of Econometrics, 150*, 182–192.

Ashbaugh, H., LaFond, R., & Mayhew, B. W. (2003). Do nonaudit services compromise auditor independence? Further evidence. *The Accounting Review, 78*(3), 611–639.

Ashbaugh-Skaife, H., Collins, D. W., Kinney, W. R., & LaFond, R. (2008). Internal control deficiencies, remediation and accrual quality. *The Accounting Review, 83*, 217–250.

Asquith, P., & Mullins, D. W., Jr. (1983). The impact of initiating dividend payments on shareholders' wealth. *Journal of Business, 56*, 77–96.

Assef, J. G., & Santos, M. (2005). Stock Options and Managerial Optimal Contracts. *Economic Theory, 26*, 813–837.

Association for Investment Manage and Research (2001). Research objectivity standards will create competitive market pressure for brokerage firms to support analyst objectivity. http://www.cfainstitute.org/pressroom/01releases/01congress.html.

Astola, J., & Kuosmanen, P. (1997). *Fundamentals of nonlinear filtering*. Boca Raton: CRC Press.

Atan, R., Jasni, N. R., & Shahwan, Y. (2010). The impact of IFRS 2 "share-based payment" on Malaysian companies. *Review of Pacific Basin Financial Markets and Policies, 13*(3), 449–468.

Athreya, K. B. (1987). Bootstrap of the mean in the infinite variance case. *Annals of Statistics, 15*, 724–731.

Atkins, D. E., et al. (2003). Revolutionizing Science and Engineering Through Cyberinfrastructure. Report of the National Science Foundation Blue-Ribbon Advisory Panel on Cyberinfrastructure. National Science Foundation.

Atkinson, A. C. (1982). The simulation of generalized inverse Gaussian and hyperbolic random variables. *SIAM Journal of Scientific & Statistical Computing, 3*, 502–515.

Atsulakis, G. S., & Valavanis, K. P. (2009). Surveying stock market forecasting techniques, Part II, soft computing methods. *Export Systems Applications, 3*, 5932–41.

Attfield, C. (1983). Consistent estimation of certain parameters in the unobservable variable model when there is a specification error. *Review of Economics and Statistics, 65*, 164–167.

Au, S., Duan, R., Hesar, S., & Jiang, W. (2010). A framework of irregularity enlightenment for data pre-processing in data mining. *Annals of Operations Research, 174*(1), 47–66.

Aussenegg, W., & Cech, C. (2011). Simple time-varying copula estimation. In: Barczak, A. S. and Dziwok, E. (Ed.), *Mathematical Econometrical and Computational Methods in Finance and Insurance*, University of Economics in Katowice, Katowice, forthcoming.

Avery, C., & Zemsky, P. (1998). Multidimensional Uncertainty and Herd Behavior in Financial Markets. *American Economic Review, 88*, 724–748.

Avramov, D. (2002). Stock Return Predictability and Model Uncertainty. *Journal of Financial Economics, 64*, 423–458.

Ayvaz, E., & Pehlivanli, D. (2011). The use of time driven activity based costing and analytic hierarchy process method in the balanced scorecard implementation. *International Journal of Business and Management, 6*, 146–158.

Azizpour, S., & Giesecke, K. (2008). Self-exciting corporate defaults: Contagion vs. Frailty. Working Paper.Stanford University.

Backus, D. K., & Gregory, A. W. (1993). Theoretical relations between risk premiums and conditional variance. *Journal of Business and Economic Statistics, 11*, 177–185.

Bae, K. H., Karolyi, G. A., & Stulz, R. M. (2003). A new approach to measuring financial contagion. *The Review of Financial Studies, 16*(3), 717–763.

Bae, S. C. (1990). Interest Rate Changes and Common Stock Returns of Financial Institutions: Revised. *Journal of Financial Research, 13*, 71–79.

Bae, S. C., Klein, D. P., & Padmaraj, R. (1994). Event risk bond covenants, agency costs of debt and equity, and stockholder wealth. *Financial Management, 23*, 28–41.

Bae, S. C., Klein, D. P., & Padmaraj, R. (1997). Firm characteristics and the presence of event risk covenants in bond indentures. *Journal of Financial Research, 20*, 373–88.

Bagehot, W. (1971). The Only Game in Town. *Financial Analysis Journal, 27*, 12–14.

Baginski, S., & Hassell, J. (1990). The Market Interpretation of Management Earnings Forecasts as a Predictor of Subsequent Financial Forecast Revision. *The Accounting Review, 65*, 175–90.

Bagnoli, M., & Bergstrom, T. (2005). Log-concave Probability and its Applications. *Economic Theory, 26*, 445–469.

Bai, J. (1997). Estimation of a changepoint in multiple regression models. *Review of Economics and Statistics, 79*, 551–563.

Bai, J. (1999). Likelihood ratio tests for multiple structural changes. *Journal of Econometrics, 91*, 299–323.

Bai, J. (2003). Testing Parametric Conditional Distributions of Dynamic Models. *Review of Economics and Statistics, 85*, 531–549.

Bai, J., & Ng, S. (2008). Forecasting economic time series using targeted predictors. *Journal of Econometrics, 146*, 304–317.

Bai, J., & Perron, P. (1996). Computation and analysis of multiple structural change models. Manuscript in Preparation, Universite de Montreal.

Bai, J., & Perron, P. (1998a). Estimation and testing linear models with multiple structural changes. *Econometrica, 66*, 47–78.

Bai, J., & Perron, P. (1998b). Testing for and estimation of multiple structural changes. *Econometrica, 66*, 817–858.

Bai, J., & Perron, P. (2003). Computation and analysis of multiple structural change models. *Journal of Applied Econometrics, 18*, 1–22.

Bai, J., Lumsdaine, R., & Stock, J. H. (1998). Testing for and dating common breaks in multivariate time series. *The Review of Economic Studies, 65*, 395–432.

Bai, Z. D., Li, H., Liu, H. X., & Wong, W. K. (2011a). Test statistics for prospect and markowitz stochastic dominances with applications. *Econometrics Journal, 14*, 278–303.

Bai, Z. D., Li, H., Wong, W. K., & Zhang, B. Z. (2011b). Multivariate causality tests with simulation and application. *Statistics and Probability Letters, 81*, 1063–1071.

Bai, Z. D., Liu, H. X., & Wong, W. K. (2009a). Enhancement of the applicability of Markowitz's portfolio optimization by utilizing random matrix theory. *Mathematical Finance, 19*, 639–667.

Bai, Z. D., Liu, H. X., & Wong, W. K. (2009b). On the markowitz mean-variance analysis of self-financing portfolios. *Risk and Decision analysis, 1*, 35–42.

Bai, Z. D., Liu, H. X., & Wong, W. K. (2011c). Asymptotic properties of eigen-matrices of a large sample covariance matrix. *Annals of Applied Probability, 21*, 1994–2015.

Bai, Z. D., Wang, K. Y., & Wong, W. K. (2011d). Mean-variance ratio test, a complement to coefficient of variation test and Sharpe ratio test. *Statistics and Probability Letters, 81*, 1078–1085.

Bai, Z. D., Wong, W. K., & Zhang, B. Z. (2010). Multivariate linear and non-linear causality tests. *Mathematics and Computers in Simulation, 81*, 5–17.

Bai, Z. D., Hui, Y. C., Wong, W. K., & Zitikis, R. (2012). Evaluating prospect performance: making a case for a non-asymptotic UMPU test. *Journal of Financial Econometrics*, (forthcoming).

Bailey, A. D., & Boe, W. J. (1976). Goal and resource transfers in the multigoal organization. *The Accounting Review, 51*, 559–573.

Bailey, W., & Stulz, R. (1989). The Pricing of Stock Index Options in a General Equilibrium Model. *Journal of Financial and Quantitative Analysis, 24*, 1–12.

Bailey, W., & Joseph, L. (1992). Evaluating the Diversification Benefits of the New Country Funds. *Journal of Portfolio Management, 18*, 74–80. Spring.

Bailey, W., & Jagtiani, J. (1994). Foreign Ownership Restrictions and Stock Prices in the Thai Capital Market. *Journal of Financial Economics, 36*, 57–87.

Baillie, R., & DeGennaro, R. P. (1990a). Stock returns and volatility. *Journal of Financial and Quantitative Analysis, 25*, 203–214.

Baillie, R. T., & DeGennaro, R. P. (1990b). Stock Returns and Volatility. *Journal of Financial and Quantitative Analysis, 25*, 203–215.

Baillie, R. T., & Myers, R. (1991). Bivariate GARCH estimation of the optimal commodity futures hedge. *Journal of Applied Econometrics, 6*, 109–124.

Baillie, R. T., & Chang, S. S. (2011). Carry trades, momentum trading and the forward premium anomaly. *Journal of Financial Markets, 14*, 441–464.

Baillie, R. T., Bollerslev, T., & Mikkelsen, H. O. (1996). Fractionally integrated generalized autoregressive conditional heteroskedasticity. *Journal of Econometrics, 74*, 3–30.

Baillie, R. T. (1996). Long memory processes and fractional integration in econometrics. *Journal of Econometrics, 73*, 5–59.

Baillie, R., & Bollwerslev, T. (1989). Common stochastic trends in a system of exchange rates. *Journal of Finance, 44*, 167–181.

Baily, W., Li, H., Mao, C., & Zhong, R. (2003). Regulation Fair Disclosure and earnings information: Market, analysts, and corporate response. *Journal of Finance, 58*, 2487–2514.

Bajari, P., & Lanier Benkard, C. (2005). Demand Estimation with Heterogeneous Consumers and Unobserved Product Characteristics: A Hedonic Approach. *Journal of Political Economy, 113*, 6.

Baker Kent, H., Haslem, J. A., & Smith, D. M. Performance and Characteristics of Actively Managed Institutional Equity Mutual Funds, Electronic copy source: http://ssrn.com/abstract=1124577.

Baker, G. P. (1992). Incentive contracts and performance measurement". *Journal of Political Economy, 100*, 598–614.

Baker, H. K., & Powell, G. E. (1999). How corporate managers view dividend policy. *Quarterly Journal of Business & Economics, 38*, 17–35.

Baker, H. K., Farrelly, G. E., & Edelman, R. B. (1985). A survey of management views on dividend policy. *Financial Management, 14*, 78–84.

Baker, M., & Stein, J. C. (2004). Market liquidity as a sentiment indicator. *Journal of Financial Markets, 7*, 271–300.

Bakke, T., & Whited, T. (2009). What Gives? A Study of Firms' Reactions to Cash Shortfalls. Working paper, University of Rochester.

Bakshi, G., Cao, C., & Chen, Z. (1997). Empirical performance of alternative option pricing models. *Journal of Finance, 52*(5), 2003–2049.

Bakshi, G., & Chen, Z. (1997a). An Alternative Valuation Model for Contingent Claims. *Journal of Financial Economics, 44*, 123–165.

Bakshi, G., Cao, C., & Chen, Z. (2000a). Do Call Prices and the Underlying Stock Always Move in the Same Direction? *Review of Financial Studies, 13*, 549–584.

Bakshi, G., Cao, C., & Chen, Z. (2000b). Pricing and Hedging Long-Term Options. *Journal of Econometrics, 94*, 277–318.

Bala, R., & Matthew, C. H. Y. (2003). Evaluating Mutual Funds in an Emerging Market: Factors that matter to Financial Advisors, International Journal of Bank Marketing, March, PP.122–136.

Balestra, P., & Varadharajan-Krishnakumar, J. (1987). Full information estimations of a system of simultaneous equations with error components structure. *Econometric Theory, 3*, 223–246.

Bali, T. G. (2008). The intertemporal relation between expected returns and risk. *Journal of Financial Economics, 87*, 101–131.

Bali, T. G., & Theodossiou, P. (2007). A conditional-SGT-VaR approach with alternative GARCH models. *Annals of Operations Research, 151*, 241–267.

Bali, T. G., & Peng, L. (2006). Is there a risk-return tradeoff? Evidence from high frequency data. *Journal of Applied Econometrics, 21*, 1169–1198.

Bali, T. G., & Engle, R. F. (2010). The intertemporal capital asset pricing model with dynamic conditional correlations. *Journal of Monetary Economics, 57*, 377–390.

Bali, T. G., Mo, H., & Tang, Y. (2008). The Role of Autoregressive Conditional Skewness and Kurtosis in the Estimation of Conditional VaR. *Journal of Banking & Finance, 32*, 269–282.

Ball, C., & Torous, W. (1984). The maximum likelihood estimation of security price volatility: Theory, evidence, and application to option pricing. *Journal of Business, 57*, 97–112.

Ball, C., & Torous, W. (1999). The Stochastic Volatility of Short-Term Interest Rates: Some International Evidence. *Journal of Finance, 54*, 2339–2359.

Ball, L., Sheridan, N., (2005). Does Inflation Targeting Matter? In: Bernanke, B., Woodford, M. (Eds.), The Inflation-Targeting Debate. University of Chicago Press.

Ball, R., & Brown, P. (1968). An Empirical Evaluation of Accounting Income Numbers. *Journal of Accounting Research*, 159–178.

Ball, R., & Watts, R. (1972). Some Time-Series of Accounting Income. *Journal of Finance, 27*, 663–82.

Ball, R., & Kothari, S. (1989). Nonstationary expected returns: implications for tests of market efficiency and serial correlation in returns. *Journal of Financial Economics, 25*, 51–74.

Baltagi, B. (2005). Econometric Analysis of Panel Data, 3rd ed. Wiley Press.

Baltagi, B. (2008). Econometric Analysis of Panel Data. 4th ed., Wiley.

Baltagi, B. H. (1981). Simultaneous equations with error components. *Journal of Econometrics, 17*, 189–200.

Baltagi, B. H. (2001). *Econometric analysis of panel data*. New York: Wiley.

Baltagi, B. H., & Wu, P. X. (1999). Unequally spaced panel data regressions with Ar(1) disturbances. *Economic Theory, 15*, 814–823.

Baltagi, B. H., & Li, Q. (1992). A note on the estimation of simultaneous equations with error components. *Econometric Theory, 8*, 113–119.

Baltagi, B. H., & Li, Q. (1994). Estimating error component models with general MA (q) disturbances. *Econometric Theory, 10*, 396–408.

Baltagi, B. H., & Li, Q. (1995). Testing AR(1) against MA(1) disturbances in an error component model. *Journal of Econometrics, 68*, 133–151.

Balvers, R. J., Cosimano, T. F., & McDonald, B. (1990). Predicting stock returns in an efficient market. *Journal of Finance, 45*, 1109–1128.

Bandi, F. (2002). Short-Term Interest Rate Dynamics: A Spatial Approach. *Journal of Financial Economics, 65*, 73–110.

Bandyopadhyay, S., Brown, L., & Richardson, G. (1995). Analysts' use of earnings forecasts in predicting stock returns: forecast horizon effects. *International Journal of Forecasting, 11*, 429–445.

Banerjee, A. (1992). A Simple Model of Herd Behavior. *Quarterly Journal of Economics, 107*, 797–818.

Banerjee, A., Dolado, J. J., Galbraith, J. W., & Hendry, D. F. (1993). *Co-integration, Error Correction and the Econometric Analysis of Non-Stationary Data*. Oxford: Oxford University Press.

Bangia, A., Diebold, F., Kronimus, A., Schagen, C., & Schuermann, T. (2002). Rating migrations and the business cycle, with applications to credit portfolio stress testing. *Journal of Banking and Finance, 26*, 445–474.

Bank for International Settlements Staff, (2004). International convergence of capital measurement and capital standards: A revised framework, BIS Staff Study.

Banker, R., & Datar, S. (1989). Sensitivity, precision, and linear aggregation of signals for performance evaluation. *Journal of Accounting Research, 27*(1), 21–39.

Bannister, J., & Newman, H. (1996). Accrual usage to manage earnings toward financial analysts'-forecasts. *Review of Quantitative Finance and Accounting, 7*, 259–278.

Bannouh, K., van Dijk, D., & Martens, M. (2009). Range-based covariance estimation using high-frequency data: the realized co-range. *Journal of Financial Econometrics, 7*, 341–372.

Bansal, R., & Zhou, H. (2002). Term Structure of Interest Rates with Regime Shifts. *Journal of Finance, 57*, 1997–2043.

Bansal, R., Hsieh, D. A., & Viswanathan, S. (1993). A New Approach to International Arbitrage Pricing. *Journal of Finance, 48*, 1719–1747.

Banyi, M., Dyl, E., & Kahle, K. (2008). Errors in estimating share repurchases. *Journal of Corporate Finance, 14*, 460–474.

Bapna, R., Goes, P., & Gupta, A. (2005). Pricing and allocation for quality-differentiated online services. *Management Science, 51*, 1141–1150.

Barber, B., Lehavy, R., McNichols, M., & Trueman, B. (2001). Can investors profit from the prophets? Security analyst recommendations and stock returns. *Journal of Finance, 56*, 531–563.

Barber, B. M., & Lyon, J. D. (1997). Detecting long-run abnormal stock returns: the empirical power and specification of test statistics. *Journal of Financial Economics, 43*, 341–372.

Barber, B., Lehavy, R., & Trueman, B. (2000). Are all brokerage houses created equal? Testing for systematic differences in the performance of brokerage house stock recommendations. University of California at Davis and University of California at Berkeley (unpublished), March.

Barndorff-Nielsen, O., & Shephard, N. (2005). How accurate is the asymptotic approximation to the distribution of realised volatility?, Identification and Inference for Econometric Models, Essays in Honor of Tom Rothenberg, (edited by Don W.K. Andrews and Jim H. Stock), Cambridge University Press, Chapter 13, 306–331.

Barndorff-Nielsen, O., Kent, J., & Sorensen, M. (1982). Normal Variance-Mean Mixtures and z Distributions. *International Statistical Review, 50*, 145–159.

Barndorff-Nielsen, O. E. (1977). Expoenentially decreasing distributions for the logarothm of partical size. *Proceedings of the Royal Society London A, 353*, 401–419.

Barndorff-Nielsen, O. E. (1978). Hyperbolic distributions and distributions on hyperbolae. *Scandinavian Journal of Statistics, 5*, 151–157.

Barndorff-Nielsen, O. E. (1995). Normal inverse Gaussian distributions and the modeling of stock returns. Research Report no. 300, Department of Theoretical Statistics, Aarhus University.

Barndorff-Nielsen, O. E., & Shephard, N. (2001). Non-Gaussian Ornstein-Uhlenbeck based models and some of their uses in financial economics. *Journal of the Royal Statistical Society B, 63*, 167–241.

Barnea, A., Haugen, R. A., & Senbet, L. W. (1980). A rationale for debt maturity structure and call provisions in the agency theoretic framework. *Journal of Finance, 25*, 1223–34.

Barone-Adesi, G., & Whaley, R. (1987). Efficient Analytic Approximation of American Option Values. *Journal of Finance, 42*, 301–320.

Barr, R. S., & Siems, T. F. (1994). Predicting Bank Failure using DEA to Quantify Management Quality. *Federal Reserve Bank of Dallas Financial Industry Studies, 1*, 1–31.

Barsky, R. B. (1989). Why don't the prices of stocks and bonds move together? *American Economic Review, 79*, 1132–1145.

Bartels, L. M. (1991). Instrumental and 'quasi-instrumental' variables. *American Journal of Political Science, 35*, 777–800.

Bartley, J. W., & Cameron, A. B. (1991). Long-run earrings forecasts by managers and financial analysts. *Journal Of Business Finance &Accounting*, January, 21–41.

Bartov, E., & Mohanram, P. (2004). Private information, earnings manipulations, and executive stock-option exercises. *The Accounting Review, 79*(4), 889–920.

Barua, S. K., & Varma, J. R. (1993). *"Speculative Dynamics: The Case of Master shares", Advances in Financial Planning and Forecasting* (Vol. 5). Greenwitch CT, USA: Jai Press.

Barua, A., Kriebel, C., & Mukhopadhyay, T. (1995). Information technology and business value: an analytic and empirical investigation. *Information Systems Research, 6*, 3–23.

Bar-Yosef, S., & Kolodny, R. (1976). Dividend policy and capital market theory. *Review of Economics & Statistics, 58*, 181–190.

Basak, S., & Shapiro, A. (2001). Value- at-Risk Based Risk Management: Optimal Policies and Asset Prices. *The Review of Financial Studies, 14*(2), 371–405.

Basel Committee on Banking Supervision, (1995). Basel Capital Accord: Treatment of Potential Exposure for Off-Balance-Sheet Items, BIS, Basel, Switzerland.
Basel Committee on Banking Supervision, (1998a). Framework for the Evaluation of Internal Control Systems, BIS, Basel, Switzerland.
Basel Committee on Banking Supervision, (1998b). Enhancing Bank Transparency, BIS, Basel, Switzerland.
Basel Committee on Banking Supervision, (1998c). Operational Risk Management, BIS, Basel, Switzerland.
Basel Committee on Banking Supervision, (1999). Credit Risk Modeling: Current Practices and Applications, BIS, Basel, Switzerland.
Basel Committee on Banking Supervision. (2001). The Internal Ratings Based Approach. http://www.bis.org/publ/bcbsca05.pdf.
Basel Committee on Banking Supervision. (2004). Bank Failures in Mature Economies. Working Paper 13.
Basel Committee on Banking Supervision. (2006). International Convergence of Capital Measurement and Capital Standards: A Revised Framework (Basel II).
Basso, A., & Pianca, P. (1997). Decreasing absolute risk aversion and option pricing bounds. *Management Science, 43*, 206–216.
Bastide, D., Benhamou, E., & Ciuca, M. (2007). A Comparative Analysis of Basket Default Swaps Pricing Using the Stein Method. Working Paper, Pricing Partners.
Basu, S. (1977). Investment Performance of Common Stocks in Relation to Thei Price-Earnings Ratio: A Test of the Efficient Market Hypothesis. *Journal of Finance*, 663–682
Batchelor, R., & Peel, D. A. (1998). Rationality testing under asymmetric loss. *Economics Letters, 61*, 49–54.
Batchelor, R., & Dua, P. (1990a). Forecaster ideology, forecasting technique, and the accuracy of economic forecasts. *International Journal of Forecasting, 6*, 3–10.
Batchelor, R., & Dua, P. (1990b). Product differentiation in the economic forecasting industry. *International Journal of Forecasting, 6*, 311–316.
Batchelor, R., & Dua, P. (1992). Conservatism and consensus-seeking among economic forecasters. *Journal of Forecasting, 11*, 169–181.
Bates, D. S. (1996a). Jumps and stochastic volatility: Exchange rate processes implicit in Deutsche Mark options. *Review of Financial Studies, 9*, 69–107.
Bates, D. (1996b). Testing Option Pricing Models. In G. S. Maddala & C. R. Rao (Eds.), *Statistical Methods in Finance (Handbook of Statistics)* (14th ed., pp. 567–611). Amsterdam: Elsevie.
Bates, D. (1996). Jumps and Stochastic Volatility: Exchange Rate Processes Implicit in Deutschemark Options. *Review of Financial Studies, 9*(1), 69–108.
Bates, D. (2000). Post-87 Crash Fears in S&P 500 Futures Options. *Journal of Econometrics, 94*, 181–238.
Bates, J. M., & Granger, C. W. J. (1969). The Combination of Forecasts. *Operations Research Quarterly, 20*, 451–468.
Bauer, G. H., & Vorkink, K. (2011). Forecasting Multivariate Realized Stock Market Volatility. *Journal of Econometrics, 160*, 93–101.
Baum, C., Schaffer, M., & Stillman, S. (2003). Instrumental Variables and GMM: Estimation and Testing. *Stata Journal, 3*, 1–31.
Bauman, M. P., & Shaw, K. W. (2006). Stock option compensation and the likelihood of meeting analysts' quarterly earnings targets. *Review of Quantitative Finance and Accounting, 26*, 301–319.
Bauman, W. S., & Miller, R. E. (1995). Portfolio performance rankings in stock market cycles. *Financial Analysts Journal, 51*, 79–87.
Baumol, W. J., & Fabian, T. (1964). Decomposition, pricing for decentralization and external economies. *Management Science*, 1–32.
Baumol, W. J., & Quandt, R. E. (1965). Investment and discount rate under capital rational—A programming approach. *Economic Journal, 75*(298), 173–329.

Baur, D. G., & Lucey, B. M. (2010). Is gold a hedge or a safe haven? An analysis of stocks, bonds and gold. *Financial Review, 45*(2), 217–29.

Baur, D. G., & McDermott, T. K. (2010). Is gold a safe haven? International evidence. *Journal of Banking and Finance, 34*(8), 1886–1898.

Bauwens, L., Laurent, S., & Rombouts, J. V. K. (2006). Multivariate GARCH models: A survey. *Journal of Applied Econometrics, 21*, 79–109.

Bawa, V. S. (1975). Optimal rules for ordering uncertain prospects. *Journal of Financial Economics, 2*, 95–121.

Bawa, V. S. (1978). Safety-first, stochastic dominance, and optimal portfolio choice. *Journal of Financial and Quantitative Analysis, 13*, 255–271.

Bayer, P., Keohane, N., & Timmins, C. (2009). Migration and hedonic valuation: The case of air quality. *Journal of Environmental Economics and Management, 58*(1), 1–14. July.

BCBS. (2001). Working paper on the regulatory treatment of operational risk. Basel Committee on Banking Supervision, Basel, Switzerland. http://www.bis.org.

BCBS. (2003). The 2002 loss data collection exercise for operational risk: Summary of the data collected. Basel Committee on Banking Supervision, Basel, Switzerland. http://www.bis.org.

Beasley, M. S. (1996). An empirical analysis of the relation between the board of director composition and financial statement fraud. *The Accounting Review, 71*, 443–465.

Beattie, D. L. (1980). Conglomerate Diversification and Performance: A Survey and Applied Economics.251–273.

Beaver, W. (1966). Financial ratios as predictors of failure. *Journal of Accounting Research, 4*, 71–111.

Beaver, W., & Ryan, G. (1987). The Information Content of Security Prices: A Second Look. *Journal of Accounting and Economics, 9*, 139–57.

Beaver, W., Lambert, R., & Morse, D. (1980). The Information Content of Security Prices. *Journal of Accounting and Economics, 2*, 3–28.

Beck, T., Demirguc-Kunt, A., & Levine, R. (2000a). A new database on financial development and structure. *World Bank Economic Review, 14*, 597–605.

Beck, T., Levine, R., & Loayza, N. (2000b). Finance and the sources of growth. *Journal of Financial Economics, 58*, 261–300.

Becker, B. E., & Huselid, M. A. (1992). The incentive effects of tournament compensation systems. *Administrative Science Quarterly, 37*, 336–350.

Beckers, S. (1980). The Constant Elasticity of Variance Model and US Implications for Option Pricing. *Journal of Finance, 35*, 661–673.

Beckers, S. (1983). Variances of security price returns based on highest, lowest, and closing prices. *Journal of Business, 56*, 97–112.

Beckers, S., (1981). Standard Deviation Implied in Option Pricices as Predictors of Future Stock Price Variability. *Journal of Banking and Finance 5*, 363–381, (September).

Beckmann, D., Lutje, T., & Rebeggiani, L. (2007). Italian Asset Managers' Behavior: Evidence on Overconfidence, Risk Taking, and Gender, Discussion Paper No. 358, Leibniz Universitat Hannover, Department of Economics, ISSN: 0949-9962.

Beder, T. S. (1995). VAR: Seductive but Dangerous. Financial *Analysts Journal*, *51*:5(September-October), 12–24.

Begley, J., & Feltham, G. A. (2002). The relations between market value, earnings forecasts, and reported earnings. *Contemporary Accounting Research, 19*, 1–48.

Bekaert, G., & Urias, M. (1996). Diversification, Integration and Emerging Market Closed-end Funds. *Journal of Finance, 51*, 835–869.

Bekaert, G., & Liu, J. (2004). Conditioning information and variance bounds on pricing kernels. *Review of Financial Studies, 17*, 339–378.

Bekaert, G., & Hodrick, R. (1992). Characterizing predictable components in excess returns on equity and foreign exchange markets. *Journal of Finance, 47*, 467–509.

Bekaert, G. (1995). Market Integration and Investment Barriers in Equity Markets. *World Bank Economic Review, 9*, 75–107.

Bell, D., & Ludwig, A. S. (2005). Grid Service Discovery in the Financial Markets Sector., 7th Intertational Conference Information Technology Interfaces , Cavtat, Croatia.

Bellman, R. (1961). Adaptive control processes: a guided tour. *Pattern Recognition Letters, 29*, 1351–1357.

Bellman, R. (1990). *Adaptive Control Process: A Guided Tour*. Princeton, NJ: Princeton University Press.

Bellman, R. (2003). *Dynamic Programming*. Mineola, NY: Dover Publication.

Belsley, D., Kuh, E., & Welsch, R. (1980). *Regression diagnostics: identifying influential data and sources of collinearity*. New York: Wiley.

Benartzi, S., Michaely, R., & Thaler, R. (1997). Do changes in dividends signal the future or the past? *Journal of Finance, 52*, 1007–1034.

Bender, R., & Vorst, T. (2001). Options on Dividends Paying Stocks. Proceeding of the 2001 International Conference on Mathematical Finance, Shanghai, China.

Beneish, D., Billings, M., & Hodder, L. (2008). Internal control weaknesses and information uncertainty. *The Accounting Review, 83*, 665–704.

Benfratello, L., Piacenza, M., & Sacchetto, S. (2009). Taste or reputation: what drives market prices in the wine industry? Estimation of a hedonic model for Italian premium wines. *Applied Economics, 41*(17), 2197–2209.

Benninga, S., & Smirlock, M. (1985). An empirical analysis of the delivery option, marking to market, and the pricing of Treasury bond futures. *Journal of Futures Markets, 5*(3), 361–374.

Benninga, S. (2000). *Financial Modeling*. Cambridge: MIT Press.

Beran, J. (1991). M-estimators of location for data with slowly decaying serial correlations. *Biometrika, 76*, 261–269.

Beran, J., & Terrin, N. (1996). Testing for a change of the long-memory parameter. *Biometrika, 83*, 627–638.

Beresford, D., Johnson, L., & Reither, C. (1996). Is a second income statement needed? *Journal of Accountancy, 176*(April) 69–72.

Berg, D., & Henrik Bakken. (2006). A Copula Goodness-of-Fit Test Based on the Probability Integral Transform. Working Paper at http://www.danielberg.no/publications/Btest.pdf.

Berger, A. N., DeYoung, R., Flannery, M. J., Lee, D., & Oztekin, O. (2008). How do large banking organizations manage their capital ratios? *Journal of Financial Services Research, 343*, 123–149.

Berger, A. N., Herring, R. J., & Szego, G. P. (1995). The role of capital in financial institutions. *Journal of Banking and Finance, 24*, 1383–1398.

Berger, A. N., Miller, N. H., Petersen, M. A., Rajan, R. G., & Stein, J. C. (2005). Does function follow organizational form? evidence from the lending practices of large and small banks. *Journal of Financial Economics, 76*, 237–269.

Bergman, N. K., & Jenter, D. (2007). Employee sentiment and stock option compensation. *Journal of Financial Economics, 84*, 667–712.

Bergstresser, D., & Philippon, T. (2006). CEO incentives and earnings management. *Journal of Financial Economics, 80*, 511–529.

Berkowitz, D., Caner, M., & Ying, F. (2012). The Validity of Instruments Revisited. *Journal of Econometrics., 166*, 255–66.

Berkowitz, J., Christoffersen, P., & Pelletier, D. (2009). Evaluating value-at-risk models with desk-level data, Working Paper 010, North Carolina State University, Department of Economics.

Bernard, R. H. (1969). Mathematical programming models for capital budgeting—A survey generalization and critique. *Journal of Financial and Quantitative Analysis, 4*, 111–158.

Bernard, V., & Thomas, J. (1989). Post-earnings-announcement drift: Delayed price response or risk premium. *Journal of Accounting Research, 27*, 1–36.

Bernard, V. (1995). The Feltham-Ohlson Framework: Implications for Empiricists. *Contemporary Accounting Research, 11*(2), 733–747.

Bernhardt, D., Campello, M., & Kutsoati, E. (2006). Who herds? *Journal of Financial Economics, 80*, 657–675.

Bernheim, B. D., & Wantz, A. (1995). A tax-based test of the dividend signaling hypothesis. *American Economic Review, 85*, 532–551.

Bertoni, A., Brtmetti, G., & Cesari, C. (2005). Mutual Fund Bench Marking and Market Bubbles-A Behavioural Approach. *Transition Studies Review, 12*(1), 36–43.

Bertrand, M., & Schoar, A. (2003). Managing with Style: The Effect of Managers on Firm Policies. *Quarterly Journal of Economics, 118*, 1169–1208.

Bertrand, M., & Mullainathan, S. (2005). Bidding for Oil and Gas Leases in the Gulf of Mexico: A Test of the Free Cash Flow Model? mimeo, University of Chicago and MIT.

Bertrand, M., Duo, E., & Mullainathan, S. (2004). How Much Should We Trust Differences-in-Differences Estimates? *Quarterly Journal of Economics, 119*, 249–275.

Bessler, D., & Ruffley, R. (2004). Prequential analysis of stock market returns. *Applied Economics, 36*(5), 399–412.

Bettis, C. J., Bizjak, J. M., & Lemmon, M. L. (2005). Exercise behavior, valuation, and the incentive effects of employee stock options. *Journal of Financial Economics, 76*, 445–470.

Bevan, A. A., & Danbolt, J. (2002). Capital structure and its determinants in the UK—a decompositional analysis. *Applied Financial Economics, 12*, 159–170.

Beyer, A. (2008). Financial analysts' forecast revisions and managers' reporting behavior. *Journal of Accounting and Economics, 46*, 334–348.

Bharadwaj, A. S. (2000). A resource-based perspective on information technology capability and firm performance: an empirical investigation. *MIS Quarterly, 24*, 169–196.

Bharath, S. T., Pasquariello, P., et al. (2009). Does asymmetric information drive capital structure decisions? *Review of Financial Studies, 22*(8), 3211–3243.

Bhardwaj, G., Corradi, V., & Swanson, N. R. (2008). A Simulation Based Specification Test for Diffusion Processes. *Journal of Business and Economic Statistics, 26*, 176–93.

Bhargava, A., & Sargan, J. D. (1983). Estimating dynamic random effects models from panel data covering short time periods. *Econometrica, 51*, 1635–1660.

Bhargava, A. (1991). Identification and panel data models with endogenous regressors. *Revies of Economic studies, 58*, 129–140.

Bhargava, A. (1994). Modeling the health of Filipino children. *Journal of the Royal Statistical Society A, 157*, 417–432.

Bhargava, A. (2008). Randomized controlled experiments in health and social sciences. *Economics and Human Biology, 6*, 293–298.

Bhargava, A. (2010). An econometric analysis of the dividends and share repurchases by U.S. firms. *Journal of the Royal Statistical Society A, 173*, 631–656.

Bhargava, A. (2012). Executive compensation, share repurchases and investment expenditures: Econometric evidence from U.S. firms. *Review of Quantitative Finance and Accounting*, in press.

Bhaskar, I. C., & McNamee, P. (1983). Multiple objectives accounting and finance. *Journal of Business Finance and Accounting, 10*, 595–621.

Bhaskar, K. A. (1979). Multiple objective approach to capital budgeting. *Accounting and Business Research, 10*, 25–46.

Bhatia, A.V. (2002). Sovereign Credit Risk Ratings: An Evaluation. IMF Working Paper WP/ 03/170.

Bhattacharya, N. (2001). Investors' trade size and trading responses around earnings announcements: An empirical investigation. *Accounting Review, 76*, 221–244.

Bhattacharya, S., & Thakor, A. V. (1993). Contemporary banking theory. *Journal of Financial Intermediation, 3*, 2–50.

Bhattacharya, S., & Pfleiderer, P. (1983). A note on performance evaluation. technical report 714, Stanford University, Graduate school of business.

Bhattacharya, S. (1979). Imperfect information, dividend policy, and "the bird in the hand" fallacy. *Bell Journal of Economics, 10*, 259–270.

Bhattacharyya, M., Chaudhary, A., & Yadav, G. (2008). Conditional VaR estimation using Pearson's type IV distribution. *European Journal of Operational Research, 191*, 386–397.

Bhojrar, S., Lee, C. M. C., & Oler, D. K. (2003). What is my line? A comparison of industry classification schemes for capital market research. *Journal of Accounting Research., 41*, 745–774.

Bhuiyan, R., (2012). Inflationary expectations and monetary policy: evidence from Bangladesh. Empirical Economics, Online First™, 26 April 2012.

Bhushan, R. (1989). Firm characteristics and analyst following, *Journal of Accounting & Economics*, 255–274.

Bianchi, G., & Sorrentino, R. (2007). Electronic filter simulation & design. McGraw-Hill, ISBN 0071494677.

Bibby, B. M., & Sørensen, M. (2003). Hyperbolic Processes in Finance. In: S.T. Rachev (Ed.), *Handbook of Heavy Tailed Distributions in Finance* (pp 211–248). Elsevier Science.

Bickel, P., & Levina, E. (2008). Regularized estimation of large covariance matrices. *Annals of Statistics, 36*, 199–227.

Biddle, G., & Choi, J. H. (2006). Is comprehensive income useful? *Journal of Contemporary Accounting & Economics, 2*(1), 1–32.

Bielecki, T. R., Crepey, S., Jeanblanc, M., & Zargari, B. (2012). Valuation and Hedging of CDS Counterparty Exposure in a Markov Copula Model. *International Journal of Theoretical and Applied Finance, 15*(1), 85–123.

Bierman, H., Jr. (1995). Bubbles, theory and market timing. *Journal of Portfolio Management, 22*, 54–56.

Bikbov, R., & Mikhail Chernov. (2004). Term Structure and Volatility: Lessons from the Eurodollar Markets, Working Paper, Columbia GSB

Bikhchandani, S., Hirshleifer, D., & Welch, I. (1992). A Theory of Fads, Fashion, Custom, and Cultural Change as Informational Cascades. *Journal of Political Economy, 100*, 992–1026.

Billett, M. T., King, T. D., & Mauer, D. C. (2007). Growth opportunities and the choice of leverage, debt maturity, and covenants. *Journal of Finance, 62*, 697–730.

Bingham, N. H., & Kiesel, R. (2001). Modelling asset returns with hyperbolic distribution. In: J. Knight and S.Satchell (Eds.), *Return Distribution on Finance*. (pp. 1–20). Butterworth-Heinemann.

Binkley, C. S., Raper, C. F., & Washburn, C. L. (1996). Institutional ownership of U.S. timberland: History, rationale, and implications for forest management. *Journal of Forestry, 94*, 21–28.

Binkley, C. S., Washbrun, C. L., & Aronow, M. E. (2003). *The NCREIF timberland property index: Research notes 2003*. Boston, MA: Hancock Timber Resource Group.

Biorn, E. (2000). Panel Data with Measurement Errors: Instrumental Variables and GMM Procedures Combining Levels and Differences. *Econometric Reviews, 19*, 391–424.

Bishop, K., & Murphy, A. (2011). Estimating the Willingness to Pay to Avoid Violent Crime: A Dynamic Approach. *The American Economic Review, 101*(3), 625–629.

Bitler, M. P., Moskowitz, T. J., & Vissing-Jorgensen, A. (2005). Testing Agency Theory with Entrepreneur Extra-Effort and Wealth. *Journal of Finance, 60*(2), 539–576.

Bizjak, J., Brickley, J., & Coles, J. (1993). Stock-based incentive compensation and investment behavior. *Journal of Accounting and Economics, 16*(1–3), 349–372.

Black, F. (1976a). The pricing of commodity contacts. *Journal of Financial Economics, 3*, 167–179.

Black, F. (1975). Fact and fantasy in the use of options. *Financial Analysts Journal, 31*(4), 36–41. +61-72.

Black, F. (1993). Choosing accounting rules. *Accounting Horizons 7*(March) 1–17.

Black, F., & Scholes, M. (1974). The effects of dividend yield and dividend policy on common stock prices and returns. *Journal of Financial Economics, 1*, 1–22.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of Political Economy, 81*(3), 637–654.

Black, F., Jensen, M. C., & Scholes, M. (1972). The capital asset pricing model: some empirical tests. In: Jensen, ed., *Studies in the theory of capital Markets* (pp. 79–121). New York: Praeger.

Black, S. W. (1976b). Rational Response to Shocks in a Dynamic model of Capital Asset Pricing. *American Economic Review, 66*, 767–779.

Blæsid, P. (1981). The two-dimensional hyperbolic distribution and related distribution with an application to Johannsen's bean data. *Biometrika, 68*, 251–263.

Blake, C. R., & Morey, M. R. (2000). Morningstar ratings and mutual fund performance. *Journal of Financial & Quantitative Analysis, 35*(3), 451–483.

Blanchard, O., Lopez-de-Silanes, F., & Shleifer, A. (1994). What Do Firms Do with Cash Windfalls? *Journal of Financial Economics, 36*, 337–360.

Blankley, A., Lamb, R., & Schroeder, R. (2000). Compliance with SEC disclosure requirements about market risk. *Journal of Derivatives, 7*, 39–50. Spring.

Blasco, N., & Ferreruela, S. (2007). Comportamiento Imitador en el Mercado Bursátil Español: Evidencia Intradiaria. *Revista de Economía Financiera, 13*, 56–75.

Blasco, N., & Ferreruela, S. (2008). Testing intentional herding in familiar stocks: an experiment in an international context. *The Journal of Behavioral Finance, 9*(2), 72–84.

Blasco, N., Corredor, P., & Ferreruela, S. (2011). Detecting intentional herding: what lies beneath intraday data in the Spanish stock market. *Journal of the Operational Research Society, 62*, 1056–1066.

Blau, B. M., & Fuller, K. P. (2008). Flexibility and dividends. *Journal of Corporate Finance, 14*, 133–152.

Blomquist, G., & Lawrence Worley. (1982). Specifying the Demand for Housing Characteristics: The Exogeniety Issue. In: D.B. Diamond and G. Tolley (eds.), *The Economics of Urban Amenities*. Academic Press.

Blossfeld, H. P., & Rohwer, G. (1995). *Techniques of event history modelling*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

Blossfeld, H.-P., Hamerle, A., & Mayer, K. U. (1989). *Event history analysis: Statistical theory and application in the social sciences*. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

Blume, M. E., & Friend, I. (1973). A new look at the capital asset pricing model. *Journal of Finance, 28*, 19–33.

Blume, M. E., Lim, F., & Mackinlay, A. C. (1998). The declining credit quality of U.S. corporate debt: Myth or Reality? *Journal of Finance, 53*, 1389–1413.

Blume, M., Keim, D., & Patel, S. (1991). Returns and volatility of low-grade bonds 1977-1989. *Journal of Finance, 46*(1), 49–74.

Blundell, R., & Bond, S. (1998). Initial conditions and moment restrictions in dynamic panel data models. *Journal of Econometrics, 87*, 115–143.

Blundell, R., Bond, S., Devereux, M., & Schiantarelli, F. (1992). Investment and Tobin's Q: Evidence from Company Panel Data. *Journal of Econometrics, 51*, 233–257.

Boardman, W., Reinhart, I., & Celec, S. E. (1982). The role of payback period in the theory and application of duration to capital budgeting. *Journal of Business Finance and Accounting, 9*, 511–522.

Bodie, Z. (1976). Common stocks as a hedge against inflation. *Journal of Finance, 31*, 459–740.

Bodie, Z., & Taggart, R. A. (1978). Future investment opportunities and the value of the call provision on a bond. *Journal of Finance, 33*, 1187–1200.

Bodurtha, J. N., Dond-Soon, K., Lee, M. C., & Charles, C. (1995). Closed-end Country Funds and U.S. Market Sentiment. *Review of Financial Studies Fall, 8*, 881–919.

Boehmer, E., Jones, C. M., & Zhang, X. (2010). Shackling short sellers: The 2008 shorting ban, Working paper.

Boehmer, E., Gramming, J., & Theissen, E. (2007). Estimating the probability of informed trading-does trade misclassification matter? *Journal of Financial Markets, 10*, 26–47.

Bohl, S., Martin, T., & Klein, A. (2011). Short Selling Bans and Institutional Investors' Herding Behavior: Evidence from the Global Financial Crisis. Working paper. Available at: http://ssrn.com [Accessed 27 December 2011].

Bollen, N. P., Smith, T., & Whaley, R. E. (2004). Modeling the Bid/Ask Spread: Measuring the Inventory-Holding Premium. *Journal of Financial Economics, 72*, 97–141.

Bollerslev, T. (1986). Generalized autoregressive conditional heteroscedasticity. *Journal of Econometrics, 31*, 307–327.

Bollerslev, T. (1987). A Conditionally Heteroskedastic Time Series Model for Speculative Prices and Rates of Return. *Review of Economics and Statistics, 69*, 542–547.

Bollerslev, T. (1990). Modeling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. *Review of Economics and Statistics, 72*, 498–505.
Bollerslev, T., & Mikkelsen, H. O. (1996). Modeling and pricing long memory in stock market volatility. *Journal of Econometrics, 73*, 151–184.
Bollerslev, T., & Mikkelsen, H. O. (1999). Long-term equity anticipation securities and stock market volatility dynamics. *Journal of Econometrics, 92*, 75–99.
Bollerslev, T., & Wright, J. H. (2000). Semiparametric estimation of long-memory volatility dependencies: the role of high-frequency data. *Journal of Econometrics, 98*, 81–106.
Bollerslev, T., & Wooldridge, J. M. (1992a). Quasi-maximum likelihood estimation and inference in dynamic models with time varying covariances. *Econometric Reviews, 11*, 143–172.
Bollerslev, T., & Wooldridge, J. M. (1992b). Quasi-maximum likelihood estimation and inference in dynamic models with time varying covariances. *Econometric Reviews, 11*, 143–172.
Bollerslev, T. R., Engle, F., & Nelson, D. B. (1994). ARCH Model. In Engle and McFadden, *Handbook of Econometrics*, vol IV. Amsterdam: North Holland.
Bollerslev, T., & Zhou, H. (2006). Volatility puzzles: a simple framework for gauging return-volatility regressions. *Journal of Econometrics, 131*, 123–150.
Bollerslev, T., & Domowitz, I. (1993). Trading patterns and prices in the interbank foreign exchange market. *Journal of Finance, 48*, 1421–1443.
Bollerslev, T., Engle, R. F., & Wooldridge, J. M. (1988). A capital asset pricing model with time-varying covariances. *Journal of Political Economy, 96*, 116–131.
Bollerslev, T., Engle, R. F., & Nelson, D. B. (1993). ARCH Models. Department of Finance, Kellog Graduate School of Management, Northwestern University, mimeo.
Bollerslev, T., Chou, R. Y., & Kroner, K. F. (1992). ARCH modeling in finance: a review of the theory and empirical evidence. *Journal of Econometrics, 52*, 5–59.
Bond, S., & Cummins, J. (2000). The Stock Market and Investment in the New Economy. *Brookings Papers on Economic Activity, 13*, 61–124.
Bonfim, D. (2009). Creditrisk drivers: Evaluating the contribution of firm level information and of macroeconomic dynamics. *Journal of Banking and Finance, 33*, 281–299.
Bonham, C. S., & Cohen, R. H. (1995). Testing the rationality of price forecasts: Comment. *American Economic Review, 85*(1), 284–289.
Bonham, C. S., & Cohen, R. H. (2001). To aggregate, pool, or neither: Testing the rational-expectations hypothesis using survey data. *Journal of Business & Economic Statstics, 19*, 278–291.
Boni, L., & Womack, K. (2006). Analysts, industries, and price momentum. *Journal of Financial Quantitative Analysis, 41*, 85–109.
Bonser-Neal, C., Brauer, G., Neal, R., & Wheatly, S. (1990). International Investment Restrictions and Closed-End Country Fund Prices. *Journal of Finance, 45*, 523–547.
Bontemps, C., & Meddahi, N. (2005). Testing Normality: a GMM Approach. *Journal of Econometrics, 124*, 149–186.
Bontemps, C., Simioni, M., & Surry, Y. (2008). Semiparametric Hedonic Price Models: Assessing The Effects Of Agricultural Nonpoint Source Pollution. *Journal of Applied Econometrics, 23*(6), 825–842.
Bookstein, A., & Swanson, D. (1974). Probabilistic Models for Automatic Indexing. *Journal of the American Society for Information Science, 25*(5), 312–318.
Booth, J. R., & Officer, D. T. (1985). Expectations, Interest Rates, and Commercial Bank Stocks. *Journal of Financial Research, 8*, 51–58.
Boothe, P., & Glassman, D. (1987a). Comparing exchange rate forecasting models. *International Journal of Forecasting, 3*, 65–79.
Boothe, P., & Glassman, D. (1987b). The Statistical Distribution of Exchange Rates. *Journal of International Economics, 22*, 297–319.
Borkowski, S. C. (1990). Environmental and organizational factors affecting transfer pricing: a survey. *Journal of Management Accounting Research*, 78–99.
Born, K. E. (1983). *'International Banking in the 19th and 20th Centuries.* New York: St. Martin's Press.

Boros, E., Hammer, P. L., Ibaraki, T., & Kogan, A. (1997). Logical Analysis of Numerical Data. *Mathematical Programming, 79*, 163–190.

Boros, E., Hammer, P. L., Ibaraki, T., Kogan, A., Mayoraz, E., & Muchnik, I. (2000). An Implementation of Logical Analysis of Data. *IEEE Transactions on Knowledge and Data Engineering, 12*(2), 292–306.

Bos, M., & Vandermark, S. (2002). Finessing Fixed Dividends. *Risk, 15*, 157–158.

Botlen, N., & Busse, J. (2004). Short-term persistence in mutual fund performance. *Review of Financial Studies, 8*(2), 569–597.

Botteron, P., Chesney, M., & Gibson-Anser, R. (2003). Analyzing Firms Strategic Investment Decisions in a Real Options Framework. *International Financial Markets, Institutions and Money, 13*, 451–479.

Bouchaud, J. P., & Potters, M. (2004). *Theory of Financial Risk and Derivative Pricing: From Statistical Physics to Risk Management*. Cambridge University Press.

Boucher, T. O., & MacStravic, E. L. (1991). Multi-attribute evaluation within a present value framework and its relation to the analytic hierarchy process. *The Engineering Economist, 37*, 1–32.

Bouchet, M. H., Clark, E., & Groslambert, B. (2003). *Country Risk Assessment: A Guide to Global Investment Strategy*. Chichester, England: Wiley.

Boudoukh, J., & Whitelaw, R. (1991). The benchmark effect in the Japanese government bond market. *The Journal of Fixed Income, 1*, 52–59.

Boudoukh, J., & Whitelaw, R. (1993). Liquidity as a choice variable: A lesson from the Japanese government bond market. *Review of Financial Studies, 6*, 265–292.

Boudoukh, J., Richardson, M., & Whitelaw, R. F. (1997a). Nonlinearities in the relation between the equity risk premium and the term structure. *Management Science, 43*, 371–385.

Boudoukh, J., Richardson, M., & Whitelaw, R. F. (1997b). Investigation of a class of volatility estimators. *Journal of Derivatives, 1997*, 63–71. Spring.

Boudoukh, J., Richardson, M., & Whitelaw, R. F. (1998). The best of both worlds. *Risk, 11*, 64–67.

Bourke P., Shanmugam B. (1990). An Introduction to Bank Lending. Addison-Wesley Business Series.

Bouwman, M., & Frishkoff, P. (1995). The relevance of GAAP-based information: A case study exploring some uses and limitations. *Accounting Horizons, 9*, 22–47.

Bouye, E. V., Durrleman, A., Nikeghbali, G. R., & Roncalli, T. (2000). Copulas for finance—a reading guide and some application, Groupe de Recherche.

Bouyè, E., & Salmon, M. (2009). Dynamic copula quantile regressions and tail area dynamic dependence in Forex markets. *European Journal of Finance, 15*, 721–750.

Bouyé, E., Valdo Durrleman, Ashkan Nikeghbali, Gaël Riboulet, & Thierry Roncalli. (2000). Copulas for Finance – A Reading Guide and Some Applications. Available at SSRN: http://ssrn.com/abstract=1032533.

Bowden, R. J. (1978). *The Econometrics of Disequilibrium*. Amsterdam: North Holland.

Bowe, M., & Domuta, D. (2004). Investor herding during financial crisis: A clinical study of the Jakarta Stock Exchange. *Pacific-Basin Finance Journal, 12*(4), 387–418.

Bowen, R. M., Daley, L. A., & Huber, C. C., Jr. (1982). Evidence on the existence and determinants of inter-industry differences in leverage. *Financial Management, 4*, 10–20.

Bower, D. H., Richard, S. B., & Dennis, E. L. (1984). Arbitrage Pricing Theory and Utility Stock Returns. *Journal of Finance, 39*, 1041–1054.

Box, G. E. P., & Tiao, G. C. (1975). Intervention Analysis with Applications to Economic and Environmental Problem. *Journal of the American Statistical Association, 70*, 70–79.

Box, G. E., Jenkins, G. W., & Reinsel, G. (1994). *Time Series Analysis, Forecasting and Control* (3rd ed.). Englewood Cliffs, N. J.: Prentice Hall.

Box, G. E. P., & Cox, D. (1964). An Analysis of Transformations. *Journal of the American Statistical Association, Society Series B, 26*, 211–52.

Box, G. E. P., & Jenkins, G. M. (1976). Time Series Analysis: Forecasting and Control, Revised Edition, Oakland, CA: Holden-Day.

Boyer, B., Mitton, T., & Vorkink, K. (2010). Expected idiosyncratic skewness. *Review of Financial Studies, 23*, 169–202.

Boyer, M. M., & Fillion, D. (2007). Common and fundamental factors in stock returns of Canadian oil and gas companies. *Energy Economics, 29*, 428–453.

Boyle, P. (1989). The quality option and timing option in futures contracts. *Journal of Finance, 44*(1), 101–113.

Boyle, P., & Lin, X. (1997). Bounds on contingent claims based on several assets. *Journal of Financial Economics, 46*, 383–400.

Brace, A., Gatarek, D., & Musiela, M. (1997). The Market Model of Interest Rate Dynamics. *Mathematical Finance, 7*, 127–155.

Brander, L., & Koeste, M. (2011). The value of urban open space: Meta-analyses of contingent valuation and hedonic pricing results. *Journal of Environmental Management, 92*(10), 2763–2773.

Brandt, M., & Jones, C. (2006). Volatility forecasting with range-based EGARCH models. *Journal of Business and Economic Statistics, 24*, 470–486.

Brandt, M., & Diebold, F. (2006). A no-arbitrage approach to range-based estimation of return covariances and correlations. *Journal of Business, 79*, 61–74.

Brandt, M. W., & Kang, Q. (2004). On the relationship between the conditional mean and volatility of stock returns: A latent VAR approach. *Journal of Financial Economics, 72*, 217–257.

Brandt, M. W., David Chapman. Comparing Multifactor Models of the Term Structure. Working paper.

Bratley, P., & Fox, B. L. (1988). Algorithm 659: Implementing Sobol's Quasirandom Sequence Generator. *ACM Trans Math. Softw., 14*, 88–100.

Brav, A. (2000). Inference in long-horizon event studies: A Bayesian approach with application to initial public offerings. *Journal of Finance, 55*, 1979–2016.

Brav, A., & Gompers, P. A. (1997). Myth or reality? the long-run underperformance of initial public offerings: evidence from venture and nonventure capital-backed companies. *Journal of Finance, 52*, 1791–1821.

Brav, A., Geczy, C., & Gompers, P. A. (2000). Is the abnormal return following equity issuances anomalous? *Journal of Financial Economics, 56*, 209–249.

Brav, A., Graham, J. R., Harvey, C. R., & Michaely, R. (2005a). Payout policy in the 21st century. *Journal of Financial Economics, 77*, 483–527.

Brav, A., Lehavy, R., & Michaely, R. (2005b). Using expectations to test asset pricing models. *Financial Management, 34*(3), 31–64.

Brealey, R. A., Myers, S. C., & Allen, F. (2008). "Principles of corporate finance". McGraw-Hill, New York, 9th Ed.

Breannan, M. (1991). A Perspective on Accounting and Stock Prices. *The Accounting Review, 66*, 67–79.

Breeden, D. T. (1979). An intertemporal asset pricing model with stochastic consumption and investment opportunities. *Journal of Financial Economics, 7*, 265–296.

Breeden, D. T., & Litzenberger, R. H. (1978). Prices of state-contingent claims implicit in option prices. *Journal of Business, 51*, 621.651.

Breen, W., Glosten, L., & Jagannathan, R. (1989). Economic significance of predictable variations in stock index returns. *Journal of Finance, 44*, 1177–1189.

Breidt, F. J., Crato, N., & de Lima, P. (1998). The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics, 83*, 325–348.

Breiman, L. (1996). Heuristics of instability and stabilization in model selection. *Annals of Statistics, 24*, 2350–2383.

Breitung, J., Bruggemann, R., & Lutkepohl, H. (2004). Structural vector autoregressive modelling and impulse responses. In: H. Lutkepohl and M. Kratzig (Eds.), *Applied Time Series Econometrics*.

Brennan, M. (1993). Agency and asset prices. Finance Working Paper No. 6-93, UCLA.

Brennan, M., & Cao, H. (1997). International portfolio investment flows. *Journal of Finance, 52*, 1851–1880.

Brennan, M. J., & Schwartz, E. S. (1979). A Continuous-Time Approach to the Pricing Bonds. *Journal of Banking and Finance, 3*, 133–155.

Brennan, M. J. (1973). An approach to the valuation of uncertain income streams. *Journal of Finance, 28*, 661–674.

Brennan, M. J. (1974). An inter-temporal approach to the optimization of dividend policy with predetermined investments: Comment. *Journal of Finance, 29*, 258–259.

Brennan, M., Chordia, T., et al. (1998). Alternative factor specifications, security characteristics, and the cross-section of expected stock returns. *Journal of Financial Economics, 49*(3), 345–373.

Brennan, M. J., & Schwartz, E. S. (1985). Evaluating Natural Resource Investments. *Journal of Business, 58*, 135–157.

Brennan, M. J., & Schwartz, E. S. (1988). The case for convertibles. *Journal of Applied Corporate Finance, 1*, 55–64.

Brennan, M. J., & Subrahmanyam, A. (1996). Market Microstructure and Asset Pricing: On the Compensation for Illiquidity in Stock Returns. *Journal of Financial Economics, 41*, 441–464.

Brenner, M., & Subrahmanyam, M. G. (1988). A Simple Formula to Compute the Implied Standard Deviation. *Financial Analysts Journal, 44*(80–83).

Brenner, M., Courtadon, G., & Subrahmanyam, M. (1985). Options on the spot and options on futures. *Journal of Finance, 40*(5), 1303–1317.

Brenner, R., Harjes, R., & Kroner, K. (1996). Another Look at Alternative Models of Short-Term Interest Rate. *Journal of Financial and Quantitative Analysis, 31*, 85–107.

Breusch, T. S., & Pagan, A. R. (1980). The lagrange multiplier test and its application to model specifications in econometrics. *Review of Economic Studies, 47*, 239–253.

Breusch, T. S. (1978). Testing for auitocorrelation in dynamic linear models. *Australian Economic Papers, 17*, 334–355.

Brewer, T. L., & Rivoli, P. (1990). Politics and Perceived Country Creditworthiness in International Banking. *Journal of Money, Credit and Banking, 22*, 357–369.

Brewer, T. L., & Rivoli, P. (1997). Political Instability and Country Risk. *Global Finance Journal, 8*(2), 309–321.

Brick, I. E., Palmon, O., & Patro, D. K. (2011). The motivations for issuing putable debt: An empirical analysis. In: *Handbook of Quantitative Finance and Econometrics*, forthcoming.

Brick, I. E., & Wallingford, B. A. (1985). The relative tax benefits of alternative call features in corporate debt. *Journal of Financial and Quantitative Analysis, 20*, 95–106.

Brick, I. E., & Palmon, O. (1993). The tax advantages of refunding debt by calling, repurchasing, and putting. *Financial Management, 22*, 96–106.

Brickley, J. A., Smith, C. W., & Zimmerman, J. L. (2001). *Managerial economics and organizational architecture*. Irwin/McGraw-Hill; Burr Ridge, IL.

Brief, R., & Peasnell, K. (Eds.). (1996). *Clean Surplus: A Link Between Accounting and Finance*. New York: Garland Publishing.

Brier, G. (1950). Verification of forecasts expressed in terms of probability. *Monthly Weather Review, 78*(1), 1–3.

Brigham, E. F., & Daves, P. (2007). *Intermediate Financial management*. Thomson South: West.

Brigham, S. (1983). *Fundamentals of Financial Management* (3rd ed.). New York: Dryden Press.

Brigo, D., & Chourdakis, K. (2009). Counterparty Risk for Credit Default Swaps: Impact of Spread Volatility and Default Correlation. *International Journal of Theoretical and Applied Finance, 12*(7), 1007–1026.

Broadie, M., Sundaresan, S. (1987). The pricing of timing and quality options: An application to Treasury bond futures markets. Working Paper.

Broadie, M., & Cao, M. (2008). Improved lower and upper bound algorithms for pricing American options by simulation. *Quantitative Finance, 8*, 845–861.

Broadie, M., & Kaya, Ö. (2006). Exact simulation of stochastic volatility and other affine jump diffusion processes. *Operations Research, 54*(2), 217–231.

Broadie, M., Chernov, M., & Johannes, M. (2007). Model speci_cation and risk premia: Evidence from futures options. *Journal of Finance, 62*, 1453–1490.
Brockwell, P. J., & Davis, R. A. (1991). Time Series: Theory and methods Busetti, F. and A. M. R. Taylor. 2004. "Tests of stationarity against a change in persistence. *Journal of Econometrics, 123*, 33–66.
Brockwell, P. J., & Davis, R. A. (1991). *Time Series: Theory and Methods* (2nd ed.). New York: Springer.
Broll, U., Wahl, J. E., & Wong, W. K. (2006). Elasticity of risk aversion and international trade. *Economics Letters, 91*, 126–130.
Broll, U., Egozcue, M., Wong, W. K., & Zitikis, R. (2010). Prospect theory, indifference curves, and hedging risks. *Applied Mathematics Research Express, 2010*, 142–153.
Brook, Y., Charlton, W. T., Jr., & Hendershott, R. J. (1998). Do firms use dividends to signal large future cash flow increases? *Financial Management, 27*, 46–57.
Brooks, C. (2008). *"Introductory econometrics for finance", Chap 8*. UK: Cambridge University Press.
Brooks, C., Burke, S., Heravi, S., & Persand, G. (2005). Autoregressive Conditional Kurtosis. *Journal of Financial Econometrics, 3*, 399–421.
Broto, C., & Ruiz, E. (2004). Estimation methods for stochastic volatility models: A survey. *Journal of Economic Surveys, 18*(5), 613–649.
Brown, L. (2001). A temporal analysis of earnings surprises: Profit vs. losses. *Journal of Accounting Research, 39*(2), 221–241.
Brown, J. R., Ivkovic, Z., Smith, P. A., & Weisbenner, S. (2008). Neighbors matter: causal community effects and stock market participation. *Journal of Finance, 63*, 1509–1531.
Brown, L., & Sivakumar, K. (2003). Comparing the quality of two operating income measures. *Review of Accounting Studies, 4*, 561–572.
Brown, L. D., & Higgins, H. N. (2001). Managing earnings surprises in the U.S. versus 12 other countries. *Journal of Accounting and Public Policy, 20*, 371–398.
Brown, P., Foster, G., & Noreen, E. (1985). Security Analyst Multiyear Earnings Forecasts and the Capital Market, Monograph, American Accounting Association.
Brown, R., Chan, H., & Ho, Y. (2009). Analysts' recommendations: from which signal does the market take its lead? *Review of Quantitative Finance and Accounting, 33*, 91–111.
Brown, S. (1990). Estimating Volatility. In: Figlewski, et al, eds., *Financial Options: From Theory to Practice* (pp 516–537). Homewood Ill: Business One Irwin.
Brown, S. J., & Warner, J. B. (1980). Measuring security price performance. *Journal of Financial Economics, 8*, 205–258.
Brown, S. J., & Warner, J. B. (1985a). Using daily stock returns: the case of event studies. *Journal of Financial Economics, 14*, 3–31.
Brown, S., & Warner, J. (1985b). Using Daily Stock Returns: The Case of Event Studies. *Journal of Financial Economics, 14*, 3–31.
Brown, S. J., & Goetzmann, W. N. (1997). Mutual fund styles. *Journal of Financial Economics, 43*(3), 373–399.
Brown, S., Hillegeist, S., & Lo, K. (2004). Conference calls and information asymmetry. *Journal of Accounting and Economics, 37*, 343–366.
Brownlees, C. T., & Engle, R. (2010). Volatility, correlation and tails for systemic risk measurement, Working paper, NYU Stern School of Business.
Brownlees, C. T., & Giampiero, G. M. (2010). Comparison of volatility measures: A risk management perspective. *Journal of Financial Econometrics, 8*, 29–56.
Brownlees, C. T., & Gallo, G. M. (2006). Financial econometric analysis at ultra-high frequency: Data handling concerns. *Computational Statistics & Data Analysis, 51*, 2232–2245.
Brownstone, D., & Valletta, R. (2001). The bootstrap and multiple imputations: Harnessing increased computing power for improved statistical tests. *Journal of Economic Perspectives, 15*, 129–141.
Brunnermeier, M., & Pedersen, L. H. (2008). Market liquidity and funding liquidity. *Review of Financial Studies, 22*, 2201–2238.

Brunnermerier, M. K., Nage, S., & Pedersen, L. H. (2008). Carry trades and currency crashes, Working Paper in NBER Macroeconomics Annual.

Bryan, S., Hwang, L., & Lilien, S. (2000). CEO stock-based compensation: an empirical analysis of incentive-intensity, relative mix and economic determinants. *Journal of Business, 73*, 661–693.

Bryant, R. C. (1995). "The "exchange risk premium," uncovered interest parity, and the treatment of exchange rates in multi-country macroeconomic models." Technical report, The Brookings Institution

Brynjolfsson, E., & Hitt, L. M. (1996). Paradox lost? Firm-level evidence on the returns to information systems. *Management Science., 42*, 541–558.

Brynjolfsson, E. (1993). The productivity paradox of information technology. *Communications of the ACM, 35*, 66–77.

Brynjolfsson, E., & Yang, S. (1996). Information technology and productivity: A review of the literature. *Advances in Computers., 43*, 179–214.

Bubnys, E. L. (1990). Simulating and Forecasting Utility Stock Returns: Arbitrage Pricing Theory vs. Capital Asset Pricing Model. *The Financial Review, 29*, 1–23.

Bughin, J., Chui, M., & Manyika, J. (2010). Cloud, Big Data and Smart Assets: Ten Tech-enabled Business Trends to Watch, McKinsey Quarterly August, 7–8.

Bun, M. J. G., & Windmeijer, F. (2010). The weak instrument problem of the system GMM estimator in dynamic panel data models. *Econometrics Journal, 13*, 95–126.

Bunn, D., & Karakatsani, N. (2003). Forecasting Electricity Prices, Working Paper, London Business School.

Burgstahler, D., & Dichev, I. (1997). Earnings management to avoid earnings decreases and losses. *Journal of Accounting and Economics, 24*, 99–126.

Burmeister, E., & McElroy, M. B. (1991). The Residual Market Factor, the APT, and Mean-Variance Efficiency. *Review of Quantitative Finance and Accounting, 1*, 27–49.

Burnside, C., Eichenbaum, M., & Rebelo, S. (2007). The returns to currency speculation in emerging markets. *American Economic Review, 97*, 333–338.

Burnside, C., Eichenbaum, M., & Rebelo, S. (2009). Understanding the forward premium puzzle: A microstructure approach. *American Economic Journal: Macroeconomics, American Economic Association, 1*(2), 127–54.

Burnside, C., Eichenbaum, M., Kleshchelski, I., & Rebelo, S. (2011). Do peso problems explain the returns to the carry trade? *Review of Financial Studies, Oxford University Press for Society for Financial Studies, 24*(3), 853–891.

Buser, S. A., Chen, A. H., & Kane, E. J. (1981). Federal deposit insurance, regulatory policy, and optimal bank capital. *Journal of Finance, 36*, 51–60.

Bushman, R., & Indjejikian, R. (1993). Accounting income, stock price, and managerial compensation. *Journal of Accounting and Economics, 16*(1–3), 3–23.

Butler, A. W., & Wan, H. (2010). Stock market liquidity and the long-run stock performance of debt issuers. *Review of Financial Studies, 23*, 3966–3995.

Byard, D., Li, Y., & Weintrop, J. (2006). Corporate governance and the quality of financial analysts' information. *Journal of Accounting and Public Policy, 25*, 609–625.

Caballero, R., & Krishnamurthy, A. (2001). International and Domestic Collateral Constraints in a Model of Emerging Market Crises. *Journal of Monetary Economics, 48*(3), 513–548.

Caballero, R., & Krishnamurthy, A. (2004). Fiscal Policy and Financial Depth. National Bureau of Economic Research (NBER) Working Papers 10532.

Cabedo, J. D., & Moya, I. (2003). Estimating oil price 'Value at Risk' using the historical simulation approach. *Energy Economics, 25*, 239–253.

Cadsby, C. B. (1986). Performance hypothesis testing with the Sharpe and Treynor measures: a comment. *Journal of Finance, 41*, 1175–1176.

Caginalp, G., Porter, D., & Smith, V. (2001). Financial bubbles: Excess cash, momentum, and incomplete information. *The Journal of Psychology and Financial Markets, 2*(2), 80–99.

Cahan, S., Courtenay, S., Gronewoller, P., & Upton, D. (2000). Value relevance of mandated comprehensive income disclosures. *Journal of Business, Finance, and Accounting 27*(November/December), 1273–1301.

Cai, F., & Zheng, L. (2002). Institutional trading and stock returns, Working paper, University of Michigan Business School, Ann Arbor.

Cai, J. (1994). A Markov model of switching-regime ARCH. *Journal of Business and Economic Statistics, 12*, 309–316.

Cai, L., & Swanson, N. R. (2011). An Empirical Assessment of Spot Rate Model Stability. *Journal of Empirical Finance, 18*, 743.764.

Cai, Y., Chou, R. Y., & Li, D. (2009). Explaining international stock correlations with CPI fluctuations and market volatility. *Journal of Banking and Finance, 33*, 2026–2035.

Cai, Z., & Wang, X. (2008). Nonparametric estimation of conditional VaR and expected shortfall. *Journal of Econometrics, 147*, 120–130.

Callen, J. L., & Morel, M. (2001). Linear Accounting Valuation When Abnormal Earnings Are AR (2). *Review of Quantitative Finance and Accounting, 16*, 191–203.

Calomiris, C. W. (1999). Building an incentive-compatible safety net. *Journal of Banking and Finance, 23*, 1499–1519.

Calomiris, C. W., & Kahn, C. M. (1991). The role of demandable debt in structuring optimal banking arrangements. *American Economic Review, 81*, 497–513.

Calvert, J. (1990). Quality customer service: a strategic marketing weapon for high performance banking. *Industrial Engineering*, 54–57.

Calvert, L. E., Fischer, A. J., & Thompson, S. B. (2006). Volatility comovement: A multi-frequency approach. *Journal of Econometrics, 131*(1–2), 179–215.

Calvo, G., & Mendoza, E. (2000). Rational Contagion and the Globalization of Securities Markets. *Journal of International Economics, 51*(1), 79–113.

Cambell, J. Y., Lo, A. W., & Mackinlay, A. C. (1997). *The Econometrics of Financial Markets*. Princeton: Princeton University Press.

Camerer, C. (1989). Bubbles and fads in asset price: A review of theory and evidence. *Journal of Economic Surveys, 3*, 3–41.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2006). Robust inference with multiway Clustering, Working paper.

Cameron, A. C., Gelbach, J. B., & Miller, D. L. (2008). Bootstrap-based improvements for inference with clustered errors. *Review of Economics and Statistics, 90*, 414–427.

Campa, J. M., & Kedia, S. (2002). Explaining the diversification discount. *Journal of Finance, 57*, 1731–1762.

Campbell, J., & Shiller, R. (1988). Stock Prices, Earnings, and Expected Dividends. *Journal of Finance*, 661–676.

Campbell, J. Y., Lo, A. W., & MacKinlay, A. C. (1997). *The econometrics of financial markets*. Princeton, New Jersey: Princeton University Press.

Campbell, J. Y. (1987). Stock returns and the term structure. *Journal Financial Economics, 18*, 373–399.

Campbell, J. Y., & Hentschel, L. (1992). No News is Good News: An asymmetric model of changing volatility in stock returns. *Journal Financial Economics, 31*, 281–318.

Campbell, J. Y., & Perron, P. (1991). Pitfalls and opportunities: what acroeconomists should know about unit roots. *National Bureau of Economic Research Macroeconomics Annual*, 141–201.

Campbell, J. Y. (1991). A variance decomposition for stock returns. *Economic Journal, 101*, 157–179.

Campbell, J. Y., & Shiller, R. (1988). The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies, 1*, 195–228.

Campbell, J. Y., Grossman, S. J., & Wang, J. (1993). Trading Volume and Serial Correlation in Stock Returns. *Quarterly Journal of Economics, 108*, 905–939.

Campbell, J., Lo, A., & MacKinlay, C. (1996). *The Econometrics of Financial Markets*. Princeton: Press Princeton University.

Campbell, J. Y., & Hentschel, L. (1992). No news is good news: an asymmetric model of changing volatility in stock returns. *Journal of Financial Economics, 31*, 281–381.
Campbell, J., & Viceira, L. (1999). Consumption and portfolio decisions when expected returns are time varying. *Quarterly Journal of Economics, 114*, 433–495.
Campbell, J., & Shiller, R. (1987). Cointegration and tests of present value models. *Journal of Political Economy, 95*, 1062–1088.
Campbell, J., & Shiller, R. (1989). The dividend-price ratio and expectations of future dividends and discount factors. *Review of Financial Studies, 1*, 195–228.
Campbell, J., Chan, Y. L., & Viceira, L. (2003). A multivariate model of strategic asset allocation. *Journal of Financial Economics, 67*, 41–80.
Koedijk, C. R. R. H. K. (2001). Optimal Portfolio Selection in a Value-at- Risk Framework. *Journal of Banking and Finance, 25*, 1789–1804.
Campello, M. (2006). Debt financing: does It boost or hurt firm performance in product markets? *Journal of Financial Economics, 82*, 135–172.
Campi, J. P. (1992). It's not as easy as ABC. *Journal of Cost Management Summer*, 5–11.
Candes, E., & Tao, T. (2007). The Dantzig selector: statistical estimation when p is much larger than n (with discussion). *Annals of Statistics, 35*, 2313–2404.
Cantoni, E., & Ronchetti, E. (2001). Resistant selection of the smoothing parameter for smoothing splines. *Statistics and Computing, 11*(2), 141–146.
Cantor, R. (2004). An Introduction to Recent Research on Credit Risk Rating. *Journal of Banking and Finance, 28*, 2565–2573.
Cantor R., Packer F. (1996). Determinants and Impact of Sovereign Credit Ratings. *FRBNY Economic Policy Review*, 37–53.
Cantor, R., & Packer, F. (1997). Differences of opinion and selection bias in the credit rating industry. *Journal of Banking and Finance, 21*, 1395–417.
Cao, J., & Laksmana, I. (2010). The effect of capital market on the association between R&D spending and CEO option compensation. *Review of Quantitative Finance and Accounting, 34*, 273–300.
Cao, C., & Huang, J. (2008). Determinants of S&P 500 Index Option Returns. *Review of Derivatives Research, 10*, 1–38.
Cao, C., Chen, Z., & Griffin, J. M. (2003). The informational content of option volume prior to takeovers. *Journal of Business, 78*, 1071–1109.
Capie, F., Mills, T. C., & Wood, G. (2005). Gold as a hedge against the dollar. *Journal of International Financial Markets, Institutions and Money, 15*(4), 343–352.
Cappiello, L., Engle, R., & Sheppard, K. (2006). Asymmetric dynamics in the correlations of global equity and bond returns. *Journal of Financial Econometrics, 4*, 537–572.
Carchano, O., & Pardo, A. (2009). Rolling over stock index futures contracts. *Journal of Futures Markets, 29*, 684–694.
Carey, M., Post, M., & Sharp, S. (1998). Does corporate lending by banks and finance companies differ? Evidence on specialization in private debt contracting. *Journal of Finance, 53*, 845–878.
Carhart, M. (1997). On persistence in mutual fund performance. *Journal of Finance, 52*, 57–82.
Carlstein, E., Mueller, H. G., & Siegmund, D. (1994). *Change-point problems*. Hay ward, CA: Institute of Mathematical Statistics.
Carpenter, J. N. (2000). Does option compensation increase managerial risk appetite? *Journal of Finance, 55*, 2311–2331.
Carr, P. (1988). Valuing bond futures and the quality option. Working Paper, University of California at Los Angeles.
Carr, P., & Chen, R. (1996). Valuing bond futures and the quality option. Working paper, Rutgers University, November.
Carr, P., Geman, H., Madan, D. B., & Yor, M. (2003). Stochastic volatility for Levy processes. *Mathematical Finance, 13*, 345–382.
Carroll, R., & Härdle, W. (1989). Symmetrized nearest neighbor regression estimates. *Statistics and Probability Letters, 7*, 315–318.

Carter, R., & Manaster, S. (1990). Initial public offerings and underwriter reputation. *Journal of Finance, 45*, 1045–1068.

Carter, R., Dark, F., & Singh, A. (1998). Underwriter reputation, initial returns, and the long-run performance of IPO stocks. *Journal of Finance, 53*, 285–311.

Carty, L. V. (1997). Moody's rating migration and credit quality correlation, 1920–1996. Special Comment (July 1997), Moody's Investor Service.

Carty, L. V., & Fons, J. S. (1994). Measuring changes in corporate credit quality. *Journal of Fixed Income, 4*(1), 27–41.

Casassus, J., Collin-Dufresne, P., & Goldstein, R. S. (2005). Unspanned Stochastic Volatility and Fixed Income Derivative Pricing. *Journal of Banking and Finance, 29*, 2723–2749.

Cascio, A. J., & Clutter, M. L. (2008). Risk and required return assessments of equity timberland investments in the United States. *Forest Products Journal, 58*, 61–70.

Casella, G., & Edward, G. (1992). Explaining the Gibbs sapmler. *The Americain Statistician, 46*(3), 167–174.

Cason, T., & Noussair, C. (2001). The Experimental Study of Market Behavior, Advances in Experimental Markets, Springer Verlag, 1–14

Caton, G., & Goh, J. (2003). Are all rivals affected equally by bond rating downgrades? *Review of Quantitative Finance and Accounting, 20*, 49–62.

Caulfield, J. P. (1998). Timberland return drivers and investing styles for an asset that has come of age. *Real Estate Finance, 14*, 65–78.

Cavaglia, S. M. F. G., Verschoor, W. F. C., & Wolff, C. C. P. (1994). On the biasedness of forward foreign exchange rates: Irrationality or risk premia? *Journal of Business, 67*, 321–343.

Cavaliere, G. (2004). Unit root tests under time-varying variances. *Econometric Reviews, 23*, 259–292.

Cavaliere, G., & Taylor, A. M. R. (2006). Testing for unit roots in time series models with non-stationary volatility. *Journal of Econometrics, 140*, 919–947.

Cavanagh, C., & Sherman, R. P. (1998). Rank Estimators for Monotonic Index Models. *J. Econometrics, 84*(1), 351–381.

Cecchetti, S. G., Lam, P., & Mark, N. C. (1990). Mean reversion in equilibrium asset prices. *American Economic Review, 80*, 398–418.

Cecchetti, S. (1999). Distinguishing theories of Monetary Transmission Mechanism. Federal Reserve of Saint-Louis Review. May-June, 83–97.

Cehng-Ru, W., Chang, H.-Y., & Li-Syuan, W. (2008). A Framework of assessable mutual fund performance. *Journal of Modeling in Management, 3*(2), 125–139.

Chacko, G., & Viceira, L. M. (2003). Spectral GMM estimation of continuous-time processes. *Journal of Econometrics, 116*(1), 259–292.

Chacko, G., & Sanjiv Das. (2002). Pricing interest rate derivatives: a general approach. *Review of Financial Studies*, 195–241.

Chakravarty, S., Gulen, H., & Mayhew, S. (2004). Informed trading in stock and option markets. *Journal of Finance, 59*, 1235–1258.

Chan, C. Y., de Peretti, C., Qiao, Z., & Wong, W. K. (2012). Empirical test of the efficiency of UK covered warrants market: stochastic dominance and likelihood ratio test approach. *Journal of Empirical Finance, 19*, 162–174.

Chan, K. C., Karolyi, G. A., & Stulz, R. M. (1992a). Global financial markets and the risk premium on U.S. equity. *Journal of Financial Economics, 32*, 137–167.

Chan, K. C., Karolyi, G. A., Longstaff, F. A., & Sanders, A. B. (1992b). An Empirical Comparison of Alternative Models of the Short-Term Interest Rate. *Journal of Finance, 47*, 1209–1227.

Chan, K., Chung, P., & Johnson, H. (1993). Why option prices lag stock prices: A trading-based explanation. *Journal of Finance, 48*, 1957–1967.

Chan, K., Ikenberry, D., & Lee, I. (2004). Economic sources of gain in share repurchases. *Journal of Financial and Quantitative Analysis, 39*, 461–479.

Chan, K., Ikenberry, D., Lee, I., & Wang, Y. (2010). Share repurchases as a potential tool to mislead investors. *Journal of Corporate Finance, 16*, 137–158.

Chan, K., Peter Chung, Y., & Fong, W.-M. (2002). The informational role of stock and option volume. *Review of Financial Studies, 15*, 1049–1075.

Chan, L. K., Karceski, J., & Lakonishok, J. (1999). On Portfolio Optimization: Forecasting Covariances and Choosing the Risk Model. *Review of Financial Studies, 12*, 937–974.

Chan, N. H., & Tran, L. T. (1989). On the first order autoregressive process with infinite variance. *Econometric Theory, 5*, 354–362.

Chan, W. H., & Maheu, J. M. (2002). Conditional jump dynamics in stock market returns. *Journal of Business Economics Statistics, 20*, 377–389.

Chan, Y. (2000). IT value: The great divide between qualitative and quantitative and individual and organizational measures. *Journal of Management Information Systems, 16*, 225–261.

Chan, Y. L. (2004). Use of capital budgeting techniques and an analytic approach to capital investment decisions in Canadian municipal governments. *Public Budgeting & Finance, 24*, 40–58.

Chan, Y. L., & Lynn, B. E. (1991). Performance evaluation and the analytic hierarchy process. *Journal of Management Accounting Research*, 57–87.

Chance, D. M. (1986). Empirical tests of the pricing of index call options. *Advances in Futures and Options Research, 1*, 141–166.

Chance, D. (1996). A Generalized Simple Formula to Compute the Implied Volatility. *The Financial Review, 31*, 859–867, (November).

Chance, D. M., & Lane, W. R. (1980). A Re-examination of Interest Rate Sensitivity in the Common Stock of Financial Institutions. *Journal of Financial Research, 3*, 49–55.

Chander, R. (2006). Informational Efficiency, Parameter Stationarity and Bench Mark Consistency of Investment Performance, The ICFAI Journal of Applied Finance, March.

Chander, S., & Singh, J. (2004). Performance of Mutual Funds in India: An Empirical evidence. *The ICFAI Journal of Applied Finance*, 45–63.

Chang, R., Huang, M., Lee, F., & Lu, M. (2007). The jump behavior of foreign exchange market: Analysis of Thai Baht. *Review of Pacific Basin Financial Markets and Policies, 10*(2), 265–288.

Chang, C., & Lin, E. (2012). On the Determinants of Basis Spread for Taiwan Index Futures and the Role of Speculators. National Chiao Tung University working paper.

Chang, C. F. (1999). Determinants of capital structure and management compensation: The partial least squares approach. Ph.D. Dissertation, Rutgers University.

Chang, C. T. (2005). A modified goal programming approach for the mean-absolute deviation portfolio optimization model. *Applied Mathematics and Computation, 171*, 567–572.

Chang, C., Lee, A. C., & Lee, C. F. (2009). Determinants of capital structure choice: A structural equation modeling approach. *Quarterly Review of Economic and Finance, 49*(2), 197–213.

Chang, C., Fuh, C. D., & Hsu, Y. H. (2008). ESO compensation: the role of default risk, employee sentiment, and insider information. *Journal of Corporate Finance, 14*, 630–641.

Chang, C., Chen, L. J., & Fuh, C. D. (2011a). The pricing of risk and sentiment: a study of executive stock options. Financial Management, Conditional accepted.

Chang, T. H., Su, H. M., & Chiu, C. L. (2011b). Value-at-risk estimation with the optimal dynamic biofuel portfolio. *Energy Economics, 33*, 264–272.

Chang, E. C., & Lewellen, W. G. (1984). Market timing and mutual fund investment performance. *Journal of Business, 57*, 57–72.

Chang, E. C., Cheng, J. W., & Khorana, A. (2000). An examination of herd behavior in equity markets: An international perspective. *Journal of Banking and Finance, 24*, 1651–1679.

Chang, E., Eun, C., & Kolodny, R. (1995). International Diversification Through Closed-end Country Funds. *Journal of Banking and Finance, 19*, 1237–1263.

Chang, F., Dean, J., Ghemawat, S., Hsieh, W. C., Wallach, D. A., Burrows, M., Chandra, T., Fikes, A., & Gruber, R. E. (2006). BigTable: a distributed storage system for structured data. In: OSDI '06: Proceedings of the 7th USENIX Symposium on Operating Systems Design and Implementation, 15–15.

Chang, H. S., & Lee, C. F. (1977). Using pooled time-series and cross section data to test the firm and time effects in financial analysis. *Journal of Financial and Quantitative Analysis, 12*, 457–471.

Chao, J. C., & Swanson, N. R. (2005). Consistent estimation with a large number of weak instruments. *Econometrica, 73*, 1673–1692.

Chari, V. V., & Kehoe, P. (2004). Financial Crises as Herds: overturning the critiques. *Journal of Economic Theory, 119*(1), 128–150.

Charnes, A., & Cooper, W. W. (1961). *Management models and industrial applications of linear programming* (Vol. I & II). New York: Wiley & Sons.

Chatfield, R. E., & Moyer, R. C. (1986). Putting away bond risk: An empirical examination of the value of the put option on bonds. *Financial Management, 15*(2), 26–33.

Chava, S., & Tookes, H. (2005). Where (and how) does news impact trading? Unpublished working paper, Texas A&M University and Yale School of Management.

Chen, C. Y. (2002). Additional evidence on the association between director stock ownership and incentive compensation. *Review of Quantitative Finance and Accounting, 19*, 21–44.

Chen, K. H., & Shimerda, T. A. (1981). An empirical analysis of useful finance ratios. *Financial Management, 10*(1), 51–60.

Chen, R., & Scott, L. (1993). Maximum likelihood estimation of a multi-factor equilibrium model of the term structure of interest rates. *Journal of Fixed Income, 3*(3), 14–32.

Chen, R., & Yeh, S. K. (2002). Analytical bounds for American option prices. *Journal of Financial and Quantitative Analysis, 37*(01), 117–135.

Chen, S., & Matsumoto, D. (2006). Favorable versus unfavorable recommendations: The Impact on analyst access to management-provided information. *Journal of Accounting Research, 44*(5), 657–689.

Chen, B., & Hong, Y. (2005). Diagnostics of Multivariate Continuous-Time Models with Application to Affine Term Structure Models. working paper, Cornell University.

Chen, C. (1981). The EM approach to the multiple indicators and multiple causes model via the estimation of the latent variable. *Journal of the American Statistical Association, 76*, 702–708.

Chen, C., & Wu, C. (2009). Small trades and volatility increases after stock splits. *International Review of Economics and Finance, 18*, 592–610.

Chen, C., Chen, S., & Su, X. (2001). Is Accounting Information Value-Relevant in the Emerging Chinese Stock Market? *Journal of International Accounting, Auditing and Taxation, 10*, 1–22.

Chen, C. R., & Stockum, S. (1986). Selectivity, market timing, and random beta behavior of mutual funds: A generalized model. *Journal of Financial Research, 9*(1), 87–96.

Chen, C. W. S., Gerlach, R., & Lin, E. M. H. (2008). Volatility forecasting using threshold heteroskedastic models of the intra-day range. *Computational Statistics and Data Analysis, 52*, 2990–3010.

Chen, C. W. S., Gerlach, R., Hwang, B. B. K., & McAleer, M. (2012a). Forecasting value-at-risk using nonlinear regression quantiles and the intra-day range. *International Journal of Forecasting, 28*, 557–574.

Chen, H. Y., Lee, C. F., & Shih, W. K. (2012b). Alternative Equity Valuation Models, working paper.

Chen, F., & Sutcliffe, C. (2007). Better cross hedge with composite hedging? hedging equity portfolios using financial and commodity futures, ICMA discussion papers in finance DP2007-04. University of Reading.

Chen, G., Thulasiraman, P., & Thulasiram, R. K. (2006). Distributed Quasi-Monte Carlo Algorithm for Option Pricing on HNOWs Using mpC. In: Proceedings of the 39th Annual Simulation Symposium (ANSS'06), Huntsville, AL.

Chen, G., Firth, M., & Ning Gao, D. (2011a). The Information Content of Earnings Components: Evidence from the Chinese Stock Market. *European Accounting Review, 20*, 669–92.

Chen, H. Y. (2011). Momentum strategies, dividend policy, and asset pricing test. Ph.D. Dissertation, Rutgers University.

Chen, L. (1996). *Stochastic Mean and Stochastic Volatility - A Three-Factor Model of Term Structure of Interest Rates and its Application to the Pricing of Interest Rate Derivatives*. Oxford, UK: Blackwell Publishers.

Chen, L. D., Lesmond, A., & Wei, J. (2007a). Corporate yield spreads and bond liquidity. *Journal of Finance, 62*, 119–149.

Chen, L. J., & Fuh, C. D. (2011). A study of employee stock options and the exercise decision. *International Research Journal of Finance and Economics, 67*, 142–159.

Chen, L., Yen, M., Wu, H., Liao, C., Liou, D., Kuo, H., & Chen, T. H. (2005). Predictive survival model with time-dependent prognostic factors: Development of computer-aided SAS Macro program. *Journal of Evaluation in Clinical Practice, 11*(2), 181–193.

Chen, N. F., Roll, R. R., & Ross, S. A. (1986). Economic forces and the stock market. *Journal of Business, 59*, 383–403.

Chen, N. F., Roll, R., & Ross, S. A. (1983). Economic forces and the stock market: Testing the APT and alternative asset pricing theories. Working Paper, University of Chicago.

Chen, N., & Roll, R. (1986). Economic Forces and the Stock Market. *Journal of Business, 59*, 383–404.

Chen, N.-F. (1983). Some Empirical Tests of the Theory of Arbitrage Pricing. *Journal of Finance, 38*, 1393–1414.

Chen, Q., & Jiang, W. (2006). Analysts' weighting of private and public information. *Review of Financial Studies, 19*, 319–355.

Chen, Q., Gerlach, R., & Lu, Z. (2012c). Bayesian Value-at-Risk and expected shortfall forecasting via the asymmetric Laplace distribution. *Computational Statistics and Data Analysis, 56*, 3498–3516.

Chen, R., & Scott, L. (2001). Stochastic Volatility and Jumps in Interest Rates: An Empirical Analysis, Working paper, Rutgers University.

Chen, R. R., Lin, H. C., & Palmon, O. (2006). Explaining the volatility smile: reduced form versus structural option models. 2006 FMA Annual Meeting. Salt Lake City, UT.

Chen, S. N., & Lee, C. F. (1981). The sampling relationship between Sharpe's performance measure and its risk proxy: Sample size, investment horizon and market conditions. *Management Science, 27*(6), 607–618.

Chen, S. X. (2008). Nonparametric estimation of expected shortfall. *Journal of Financial Econometrics, 6*, 87–107.

Chen, S., Matsumoto, D., & Rajgopal, S. (2011b). Is silence golden? An empirical analysis of firms that stop giving quarterly earnings guidance. *Journal of Accounting and Economics, 51*(1–2), 134–150.

Chen, S. S., Lee, C. F., & Shrestha, K. (2001c). On a mean-generalized semivariance approach to determining the hedge ratio. *Journal of Futures Markets, 21*, 581–598.

Chen, S. S., Lee, C. F., & Shrestha, K. (2003). Futures hedge ratio: a review. *The Quarterly Review of Economics and Finance, 43*, 433–465.

Chen, S-S, & Wang, Y. (2012). Financial constraints and share repurchases. *Journal of Financial Economics*, (in press).

Chen, W. P., Chung, H., Lee, C. F., & Liao, W. L. (2007). Corporate governance and equity liquidity: Analysis of S&P transparency and disclosure rankings. *Corporate Governance: An International Review, 15*(4), 644–660.

Chen, X., Zou, C., & Cook, R. D. (2010). Coordinate-independent sparse sufficient dimension reduction and variable selection. *Annals of Statistics, 38*, 3696–3723.

Chen, X., Koenker, R., & Xiao, Z. (2009). Copula-based nonlinear quantile autoregression. *The Econometrics Journal, 12*, 50–67. January 2009.

Chen, X., Koenker, R., & Xiao, Z. (2009). Copula-based nonlinear quantile autoregression. *The Econometrics Journal, 12*, 50–67.

Chen, X., Fan, Y., & Patton, A. (2004). Simple tests for models of dependence between multiple financial time series, with applications to U.S. equity returns and exchange rates. Working Paper, Warwick Business School.

Chen, Z., & Lin, R. (2006). Mutual fund performance evaluation using data envelopment analysis with new risk measures. *OR Spectrum, 28*, 375–398.

Chen, Z., & Glasserman, P. (2008). Fast Pricing of Basket Default Swaps. *Operations Research, 56*(2), 286–303.

Chen, Z., & Knez, P. J. (1996). Portfolio performance measurement: Theory and applications. *The Review of Financial Studies, 9*, 511–555.

Cheng, Q., & Warfield, T. (2005). Equity incentives and earnings management. *The Accounting Review, 80*(2), 441–476.

Cheng, C., Liu, C., & Schaefer, T. (1996). Earnings permanence and the incremental information content of cash flows from operations. *Journal of Accounting Research, 34* (Spring), 173–181.

Cheng, D. C., & Lee, C. F. (1986). Power of alternative specification errors tests in identifying misspecified market models. *The Quarterly Review of Economics and Business, 16*(3), 6–24.

Cheng, J. K. (1993). Managerial flexibility in capital investment decision: insights from the real-options. *Journal of Accounting Literature, 12*, 29–66.

Cheng, K. F., Chu, C. K., & Hwang, R. C. (2010). Predicting bankruptcy using the discrete-time semiparametric hazard model. *Quantitative Finance, 10*, 1055–1066.

Chernoff, H., & Zacks, S. (1964). Estimating the current mean of a normal distribution which is subjected to change in time. *The Annals of Mathematical Statistics, 35*, 999–1018.

Chernov, M., & Ghysels, E. (2000). A Study Towards a Unified Approach to the Joint Estimation of Objective and Risk Neutral Measures for the Purpose of Options Valuation. *Journal of Financial Economics, 56*, 407–458.

Chernozhukov, V., & Umantsev, L. (2001). Conditional value-at-risk: Aspects of modeling and estimation. *Empirical Economics, 26*, 271–292.

Cherubini, U., Luciano, E., & Vecchiato, W. (2004a). *Copula Methods in Finance*. Hoboken, NJ: Wiley.

Cherubini, U., Luciano, E., & Vecchiato, W. (2004b). *Copula Methods in Finance. Wiley Series in Financial Engineering*. Chichester, UK: Wiley.

Cheung, Y. (2007). An empirical model of daily highs and lows. *International Journal of Finance and Economics, 12*, 1–20.

Cheung, Y., Cheung, Y., & Wan, A. (2009). A high–low model of daily stock price ranges. *Journal of Forecasting, 28*, 103–119.

Chia, Yuxuan. (1994). Examination of mandatory management earnings forecasts and the rate of trading turnover, Unpublished thesis, Accounting Department, Politics University

Chiang, A. P., Beck, J. S., Yen, H.-J., Tayeh, M. K., Scheetz, T. E., Swiderski, R., Nishimura, D., Braun, T. A., Kim, K.-Y., Huang, J., Elbedour, K., Carmi, R., Slusarski, D. C., Casavant, T. L., Stone, E. M., & Shefield, V. C. (2006). Homozygosity Mapping with SNP Arrays Identifies a Novel Gene for Bardet-Biedl Syndrome (BBS10). *Proceeding of the National Academy of Sciences, 103*, 6287–6292.

Chiang, A. C. (1984). *Fundamental Methods of Mathematical Economics* (3rd ed.). New York, NY: McGraw-Hill.

Chiang, Ethan. (2012). Modern portfolio management with conditioning information, working paper, University of North Carolina, Charlotte.

Chiang, M., & Wang, L. (2011). Volatility contagion: A range-based volatility approach. *Journal of Econometrics, 165*, 175–189.

Chiang, M. H., Yueh, M. L., & Hsieh, M. H. (2007a). An Efficient Algorithm for Basket Default Swap Valuation. *Journal of Derivatives, 15*(2), 1–22.

Chiang, T. C., Yu, H. C., & Wu, M. C. (2009). Statistical properties, dynamic conditional correlation, scaling analysis of high-frequency intraday stock returns: Evidence from Dow-Jones and Nasdaq Indices. *Physica A, 388*(8), 1555–1570.

Chiang, T. C., Jeon, B. N., & Li, H. M. (2007b). Dynamic correlation analysis of financial contagion: Evidence from Asian markets. *Journal of International Money and Finance, 26*(7), 1206–1228.

Chiang, T. C., Tan, L., & Li, H. M. (2007c). Empirical analysis of dynamic correlations of stock returns: Evidence from Chinese A-share and B-share markets. *Quantitative Finance, 7*(6), 651–667.

Chiappori, P., & Salanie, B. (2000). Testing for asymmetric information in insurance markets. *Journal of Political Economy, 108*, 56–78.

Chib, S., Nardari, F., & Shephard, N. (2002). Markov Chain Monte Carlo methods for stochastic volatility models. *Journal of Econometrics, 108*, 281–316.

Chien, I. S., Shi, Y., & Yu, P. L. (1989). MC2 program: a Pascal program run on PC or Vax (revised version), School of Business, University of Kansas.

Childers, D. (1978). *Modern Spectrum Analysis"*. New York: IEEE Press.

Chionis, D., & MacDonald, R. (1997). Some tests of market microstructure hypotheses in the foreign exchange market. *Journal of Multinational Financial Management, 7*, 203–229.

Chiras, D., & Manaster, S. (1978). The Informational Content of Option Prices and a Test of Market Efficiency. *Journal of Financial Economics, 6*, 213–34.

Chirinko, R. (1993). Business Fixed Investment Spending: Modeling Strategies, Empirical Results, and Policy Implications. *Journal of Economic Literature, 31*, 1875–1911.

Chmielewski, M. A. (1981). Elliptically symmetric distributions: a review and bibliography. *International Statistical Review, 49*, 67–74.

Cho, J. Y., & Jung, K. (1991). Earnings Response Coefficients: A Synthesis of Theory and Empirical Evidence. *Journal of Accounting Literature, 10*, 85–116.

Cho, M. (1998). Ownership structure, investment and the corporate value: an empirical analysis. *Journal of Financial Economics, 47*, 103–121.

Choe, H., Kho, B. C., & Stulz, R. (1999). Do foreign investors destabilize stock markets? The Korean experience in 1997. *Journal of Financial Economics, 54*, 227–264.

Choi, J. J., Elyasiani, E., & Kopecky, K. (1992). The Sensitivity of Bank Stock Returns to Market, Interest, and Exchange Rate Risks". *Journal of Banking and Finance, 16*, 983–1004.

Choi, T. S., & Levary, R. R. (1989). Multi-national capital budgeting using chance-constrained goal programming. *International Journal of Systems Science, 20*, 395–414.

Chollete, L., Heinen, A., & Valdesogo, A. (2009). Modeling international financial returns with a multivariate regime-switching copula. *Journal of Financial Econometrics, 7*, 437–480.

Chollete, L., de la Penã, V., & Lu, C. (2011). "International diversification: A copula approach''. *Journal of Banking and Finance, 35*, 403–417.

Chong, Y. Y., & Hendry, D. F. (1986). Econometric evaluation of linear macroeconomic models. *Review of Economic Studies, 53*, 671–690.

Chou, H., Zaabar, R., & Wang, D. (2013). Measuring and testing the long-term impact of terrorist attacks on the US futures market. *Applied Economics, 45*, 225–238.

Chou, H., & Wang, D. (2007). Forecasting volatility of UK stock market: A test of conditional autoregressive range (CARR) model. *International Research Journal of Finance and Economics, 10*, 7–13.

Chou, R. (2005). Forecasting financial volatilities with extreme values: the conditional autoregressive range (CARR) model. *Journal of Money Credit and Banking, 37*, 561–82.

Chou, R. (2006). Modeling the Asymmetry of Stock Movements Using Price Ranges. *Advances in Econometrics, 20A*, 231–257.

Chou, R. Y., & Liu, N. (2010). The economic value of volatility timing using a range-based volatility model. *Journal of Economic Dynamics and Control, 34*, 2288–2301.

Chou, R. Y., & Liu, N. (2011). *Estimating time-varying hedge ratios with a range-based multivariate volatility model*. Working paper, Academia Sinica.

Chou, R. Y., & Cai, Y. (2009). Range-based multivariate volatility model with double smooth transition in conditional correlation. *Global Finance Journal, 20*, 137–152.

Chou, R. Y. (1988). Volatility persistence and stock valuations: Some empirical evidence using GARCH. *Journal of Applied Econometrics, 3*, 279–294.

Chou, R. Y., Wu, C., & Liu, N. (2009). Forecasting time-varying covariance with a range-based dynamic conditional correlation model. *Review of Quantitative Finance and Accounting, 33*, 327–345.

Chou, R., Engle, R. F., & Kane, A. (1992). Measuring risk aversion from excess returns on a stock index. *Journal of Econometrics, 52*, 201–224.

Chow, G. C. (1960). Tests of equality between sets of coefficients in two linear regressions. *Econometrica, 28*, 591–605.

Chow, G. (1983). *Econometrics*. New York: McGraw Hill.

Chow, Y. S. (1965). Local convergence of martingales and the law of large numbers. *The Annals of Mathematical Statistics, 36*, 552–558.

Chowdry, B. (1986). *Pricing of futures with quality option*. Working paper, University of Chicago.

Christensen, J. H. E., Diebold, F. X., & Rudebusch, G. D. (2009a). An arbitrage-free generalized Nelson-Siegel term structure model. *Econometrics Journal, 12*, 33–64.

Christensen, J. H. E., Diebold, F. X., & Rudebusch, G. D. (2011). The affine arbitrage-free class of Nelson-Siegel term structure models. *Journal of Econometrics, 164*, 4–20.

Christensen, K., & Podolskij, M. (2007). Realized range-based estimation of integrated variance. *Journal of Econometrics, 141*, 323–349.

Christensen, K., Podolskij, M., & Vetter, M. (2009b). Bias-correcting the realized range-based variance in the presence of market microstructure noise. *Finance and Stochastics, 13*, 239–268.

Christiansen, C., & Ranaldo, A. (2011). The time-varying systematic risk of carry trade strategies. *Journal of Financial and Quantitative Analysis, 46*(4), 1107–1125.

Christie, A. (1987). On cross-sectional analysis in accounting research. *Journal of Accounting and Economics, 9*, 231–58.

Christie, A. A. (1981). The Stochastic Behavior of Common Stocks Variances: Value, Leverage and Interest Rate Effects. *Journal of Financial Economics, 10*, 407–432.

Christie, W. G., & Huang, R. D. (1995). *"Following the Pied Piper: Do individual Returns Herd around the Market?" Financial Analyst Journal July-August, 31–37*.

Christoffersen, P. F. (2002). *Elements of financial risk management*. San Diego: Academic Press.

Christoffersen, P. F., & Diebold, F. X. (2006). Financial asset returns, direction-of-change, forecasting, and volatility dynamics. *Management Science, 52*, 1273–1287.

Christoffersen, P., Dorion, C., Jacobs, K., & Wang, Y. (2010a). Volatility components: affine restrictions and non-normal innovations. *Journal of Business and Economic Statistics, 28*, 483–502.

Christoffersen, P., Elkamhi, R., Feunou, B., & Jacobs, K. (2010b). Option valuation with conditional heteroskedasticity and nonnormality". *Review of Financial Studies, 23*, 2139–2189.

Christoffersen, P., Jacobs, K., Ornthanalai, C., & Wang, Y. (2008). Option valuation with long-run and short-run volatility components. *Journal of Financial Economics, 90*, 272–297.

Chu, C. C. (1984). *Alternative methods for determining the expected market risk premium: Theory and evidence*. Ph.D. Dissertation, University of Illinois at Urbana-Champaign.

Chu, C. J., Stinchcombe, M., & White, H. (1996). Monitoring structural change. *Econometrica, 64*, 1045–1065.

Chuang, H., Lee, C. F., & Zhong, Z. K. (2011). *Option bounds: a review and comparison*. Working paper.

Chung, K. H., Elder, J., & Kim, J. (2010a). Corporate governance and liquidity. *Journal of Financial and Quantitative Analysis, 45*(02), 265–291.

Chung, S. L., Hung, M. W., & Wang, J. Y. (2010b). Tight bounds on American option prices". *Journal of Banking and Finance, 34*, 77–89.

Church, K. W., & Gale, W. A. (1995). Poisson Mixtures. *Natural Language Engineering, 1*(2), 163–190.

Churchill, G. A., & Iacobucci, D. (2004). Marketing research: Methodological foundations (Ninth Ed.), South-Western College.

Churchill, G. A., Jr., & Research, M. (1983). *Methodological Foundations* (3rd ed.). New York: CBS College Publishing.

Cipollini, F., Engle, R. F., & Gallo, G. M. (2009). *A model for multivariate non-negative valued processes in financial econometrics*. Working paper.

Citron, J. T., & Neckelburg, G. (1987). Country Risk and Political Instability. *Journal of Development Economics, 25*, 385–395.

Cizek, P., Härdle, W., & Weron, R. (2011) *Statistical tools for finance and insurance*. Springer.

Claessens, S., Djankov, S., Fan, J. P. H., & Lang, L. H. P. (2002). Disentangling the incentive and entrenchment effects of large shareholdings. *Journal of Finance 2004*.

Claessens, S., Djankov, S., & Lang, L. H. P. (2000). The separation of ownership and control in East Asian corporations. *Journal of Financial Economics, 58*, 81–112.

Claessens, S., Dasgupta, S., & Glen, J. (1995). Return Behavior in Emerging Stock Markets. *World Bank Economic Review, 9*(1), 131–151.

Clarida, R., Davis, J., & Pedersen, N. (2009). Currency carry trade regimes: Beyond the Fama regression. *Journal of International Money and Finance, 28*, 1375–1389.

Clark, P. K. (1973). A subordinated stochastic process model with finite variance for speculative prices. *Econometrica, 41*, 135–155.

Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting, 5*, 559–583.

Clement, M. (1999). Analyst forecast accuracy: Do ability, resources, and portfolio complexity matter?". *Journal of Accounting and Economics, 27*, 285–303.

Clement, M., & Tse, S. (2005). Financial analyst characteristics and herding behavior in forecasting. *Journal of Finance*, 60, 307–341

Clements, B, J., Kontolemis, G. Z., & Levy, J. V. (2001). Monetary policy under EMU: differences in the transmission mechanisms? IMF Working paper 01/102.

Clements, M. P., & Smith, J. (2000). Evaluating the Forecast Densities of Linear and Nonlinear Models: Applications to Output Growth and Unemployment. *Journal of Forecasting, 19*, 255–276.

Clements, M. P., & Smith, J. (2002). Evaluating Multivariate Forecast Densities: A Comparison of Two Approaches. *International Journal of Forecasting, 18*, 397–407.

Clemons, E., & Weber, B. (1997). Information Technology and Screen-Based Securities Trading: Pricing the Stock and Pricing the Trade. *Management Science, 43*(12), 1693–1708.

Clewlow, L., & Strickland, C. (2000). Valuing Energy Options In A One-Factor Model Fitted To Forward Prices. Working paper, University of Sydney.

Cliff, M. T., & Denis, D. J. (2004). Do initial public offering firms purchase analyst coverage with underpricing?. *Journal of Finance, 59*, 2871–2901.

Coase, R. H. (1960). The problem of social cost. *Journal of Law and Economics, 3*, 1–44.

Coates, J. C., & Hubbard, G. R. (2007). Competition in the mutual fund industry: Evidence and implications for policy, Discussion paper No.592, Aug, Source: http://ssrn.com/abstract=1005426

Cochrane, J. H. (1996). A Cross-sectional Test of an Investment-Based Asset Pricing Model. *Journal of Political Economy, 104*(3), 572–621.

Cochrane, J. (1992). Explaining the variance of price-dividend ratios. *Review of Financial Studies, 5*, 243–280.

Cohen, H. (1995). Isolating the wild card option. *Mathematical Finance, 5*(2), 155–165, April

Cohen, K., Conroy, R., & Maier, S. (1985). Order Flow and the Quality of the Market, In Y. Amihud, T. Ho, & R. Schwartz (Eds.), *Market Making and the Changing Structure of the Securities Industry*, Lexington Books, pp. 93–109

Cohen, R. H., & Bonham, C. S. (2006). Specifying the forecast generating process for exchange rate survey forecasts. Working paper, Department of Economics, University of Hawaii Manoa

Cohn, R., & Lessard, D. (1981). Are markets efficient? Tests of alternative hypothesis. On the effect of inflation on stock prices: International evidence. *Journal of Finance, 36*, 277–289.

Colak, G., & Whited, T. (2007). Spin-offs, Divestitures, and Conglomerate Investment. *Review of Financial Studies, 20*, 557–595.

Coles, J. L., Lemmon, M. L., & Meschke, F. J. (2012). Structural models and endogeneity in corporate finance: the link between managerial ownership and corporate performance. *Journal of Financial Economics, 103*, 149–68.

Collin-Dufresne, P., & Goldstein, R. S. (2002). Do Bonds Span the Fixed Income Markets? Theory and Evidence for Unspanned Stochastic Volatility. *Journal of Finance, 57*, 1685–1729.

Collin-Dufresne, P., & Goldstein, R. S. (2003). Stochastic Correlation and the Relative Pricing of Caps and Swaptions in a Generalized Affine Framework. Working paper, Carnegie Mellon University.

Collin-Dufresne, P., Goldstein, R., & Hugonnier, J. (2004). A General Formula for Valuing Defaultable Securities. *Econometrica, 72*(5), 1377–1407.

Collins, A. A., Scorcu, E., & Zanola, R. (2007). "Sample Selection Bias and Time Instability of Hedonic Art Price Indexes, Quaderni , Working Papers DSE N° 610

Collins, D., & Kothari, S. (1989). An Analysis of International and Cross-Sectional Determinants of Earnings Response Coefficients. *Journal of Accounting and Economics, 11*, 143–81.

Collins, D., Kothari, S., & Rayburn, J. (1987). Firm Size and the Information Content of Prices with Respect to Earnings. *Journal of Accounting and Economics, 9*, 111–38.

Comon, P. (1994). Independent component analysis, a new concept?. *Signal Processing, 36*, 287–314.

Compustat (2008). Standard & Poor's Compustat website: http://www.compustat.com/

Comte, F., & Renault, E. (1998). Long memory in continuous-time stochastic volatility models. *Mathematical Finance, 8*, 291–323.

Conger, J., Finegold, D., & Lawler, III E. (1998). Appraising boardroom performance.

Connolly, R. A., Stivers, C. T., & Sun, L. (2005). "Stock market uncertainty and the stock-bond return relation''. *Journal of Financial and Quantitative Analysis, 40*, 161–194.

Connor, G., & Korajczyk, R. (1986). Performance measure with the arbitrage pricing theory: A new framework for analysis. *Journal of Financial Economics, 15*(3), 373–394.

Conover, W. J. (1999). *Practical Nonparametric Statistics* (6th ed.). John Wiley & Sons, New York

Conrad, J., Johnson, K. M., & Wahal, S. (2003). Institutional trading and alternative trading systems. *Journal of Financial Economics, 70*, 99–134.

Conrad, J. S., Dittmar, R. F., & Ghysels, E. (2009). Ex Ante Skewness and Expected Stock Returns. Available at SSRN: http://ssrn.com/abstract=1522218 or http://dx.doi.org/10.2139/ssrn.1522218. (December 11).

Conrad, J. S., Hameed, A., & Niden, C. (1994). Volume and autocovariances in short-horizon individual security returns. *Journal of Finance, 49*, 1305–1329.

Conrad, J., Cornell, B., Landsman, W., & Rountree, B. (2006). How do analyst recommendations respond to major news?". *Journal of Financial and Quantitative Analysis, 41*, 25–49.

Conroy, R., & Miles, M. (1989). Commercial forestland in the pension portfolio: The biological beta. *Financial Analysts Journal, 45*, 46.

Conroy, R., & Harris, R. (1987). Consensus Forecasts of Corporate Earnings: Analysts Forecasts and Time Series Methods. *Management Science, 33*, 725–38.

Constantinides, G., & Zariphopoulou, T. (2001). Bounds on derivative prices in an intertemporal setting with proportional transaction costs and multiple securities". *Mathematical Finance, 11*, 331–346.

Constantinides, G. M., Czerwonko, M., Jackwerth, J. C., & Perrakis, S. (2011). Are options on index futures profitable for risk-averse investors? empirical evidence". *Journal of Finance, 66*, 1407–1437.

Constantinides, G. M., Donaldson, J. M., & Mehra, R. (2002). Junior Can't Borrow: A New Perspective on the Equity Premium Puzzle. *Quarterly Journal of Economics, 117*, 269–296.

Constantinides, G. M., Jackwerth, J. C., & Perrakis, S. (2009). Mispricing of S&P 500 Index Options". *Review of Financial Studies, 22*, 1247–1277.

Cook, D. O., & Easterwood, J. C. (1994). Poison put bonds: An analysis of their economic role. *Journal of Finance, 49*, 1905–20.

Cook, R. D. (1977). Detection of Influential Observations in Linear Regression. *Technometrics, 19*(1), 15–18.

Cook, R. D. (1998). Regression graphics: ideas for studying regressions through graphics. Wiley Online Library.

Cook, R. D., & Weisberg, S. (1991). Sliced inverse regression for dimension reduction: Comment. *Journal of the American Statistical Association, 86*, 328–332.
Copeland, T. E. (1976). A model of asset trading under the assumption of sequential information arrival. *Journal of Finance, 34*, 1149–1168.
Copeland, T. E., Weston, J. F., & Shastri, K. (2005). *Financial theory and corporate policy* (4th ed.). Pearson Addison Wesley, Boston.
Copeland, T. E., & Galai, D. (1983). Information Effects on the Bid-Ask Spread. *Journal of Finance, 38*, 1457–1469.
Copeland, T. E., & Weston, J. F. (1979). *Financial Theory and Corporate Policy*, Addison-Wesley Publishing Company, Inc.
Core, J., Holthausen, R., & Larcker, D. (1999). Corporate governance, chief executive officer compensation, and firm performance. *Journal of Financial Economics, 51*(3), 371–406.
Core, J. E. (2000). The directors' and officers' insurance premium: An outside assessment of the quality of corporate governance. *Journal of Law, Economics, and Organization, 16*(2), 449–477.
Core, J. E., & Guay, W. (2001). Stock option plans for non-executive employees. *Journal of Financial Economics, 61*, 253–287.
Cornell, B., & Landsman, W. (1989). Security Price Response to Quarterly Earnings Announcements and Analysis Forecast Revisions. *The Accounting Review, 64*, 680–92.
Corner, I. L., Decicen, R. F., & Spahr, R. W. (1993). Multiple-objective linear programming in capital budgeting. *Advances in Mathematical Programming and Financial Planning, 3*, 241–264.
Corradi, V., & Swanson, N. R. (2006). Predictive Density and Conditional Confidence Interval Accuracy Tests. *Journal of Econometrics, 135*, 187–228.
Corradi, V., & Swanson, N. R. (2007a). Nonparametric Bootstrap Procedures for Predictive Inference Based on Recursive Estimation Schemes. *International Economic Review, 48*, 67–109.
Corradi, V., & Swanson, N. R. (2011). Predictive Density Construction and Accuracy Testing with Multiple Possibly Misspecified Diffusion Models. *Journal of Econometrics, 161*, 304–324.
Corradi, V., & Swanson, N. R. (2005). A Bootstrap Specification Test for Diffusion Processes. *Journal of Econometrics, 124*, 117–148.
Corradi, V., & Swanson, N. R. (2007b). Nonparametric Bootstrap Procedures for Predictive Inference Based on Recursive Estimation Schemes. *International Economic Review February, 2007*(48), 67–109.
Corrado, C., & Truong, C. (2007). Forecasting stock index volatility: Comparing implied volatility and the intraday high-low price range. *Journal of Financial Research, 30*, 201–215.
Corrado, C. J., & Miller, T. W. (1996). A Note on a Simple, Accurate Formula to Compute Implied Standard Deviations. *Journal of Banking and Finance, 20*, 595–603.
Corrado, C. J., & Miller, T. W. (2004). Tweaking Implied Volatility. Working paper, Auckland, New Zealand: Massey University-Albany
Cortazar, G., Schwartz, E. S., & Casassus, J. (2001). Optimal Exploration Investments under Price and Geological-Technical Uncertainty: A Real Options Model. *R&D Management, 31*, 181–190.
Costanigro, M., Mittelhammer, R., & McCluskey, J. (2009). Estimating class-specific parametric models under class uncertainty: local polynomial regression clustering in an hedonic analysis of wine markets. *Journal of Applied Econometrics, 24*(7), 1117–1135, November/December
Costanigro, M., McCluskey, J. J., & Mittelhammer, R. C. (2007). Segmenting the Wine Market Based on Price: Hedonic Regression when Different Prices mean Different Products. *Journal of Agricultural Economics, 58*(3), 454–466.
Court, A. T. (1939). Hedonic Price Indexes With Automotive Examples. The Dynamics of Automobile Demand, New York, General Motors.
Courtadon, G. (1982). The Pricing of Options on Default-Free Bonds. *Journal of Financial and Quantitative Analysis, 17*, 75–100.

Cowan, A. R., & Sergeant, A. M. A. (2001). Interacting biases, non-normal return distributions and the performance of tests for long-horizon event studies. *Journal of Banking and Finance, 25*, 741–765.

Cox, J., Ingersoll, J., & Ross, S. (1981). The relation between forward prices and futures prices. *Journal of Financial Economics, 9(4), pp. 321–346*(December).

Cox, J., Ingersoll, J., & Ross, S. (1985). A theory of the term structure of interest rates. *Econometrica, 53*(2), 385–407, March.

Cox, D. (1958). *The planning of experiments*. New York: John Wiley & Sons.

Cox, D. (1972). Regression models and life tables. *Journal of Royal Statistical Society Series B (Methodological), 34*, 187–220.

Cox, J., & Rubinstein, M. (1985). *Options markets*. NJ: Prentice-Hall.

Cox, J., & Ross, S. (1976). The valuation of options for alternative stochastic process". *Journal of Finance Economics, 3*, 145–166.

Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985). A Theory of the Term Structure of Interest Rates. *Econometrica, 53*, 385–407.

Cox, J. C., Ross, S. A., & Rubinstein, M. (1979). Option pricing: A simplified approach. *Journal of Financial Economics, 7*(3), 229–263.

Cox, J. C., Ingersoll, J. E., & Ross, S. A. (1985). An Intertemporal General Equilibrium Model of Asset Prices. *Econometrica, 53*, 363–84.

Crabbe, L. (1991). Event risk: An analysis of losses to bondholders and 'super poison put' bond covenants. *Journal of Finance, 46*, 689–706.

Crabbe, L., & Nikoulis, P. (1997). The putable bond market: Structure, historical experience, and strategies. *The Journal of Fixed Income, 7*, 47–60.

Cragg, J. (1997). Using Higher Moments to Estimate the Simple Errors-In-Variables Model. *RAND Journal of Economics, 28*, 71–91.

Cragg, J. G., & Donald, S. G. (1993). Testing identifiability and specification in instrumental variable models. *Econometric Theory, 9*, 222–240.

Crama, Y., Hammer, P. L., & Ibaraki, T. (1988). Cause-Effect Relationships and Partially Defined Boolean Functions. *Annals of Operations Research, 16*, 299–326.

Craven, P., & Wahba, G. (1979). Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation. *Numer. Math, 31*, 377–403.

CrediteSuisse Financial Products. (1997). CreditRisk+: A Credit Risk ManagementFramework. New York, NY.

CreditMetrics. (1997). The benchmark for understanding credit risk. Technical Document, NewYork, NY: JP Morgan Inc.

Crouhy, M., Galai, D., & Mark, R. (2000). A comparative analysis of current credit risk models. *Journal of Banking and Finance, 24*, 59–117.

Crowley, P. (2005). An Intuitive Guide to Wavelets for Economists, Working Paper, Bank of Finland.

Crowley, P. (2007). A guide to wavelets for economists. *Journal of Economics Surveys, 21*, 207–264.

Cumby, R. E., & Glen, J. D. (1990). Evaluating the Performance of International Mutual Funds. *Journal of Finance, 45*(2), 497–521.

Cummins, J. D., & Nye, D. J. (1972). The Cost of Capital of Insurance Companies: Comment. *Journal of Risk and Insurance, 39*(3), 487–491.

Cummins, J. D., & Harrington, S. E. (1987). *Fair Rate of Return in Property-Liability Insurance*. Norwell, MA: Kluwer Academic Publishers.

Cummins, J. D. (1990a). Asset Pricing Models and Insurance Ratemaking. *ASTIN Bulletin, 20*(2), 125–166.

Cummins, J. D. (1990b). Multi-Period Discounted Cash Flow Rate-making Models in Property-Liability Insurance. *Journal of Risk and Insurance, 57*(1), 79–109.

Cummins, J. D. (1992). Financial Pricing of Property-Liability Insurance. In G. Dionne (Ed.), *Contributions to Insurance Economics* (pp. 141–168). Boston: Kluwer Academic Publishers.

Curci, G., & Corsi, F. (2012). Discrete sine transform for multi-scale realized volatility measures. *Quantitative Finance, 12*, 263–279.

Curry, T., & Shibut, L. (2000). Cost of the S&L Crisis. *FDIC Banking Review, 2*(2).

Curtis, S. L. (2010). Intercompany licensing rates: Implications from principal-agent theory. *International Tax Journal, 36*, 25–46.

Cutler, D. M., Poterba, J. M., & Summers, L. H. (1991). Speculative dynamics". *Review of Economic Studies, 58*, 529–546.

Cutter, B., Fernandez, L., Sharma, R., & Scott, T. (2011). Dynamic Analysis Of Open Space Value Using A Repeat Sales/Hedonic Approach, Working Paper, University of Oregon.

Cvitanic, J., Polimenis, V., & Zapatero, F. (2005). Optimal Portfolio Allocation with Higher Moments. National Science Foundation Working Paper, NSF-DMS 04-03575.

Czyzyk, J., Mesnier, M. P., & Moré, J. J. (1998). The NEOS Server. *IEEE Computer Science Engineering, 5*(3), 68–75.

D'Arcy, S. P., & Garven, J. A. (1990). Property-Liability Insurance Pricing Models: An Empirical Evaluation. *Journal of Risk and Insurance, 57*(3), 391–430.

D'Arcy, S. P., & Neil, A. (1988). *Doherty, Financial Theory of Insurance Pricing*. Philadelphia: S.S. Huebner Foundation.

Dacorogna, M. M. (2001). An introduction to high-frequency finance. Academic Press.

Dacorogna, M. M., Muller, U. A., Nagler, R. J., Olsen, R. B., & Pictet, O. V. (1993). A geographical model for the daily and weekly seasonal volatility in the foreign exchange market. *Journal of International Money and Finance, 12*, 413–438.

D'Agostino, R., & Stephens, M. (1986). *Goodness-of-Fit Techniques*. New York and Basel: Marcel Dekker, Inc.

Dahlquist, M., & de Jong, F. (2008). Pseudo market timing: a reappraisal. *Journal of Financial and Quantitative Analysis, 43*, 547–580.

Dai, Q., & Singleton, K. J. (2000). Specification Analysis of Affine Term Structure Models. *Journal of Finance, 55*, 1943–1978.

Dai, Q., & Singleton, K. (2003). Term Structure Dynamics in Theory and Reality. *Review of Financial Studies, 16*, 631–678.

Dai, Q., & Singleton, K. (2000). Specification analysis of Affine term structure models. *Journal of Finance, 50*(5), 1943–1978.

Dai, Q., & Singleton, K. (2002). Expecttatio puzzles, time-varying risk premia, and affine models of the term structure. *Journal of Financial Economics, 63*(3), 415–441.

Dai, T. S. (2009). Efficient Option Pricing on Stocks Paying Discrete or Path-Dependent Dividends with the Stair Tree. *Quantitative Finance, 9*, 827–838.

Dai, T. S., & Lyuu, Y. D. (2009). Accurate Approximation Formulas for Stock Options with Discrete Dividends. *Applied Economics Letters, 16*, 1657–1663.

Daigler, R. T. (1994). *Financial Futures and Options Markets Concepts and Strategies*. New York: Harper Collins.

Daily, C. M., & Dalton, D. R. (1997). Separate, but not independent: Board leadership structure in large corporations. *Corporate Governance: An International Review, 5*(3), 126–136.

Daines, R. (2001). Does Delaware law improve firm value? *Journal of Financial Economics, 62*, 525–558.

Dang, H. (2010). Rating History, Time and The Dynamic Estimation of Rating Migration Hazard. Ph.D. Thesis, University of Sydney. http://hdl.handle.net/2123/6397. Accessed 30 June 2011.

Daniel, K., & Titman, S. (1997). Evidence on the characteristics of cross sectional variation in stock returns. *Journal of Finance, 52*, 1–33.

Danielson, M. G., Heck, J. L., & Shaffer, D. R. (2008). Shareholder theory – How opponents and proponents both get it wrong. *Journal of Applied Finance, 18*(2), 62–66.

Danielsson, J. (1994). Stochastic volatility in asset prices: estimation with simulated maximum likelihood. *Journal of Econometrics,* 375–400.

Danielsson, J., de Vries, C. G., & Jorgensen, B. N. (1998a). The Value of Value at Risk: Statistical, Financial, and Regulatory Considerations. *FRBNY Economic Policy Review , 107–108*.

Danielsson, J., Hartmann, P., & de Vries, C. G. (1998b). The Cost of Conservatism: Extreme Returns, Value-at-Risk, and the Basle 'Multiplication Factor. *Risk, 11*(January), 101–113.

Dann, L. Y., & Mikkelson, W. H. (1984). Convertible debt Issuance, capital structure change and financing-related information: Some new evidence. *Journal of Financial Economics, 13*, 157–86.

Dantzig, G. B. (1963). *Linear programming and extensions*. New Jersey: Princeton University Press.

Dantzig, G. B., & Wolfe, P. (1960). Decomposition principles for linear programs. *Operations Research,* 101–111.

Darrat, A. F., Gilley, O. W., Li, B., & Wu, Y. (2011). Revisiting the risk/return relations in the Asian Pacific markets: New evidence from alternative models. *Journal of Business Research, 64*, 199–206.

Darvas, Z. (2009). Leveraged carry trade portfolios. *Journal of Banking and Finance, 33*, 944–957.

Das, S., Levine, C., & Sivaramakrishnan, K. (1998). Earnings predictability and bias in analysts' earnings forecast. *The Accounting Review, 73*(2), 277–294.

Das, S., Levine, C. B., & Sivaramakrishnan, K. (1998). Earnings predictability and bias in analysts' earnings forecasts. Financial Analysts Journal, November/December, 35–42.

Das, S. R., Freed, L., Geng, G., & Kapadia, N. (2006). Correlated default risk. *Journal of Fixed Income, 16*(2), 7–32.

Das, S. (2002). The Surprise Element: Jumps in Interest Rates. *Journal of Econometrics, 106*, 27–65.

Das, S., Levine, C. B., & Sivaramakrishnan, K. (1998). Earnings predictability and bias in analysts' earnings forecasts. *The Accounting Review, 73*, 277–294.

Das, S., Duffie, D., Kapadia, N., & Saita, L. (2007). Common failings: How corporate defaults are correlated. *Journal of Finance, 62*, 93–117.

Datar, V. T., Naik, N. Y., & Radcliffe, R. (1998). Liquidity and stock returns: An alternative test. *Journal of Financial Markets, 1*, 203–219.

David, A. (2001). Pricing the strategic value of putable securities in liquidity crises. *Journal of Financial Economics, 59*, 63–99.

Davidson, R., & Flachaire, E. (2008). The wild bootstrap, tamed at last. *Journal of Econometrics, 146*, 162–169.

Davidson, R., & MacKinnon, J. G. (1993). *Estimation and inference in Econometrics*. New York: Oxford University Press.

Davidson, Russell and James G. Mackinnon, "Several Tests for Model Specification in the Presence of Alternative Hypotheses." Econometrica, 49 (No. 3. May), 781–793, (1981).

Davies, A., & Lahiri, K. (1995). A new framework for testing rationality and measuring aggregate shocks using panel data. *Journal of Econometrics, 68*, 205–227.

Davies, R. B. (1977). Hypothesis testing when a nuisance parameter is present only under the alternative. *Biometrika, 64*, 247–254.

Davis, M., & Lo, V. (2001). Infectious defaults. *Quantitative Finance, 1*, 382–387.

Davis, P. (2010). "A firm-level test of the CAPM." Working paper

Davison, A. C., & Hinkley, D. V. (1997). *Bootstrap methods and their application*. Cambridge University Press.

Dawson, P. J., Tiffin, A. L., & White, B. (2000). Optimal hedging ratios for wheat and barley at the LIFFE: A GARCH approach. *Journal of Agricultural Economics, 51*, 47–161.

De Angelo, H., De Angelo, L., & Stultz, R. (2006). Dividend policy and the earned/contributed capital mix: A test of the life cycle theory. *Journal of Financial Economics, 81*, 227–254.

De Jong, A., De Roon, F., & Veld, C. (1997). Out-of-sample hedging effectiveness of currency futures for alternative models and hedging strategies". *Journal of Futures Markets, 17*, 817–837.

De Jong, P., & Apilado, V. (2008). The changing relationship between earnings expectations and earnings for value and growth stocks during Reg FD. *Journal of Banking and Finance, 33*, 435–442.

De La Pena, V., Ibragimov, R., & Jordan, S. (2004). Option bounds". *Journal of Applied Probability, 41*, 145–156.

De Long, B., Shleifer, A., Summers, L., & Waldman, R. (1990). Noise trader risk in financial markets. *Journal of Political Economy, 98*, 703–738.

De Santis, Giorgio. (1995). Volatility Bounds for Stochastic Discount Factors: Tests and Implications from International Financial Markets. working paper, University of Southern California January.

de Servigny, A., & Renault, O. (2004). *Measuring and Managing Credit Risk*. New York, NY: McGraw-Hill.

Dean, J., & Ghemawat, S. (2004). *"MapReduce: Simpli_ed Data Processing on Large Clusters" OSDI '04, 137–150*.

DeAngelo, H., & Masulis, R. (1980). Optimal capital structure under corporate and personal taxation. *Journal of Financial Economics, 8*, 3–27.

DeAngelo, H., & DeAngelo, L. (2006). The irrelevance of the MM dividend irrelevance theorem. *Journal of Financial Economics, 79*, 293–315.

DeAngelo, H., DeAngelo, L., & Skinner, D. J. (1996). Reversal of fortune: Dividend signaling and the disappearance of sustained earnings growth. *Journal of Financial Economics, 40*, 341–371.

DeAngelo, H., DeAngelo, L., & Stulz, R. M. (2006). Dividend policy and the earned/contributed capital mix: A test of the life-cycle theory. *Journal of Financial Economics, 81*, 227–254.

DeBondt, G. (2000). Credit Channels and Consumption in Europe: Empirical Evidence. BIS Working Paper No 69.

Dechow, P. M., & Skinner, D. J. (2000). Earnings management: Reconciling the views of accounting academics, practitioners, and regulators. *Accounting Horizons, 14*, 235–250.

Dechow, P. M., & Sloan, R. G. (1997). Returns to contrarian investment: Tests of the naïve expectations hypotheses. *Journal of Financial Economics, 43*, 3–28.

Dechow, P. M., Hutton, A. P., & Sloan, R. G. (1999). An Empirical Assessment of the Residual Income Valuation Model. *Journal of Accounting and Economics, 26*, 1–34.

Dechow, P., Kothari, S., & Watts, R. (1998). The relation between earnings and cash flows. *Journal of Accounting and Economics, 25*, 133–168.

Dechow, P. M., & Dichev, I. (2002). The quality of accruals and earnings: The role of accrual estimation errors. *The Accounting Review, 77*, 35–59.

Dechow, P. M., Sloan, R. G., & Sweeney, A. P. (1996). Causes and consequences of earnings manipulcation: an analysis of firms subject to enforcement actions by the SEC. *Contemporary Accounting Research, 13*, 1–36.

Deckro, R. F., Spahr, R. W., & Herbert, J. E. (1985). Preference trade-offs in capital budgeting decisions. *IIE Transactions, 17*, 332–337.

Dede, E., Govindaraju, M., Gunter, D., & Ramakrishnan, L. (2011). Riding the elephant: Managing ensembles with hadoop 4th Workshop on Many-Task Computing on Grids and Supercomputers, 49–58.

Degiannakis, S., Floros, C., & Dent, P. (2012). Forecasting value-at-risk and expected shortfall using fractionally integrated models of conditional volatility: International evidence. International Review of Financial Analysis, accepted manuscrtipt.

Deheuvels, P. (1981). *A Nonparametric Test for Independence*. Institut de Statistique: Universite de Paris.

DeLong, J., Bradford, A., Schleifer, L. H., Summers, H., & Waldmann, R. T. (1989). The size and incidence of the losses from noise trading. *Journal of Finance, 44*, 681–696.

DeLong, J. B., Shleifer, A., Summers, L., & Waldman, R. (1990). Noise Trader Risk in Financial Market. *Journal of Political Economy, 98*, 703–739.

Demarta, S., & McNeil, A. (2005). The t copula and related copulas. *International Statistical Review, 73*, 111–129.

Demartines, P., & H´erault, J. (1997). Curvilinear component analysis: A self-organizing neural network for nonlinear mapping of data sets". *Neural Networks, 8*, 148–154.

Demirer, R., & Kutan, A. (2006). "Does herding behavior exist in Chinese stock markets?" Journal of International Financial Markets. *Institutions and Money, 16*(2), 123–142.

Demirer, R., Kutanand, A., & Chen, C. (2010). Do investors herd in emerging stock markets?: evidence from the Taiwanese Market. *Journal of Economic Behavior & Organization, 76*(2), 283–295.

Dempsey, S. (1989). Predisclosure Information Search Incentives, Analyst Following and Earnings Announcement Price Response. *The Accounting Review, 64*, 748–57.

Demsetz, H. (1968). The Cost of Transacting. *Quarterly Journal of Economics, 82*, 33–53.

Deng, X. T., Li, Z. F., & Wang, S. Y. (2005). A minimax portfolio selection strategy with equilibrium. *European Journal of Operational Research, 166*, 278–292.

Deng, X. T., Wang, S. Y., & Xia, Y. S. (2000). Criteria models and strategies in portfolio selection. *Advanced Modeling and Optimization, 2*, 79–104.

Denis, D., & Sarin, A. (2001). Is the market surprised by poor earnings realizations following seasoned equity offerings?". *Journal of Financial and Quantitative Analysis, 36*(2), 169–193.

Denis, D. J., & Osobov, I. (2008). Why do firms pay dividends? *International evidence on the determinants of dividend policy, Journal of Financial Economics, 89*, 62–82.

Denis, D. J., Denis, D. K., & Sarin, A. (1994). The Information Content of Dividend Changes: Cash Flow Signaling. *Overinvestment, and Dividend Clienteles, Journal of Financial & Quantitative Analysis, 29*, 567–587.

Dennis, S., Nandy, D., & Sharpe, I. G. (2000). The determinants of contract terms in bank revolving credit agreements. *Journal of Financial and Quantitative Analysis, 35*, 87–110

Dentcheva, D., & Ruszczynski, A. (2003a). Optimization under nonlinear stochastic dominance. *Comptes Rendus de l'Academie Bulgare des Sciences, 56*(7), pp. 19–25

Dentcheva, D., & Ruszczynski, A. (2003b). Optimization with stochastic dominance constraints. *SIAM Journal on Optimization, 14*, 548–566.

Dentcheva, D., & Ruszczynski, A. (2004a). Optimality and duality theory for stochastic optimization problems with nonlinear dominance constraints. *Mathematical Programming, 99*, 329–350.

Dentcheva, D., & Ruszczynski, A. (2004b). Convexification of stochastic ordering constraints. *C. R. Acad. Bulgare Sci., 57*(3), 5–10.

Dentcheva, D., & Ruszczynski, A. (2004c). Semi-infinite probabilistic optimization: first order stochastic dominance constraints. *Optimization, 53*, 583–601.

Dentcheva, D., & Ruszczynski, A. (2006a). Inverse stochastic dominance constraints and rank dependent expected utility theory. *Mathematical Programming, 108*, 297–311.

Dentcheva, D., & Ruszczynski, A. (2006b). Portfolio optimization with stochastic dominance constraints. *Journal of Banking and Finance, 30*(2), pp. 433–451.

Dentcheva, D., & Ruszczynski, A. (2008). Duality between coherent risk measures and stochastic dominance constraints in risk-averse optimization. *Pacific Journal of Optimization, 4*, 433–446.

Dentcheva, D., & Ruszczynski, A. (2010). Inverse cutting plane methods for optimization problems with second order stochastic dominance constraints. *Optimization, 59*, 323–338.

Dentcheva, D. (2005). Optimization models with probabilistic constraints. In G. Calafiore & F. Dabbene (Eds.), *Probabilistic and Randomized Methods for Design under Uncertainty*. London: Springer-Verlag.

Dentcheva, D., & Ruszczynski, A. (2003). Optimization under linear stochastic dominance. *Comptes Rendus de l'Academie Bulgare des Sciences, 56*(6), 6–11

Dermoune, A., Djehiche, B., & Rahmania, N. (2008). Consistent estimators of the smoothing parameter in the hodrick-prescott filter. *Journal of the Japan Statistical Society, 38*, 225–241.

Desai, M. A., Foley, C. F., & Hines, J. R. (2004). A multinational perspective on capital structure choice and internal capital markets. *Journal of Finance, 59*, 2451–2487.

Deuskar, P., Gupta, A., & Subrahmanyam, M. (2003). Liquidity Effects and Volatility Smiles in Interest Rate Option Markets, Working paper, New York University

Devaraj, S., & Kohli, R. (2003). Information technology payoff paradox and system use: Is actual usage the missing link?". *Management Science., 49*, 273–289.

Devenow, A., & Welch, I. (1996). Rational Herding in Financial Economics. *European Economic Review, 40*, 603–615.

DeWayne, L. S. (2009). Developing a lean performance score. *Strategic Finance, 91*, 34–39.

Dhaliwal, D., Subramanyam, K., & Trezevant, R. (1999). Is comprehensive income superior to net income as a measure of firm performance?". *Journal of Accounting and Economics, 26* (January), 43–67.

Dhrymes, P. J., & Kurz, M. (1967). Investment, dividend, and external finance behavior of firms, in Determinants of Investment Behavior, National Bureau of Economic Research.

Dhrymes, P. J., Friend, I., & Bulent Gultekin, N. (1984). A Critical Reexamination of the Empirical Evidence on the Arbitrage Pricing Theory. *Journal of Finance, 39*(2), 323–346, (June).

Diamond, D. W. (1984). Financial intermediation and delegated monitoring. *Review of Economic Studies, 51*, 393–414.

Diamond, D. W., & Dybvig, P. H. (1983). Bank runs, deposit insurance, and liquidity. *Journal of Political Economy, 91*, 401–419.

Diamond, D. W., & Rajan, R. (2000). A theory of bank capital. *Journal of Finance, 55*, 2431–2465.

Diamond, D. B., Jr., & Smith, B. (1985). Simultaneity in the Market for Housing Characteristics". *Journal of Urban Economics, 17*, 280–92.

Diba, B. T., & Grossman, H. (1987). On the inception of rational bubbles". *Quarterly Journal of Economics, 102*, 197–700.

Diba, B. T., & Grossman, H. (1985). Rational bubbles in stock prices?, NBER working paper no. 1779.

Diba, B.T., & Grossman, H. I. (1988). Explosive rational bubbles in stock prices? *The American Economic Review, 78*, 520–530.

Dickey, D. A., & Fuller, W. A. (1979). Distribution of the estimators for autoregressive time series with a unit root. *Journal of the American Statistical Association, 74*, 427–31.

Dickey, D. A., & Fuller, W. A. (1981). Likelihood ratio statistics for autoregressive time series with a unit root. *Econometrica, 49*, 1057–1072.

Diebold, F. X. (1986). Comment on modeling the persistence of conditional variance. *Econometric Reviews, 5*, 51–56.

Diebold, F. X. (1989). Forecast combination and encompassing: Reconciling Two Divergent Literatures. *International Journal of Forecasting, 5*, 589–592.

Diebold, F. X., & Inoue, A. (2001). Long memory and regime switching. *Journal of Econometrics, 105*, 131–159.

Diebold, F. X., & Lopez, J. A. (1996). Forecast Evaluation and Combination. In G. S. Maddala & C. R. Rao (Eds.), *Handbook of Statistics* (pp. 241–268). Amsterdam: North-Holland.

Diebold, F. X., & Pauly, P. (1987). Structural change and the combination of forecasts. *Journal of Forecasting, 6*, 21–40.

Diebold, F. X., & Pauly, P. (1990). The use of price information in forecast combination. *International Journal of Forecasting, 6*, 503–508.

Diebold, F. X., & Rudebusch, G. D. (1991). Forecasting output with the composite leading index: an ex ante analysis. *Journal of the American Statistical Association, 86*, 603–610.

Diebold, F. X., & Li, C. (2006). Forecasting the Term Structure of Government Bond Yields. *Journal of Econometrics, 130*, 337–364.

Diebold, F. X., & Mariano, R. S. (1995). Comparing Predictive Accuracy. *Journal of Business and Economic Statistics, 13*, 253–263.

Diebold, F. X., Tay, A. S., & Wallis, K. D. (1998). "Evaluating Density Forecasts of In‡ation: The Survey of Professional Forecasters." in Festschrift in Honor of C.W.J. Granger, eds. R.F. Engle and H. White, Oxford University Press, Oxford.

Diebold, F. X., Li, C., & Yue, V. (2008). Global yield curve dynamics and interactions: a generalized Nelson-Siegel approach. *Journal of Econometrics, 146*, 351–363.

Diebold, F. X., Hahn, J., & Tay, A. S. (1999). Multivariate Density Forecast Evaluation and Calibrationin Financial Risk Management: High Frequency Returns on Foreign Exchange. *Review of Economics and Statistics, 81*, 661–673.

Diebold, F. X., Ji, L., & Li, C. (2006). A three-factor yield curve model: non-affine structure, systematic Risk Sources, and Generalized Duration. In L. R. Klein (Ed.), *Macroeconomics, Finance and Econometrics: Essays in Memory of Albert Ando 240-274*. Cheltenham: U.K. Edward Elgar.

Diebold, F. X., Gunther, T., & Tay, A. S. (1998). Evaluating Density Forecasts with Applications to Finance and Management. *International Economic Review, 39*, 863–883.

Diebold, F., Gardeazabai, J., & Yilmaz, K. (1994). On cointegration and exchange rate dynamics. *Journal of Finance, 49*, 727–735.

Diewert, W. E. (2002), Hedonic Producer Price Indexes and Quality Adjustment, Discussion Paper No.: 02-14, University of British Columbia, Vancouver

Dimson, E., & Mussavian, M. (1998). A brief history of Market efficiency. *European Financial Management, 4*(1), 91–103.

Dimson, E., & Mussavian, M. (1999). Three centuries of asset pricing". *Journal of Banking and Finance, 23*(12), 1745–1769.

Dimson, E. (1979). Risk Measurement when Shares are Subject to Infrequent Trading. *Journal of Financial Economics, 7*, 197–266.

Ding, Z., Granger, C. W. J., & Engle, R. F. (1993). A long memory property of stock market returns and a new model. *Journal of Empirical Finance, 1*, 83–106.

Dittmann, I., & Maug, E. (2007). Lower salaries and no options? On the optimal structure of executive pay. *Journal of Finance, 62*, 303–343.

Dittmann, I., & Yu, K. C. (2011). How Important are Risk-Taking Incentives in Executive Compensation? Working Paper, Erasmus University, Rotterdam, The Netherlands

Dittmann, I., Maug, E., & Spalt, O. (2010). Sticks or carrots: Optimal CEO compensation when managers are loss-averse. *Journal of Finance, 65*, 2015–2050.

DiValentino, L. M. (1994). "Preface." Financial Markets, Institutions & Instruments. Estimating Cost of Capital: Methods and Practice V. 3. N. 3, 6–8 (August).

Diwan, I., Errunza, V., & Senbet, L. W. (1995). Diversification Benefits of Country Funds. In Mike Howell (Ed.). *Investing in Emerging Markets,* Euromoney/World Bank.

Djankov, S., La Porta, R., Lopez-de-Silvanes, F., & Shleifer, A. (2008). The Law and economics of self-dealing. *Journal of Financial Economics, 88*, 430–465.

Doetsch, G. (1970). *Introduction to the theory and application of the Laplace transformation*. New York: Springer-Verlag.

Donoho, D. L., & Johnstone, I. M. (1994). Ideal spatial adaptation by wavelet shrinkage. *Biometrika, 81*(3), 425–455.

Donovan, G., & Butry, D. (2010). Trees in the city: Valuing street trees in Portland, Oregon. *Landscape and Urban Planning, 94*(2), 28 February, Pages 77–83

Doong, S. C., Yang, S. Y., & Wang, A. T. (2005). The dynamic relationship and pricing of stocks and exchange rates: Empirical evidence from Asian emerging markets". *Journal of the American Academy of Business, 7*(1), 118–123.

Dopuch, N., & Drake, D. F. (1964). Accounting implications of a mathematical programming approach to the transfer price problem. *Journal of Accounting Research, 10–24*.

Doran, H. E., & Rambaldi, A. N. (1997). Applying linear time-varying constraints to econometric models: With an application to demand systems. *Journal of Econometrics, 79*, 83–95.

Doran, H., & Schmidt, P. (2006). GMM Estimators with Improved Finite Sample Properties Using Principal Components of the Weighting Matrix, with an Application to the Dynamic Panel Data Model. *Journal of Econometrics, 133*, 387–409.

Dowd, K. (1998). Beyond Value at Risk: The new science of risk management, John Wiley & Sons.

Dowd, K. (2005). *Copulas and coherence* (pp. 123–127). Fall: Journal of Portfolio Management.

Dowd, K. (2000). Assessing VaR accuracy. *Derivatives Quarterly, Spring, 2000*, 61–63.

Dowd, K. (2002). Measuring Market Risk. Wiley

Dowell, R. (1985). Risk Preference and the Work-Leisure Trade-off. *Economic-Inquiry, 23*, 691–701.

Downing, C., Underwood, S., & Xing, Y. (2009). An intraday analysis of the relative informational efficiency of stocks and bonds. *Journal of Financial and Quantitative Analysis, 44*, 1081–1102.

Doyle, J., Lundholm, R., & Soliman, M. (2003). The predictive value of expenses excluded from pro forma earnings. *Review of Accounting Studies, 8*, 145–174.

Doyle, J., Ge, W., & McVay, S. (2007a). Accruals quality and internal control over financial reporting. *The Accounting Review, 82*, 1141–1170.

Doyle, J., Ge, W., & McVay, S. (2007b). Determinants of weaknesses in internal control over financial reporting and the implications for earnings quality. *Journal of Accounting and Economics, 44*, 193–223.

Dreiman, M., & Follain, J. R. (2000). *Drawing Inferences about Housing Supply Elasticity from House Price Responses to Income Shocks"*. Processed: Freddie Mac.

Driessen, J. (2005). Is default event risk priced in corporate bonds?". *Review of Financial Studies, 18*, 165–195.

Droms, W. G., & Walker, D. A. (1994). Investment Performance of International Mutual Funds. *Journal of Financial Research, 17*, Spring, 1–14.

Drost, F. C., & Nijman, T. E. (1993). Temporal Aggregation of GARCH Processes. *Econometrica, 61*, 909–27.

Drucker, P. F. (1993). We need to measure. *not count. Wall Street Journal April 13. Also Expert Choice Voice*(July).

Duan, J., & Simonato, J. (1999). Estimating and testing exponential-affine term structure models by Kalman Filter. *Review of Quantitative Finance and Accounting, 13*, 111–135.

Duan, J., Moreau, A. F., & Sealey, C. W. (1992). Fixed rate deposit insurance and risk-shifting behavior at commercial banks. *Journal of Banking and Finance, 16*, 715–742.

Duan, J.-C. (1995). The GARCH option pricing model. *Mathematical Finance, 5*, 13–32.

Duan, J.-C. (1997). Augmented GARCH(p, q) process and its diffusion limit. *Journal of Econometrics, 79*, 97–127.

Duan, J. C. (2003). A Specification Test for Time Series Models by a Normality Transformation Working Paper, University of Toronto

Duarte, J., & Young, L. (2009). Why is PIN priced? *Journal of Financial Economics, 91*(2), 119–138.

Duarte, J. (2004). Evaluating an Alternative Risk Preference in Affine Term Structure Models. *Review of Financial Studies, 17*, 379–404.

Duarte, J. (2008). Mortgage-Backed Securities Refinancing and the Arbitrage in the Swaption Market. *Review of Financial Studies, 21*, 1689–1731.

Duesenberry, J. (1949). *Income, saving, and the theory of consumer behavior*. Cambridge: Harvard University Press.

Duffe, D., Pan, J., & Singleton, K. (2000). Transform analysis and asset pricing for affine jump-diffusions. *Econometrica, 68*, 1343–1376.

Duffee, G. R. (1999). Estimating the price of default risk. *The Review of Financial Studies, 12*, 197–226.

Duffee, G., Kupiec, P., & White, A. P. (1990). A Primer on Program Trading and Stock Price Volatility: A Survey of the Issues and Evidence. FEDS Board of Governors of the Federal Reserve System, Working Paper No 109 (Feb).

Duffee, G. R. (2002). Term Premia and Interest Rate Forecasts in Affine Models. *Journal of Finance, 57*, 405–443.

Duffie, D., & Singleton, K. (1993). Simulated Moment Estimation of Markov Models of Asset Prices. *Econometrica, 61*, 929–952.

Duffie, D., & Singleton, K. J. (1999a). Modelingterm structures of defaultable bonds. *Review of Financial StudiesSpecial, 12*(4), 687–720.

Duffie, D., & Singleton, K. J. (1999b). Simulating Correlated Defaults, Working Paper, Graduate School of Business, Stanford University

Duffie, D., & Kan, R. (1996). A Yield-Factor Model of Interest Rates. *Mathematical Finance, 6*, 379–406.

Duffie, D. (2002). *Dynamic Asset Pricing Theory*. New Jersey: Princeton University Press.
Duffie, D., & Pan, J. (1997). An overview of value at risk. *Journal of Derivatives, 4*, 7–49.
Duffie, D., Pan, J., & Singleton, K. (2000). Transform analysis and option pricing for affine jump-diffusions. *Econometrica, 68*, 1343–1376.
Duffie, D., & Lando, D. (2001). Term Structures of Credit Spreads with Incomplete Accounting Information. *Econometrica, 69*(3), 633–664.
Dufour, J. (2003). Identification, weak instruments and statistical inference in Econometrics. *Canadian Journal of Economics, 36*, 767–808.
Dufour, R., & Maag, U. R. (1978). Distribution Results for Modified Kolmogorov-Smirnov Statistics for Truncated or Censored Samples. *Technometrics, 20*, 29–32.
Dufwenberg, M., Lindqvist, T., & Moore, E. (2005). Bubbles and Experience: An Experiment". *The American Economic Review, 95*, 1731–1737.
Dumas, B., Fleming, J., & Whaley, R. E. (1998). Implied volatility functions: empirical tests. *Journal of Finance, 53*, 2059–2106.
Dumas, B., Fleming, J., & Whaley, R. (1998). Implied Volatility Functions: Empirical Tests, *Journal of Finance, 53*(6), 2059–2106.
Dunbar, C. G. (1995). The use of warrants as underwriter compensation in initial public offerings. *Journal of Financial Economics, 38*, 59–78.
Duncan, J., & Isaac, R. M. (2000). "Asset Markets: How They Are Affected by Tournament Incentives for Individuals", American Economic Review. *American Economic Association, 90*(4), 995–1004.
Dunn, P. C., & Theisen, R. D. (1983). How consistently do active managers win?". *Journal of Portfolio Management, 9*, 47–51.
Duong, D., & Swanson, N. R. (2010). Empirical Evidence on Jumps and Large Fluctuations in Individual Stocks Working Paper, Rutgers University
Duong, D., & Swanson, N. R. (2011). Volatility in Discrete and Continuous Time Models: A Survey with New Evidence on Large and Small Jumps. *Missing Data Methods: Advances in Econometrics, 27B*, 179–233.
Durand, R., Newby, R., & Sanghani, J. (2006). An Intimate Portrait of the Individual Investor, working paper
Durbin, J. (1954). Errors in variables. *Review of the International Statistical Institute, 22*, 23–32.
Durbin, J., & Koopman, S. J. (2001). *Time Series Analysis by State Space Methods*. New York, NY: Oxford University Press.
Durrett, R. (2005). *Probability: Theory and Examples* (3rd ed.). Brooks/Cole, Belmont, CA: Thomson.
Duru, A., & Reeb, D. (2002). International diversification and analysts' forecast accuracy and bias. *The Accounting Review, 77*(2), 415–433.
Duru, A., & Reeb, D. (2002). International diversification and analysts' forecast accuracy and bias. *The Accounting Review, 77*, 415–433.
Dutoit, L. C. (2007). Heckman's selection model, endogenous and exogenous switching models, a survey, The Selected Works of Laure C Dutoit Available at: http://works.bepress.com/laure_dutoit/3
Dybvig, P., & Ross, S. (1985). Differential information and performance measurement using a security market line. *Journal of Finance, 40*, 383–399.
Dyer, R. F., & Forman, E. H. (1991). *An analytic approach to marketing decisions*. Englewood, NJ: Prentice Hall.
Eames, M., & Glover, S. (2003). Earnings predictability and the direction of analysts' earnings forecast errors. *The Accounting Review, 78*, 707–724.
Easley, D., Hvidkjaer, S., & O'Hara, M. (2002). Is information risk a determinant of asset returns? *Journal of Finance, 57*, 2185–2221.
Easley, D., López de Prado, M., & O'Hara, M. (2012). Flow toxicity and liquidity in a high frequency world. *Review of Financial Studies, 25*(5), 1457–1493.
Easley, D., Kiefer, N., O'Hara, M., & Paperman, J. B. (1996). Liquidity, information and infrequently traded stocks. *Journal of Finance, 51*, 1405–1436.

Easley, D., Engle, R. F., O'Hara, M., & Wu, L. (2008). Time-varying arrival rates of informed and uninformed traders. *Journal of Financial: Econometrics.*

Easley, D., O'Hara, M., & Srinivas, P. S. (1998). Option volume and stock prices: evidence on where informed traders trade. *Journal of Finance, 53*, 431–466.

Easterbrook, F. H. (1984). Two agency-cost explanations of dividends. *American Economic Review, 74*, 650–659.

Easton, P., Taylor, G., Shroff, P., & Sougiannis, T. (2002). Using forecasts of earnings to simultaneously estimate growth and the rate of return on equity investment. *Journal of Accounting Research, 40*, 657–676.

Easton, P., & Zmijewski, M. (1989). Cross-Sectional Variation in the Stock Market Response to Accounting Earnings Announcements. *Journal of Accounting and Economics, 29*, 19–36.

Easton, P., & Harris, T. (1991). Earnings as an Explanatory Variable for Returns". *Journal of Accounting Research, 29*, 19–36.

Eaton, J., Gersovitz, M., & Stiglitz, J. E. (1986). The Pure Theory of Country Risk. *European Economic Review, 30*, 481–513.

Ebens, H. (1999). Realized stock volatility. Working paper, John Hopkins University

Eberhart, A., Maxwell, W., & Siddique, A. (2004). An examination of long-term abnormal stock returns and operating performance following R&D increases. *Journal of Finance, 59*(2), 623–650.

Eberhart, A. C., Altman, E. I., & Aggarwal, R. (1999). The equity performance of firms emerging from bankruptcy. *Journal of Finance., 54*, 1855–1868.

Eberlein, E., & Keller, U. (1995). Hyperbolic distributions in finance". *Bernoulli, 1*, 281–299.

Eberlein, E., Keller, U., & Prause, K. (1998). New insights into smile, mispricing and value at risk: the hyperbolic model". *Journal of Business, 71*, 371–406.

Eccles, R. G. (1983). Control with fairness in transfer pricing. *Harvard Business Review,* 149–161.

Eccles, R. G. (1991). The performance measurement manifesto. *Harvard Business Review,* January-February, 131–137.

Eccles, R. G., & Pyburn, P. J. (1992). Creating a comprehensive system to measure performance. *Management Accounting* October, 41–44.

Eckbo, B. E., & Norli, O. (2005). Liquidity risk, leverage, and long-run IPO returns. *Journal of Corporate Finance, 11*, 1–35.

Eckbo, B. E., Masulis, R. W., & Norli, O. (2000). Seasoned public offerings: resolutions of the 'new issues puzzle'. *Journal of Financial Economics, 56*, 251–291.

Eckbo, B. E. (1986). Valuation effects of corporate debt offerings. *Journal of Financial Economics, 5*, 119–52.

Economou, F., Kostakis, A., & Philippas, N. (2011). Cross-Country Effects in Herding Behavior: Evidence from Four South European Markets. *Journal of International Financial Markets, Institutions and Money, 21*(3), 443–460.

Eddie, W., Cheng, L., & Li, H. (2001). Analytic hierarchy process: an approach to determine measures for business performance. *Measuring Business Excellence, 5*, 30–36.

Ederington, L. H. (1985). Classification models and bond ratings. *The Financial Review, 20*, 237–261.

Ederington, L. H. (1979). The hedging performance of the new futures markets. *Journal of Finance, 34*, 157–170.

Ederington, L. H., & Goh, J. C. (1999). Cross-sectional Variation in the Stock Market Reaction to Bond Rating Changes. *The Quarterly Review of Economics and Finance, 39*(1), 101–112.

Ederington, L. H., & Lee, J. H. (1993). How markets process information: news releases and volatility. *Journal of Finance, 49*, 1161–1191.

Edwards, F. R. (1988). Futures Trading and Cash Market Volatility: Stock Index and Interest rate futures. *Journal of Futures Markets, 8*, 421–39.

Efendi, J., Srivastava, A., & Swanson, E. (2007). Why do corporate managers misstate financial statements? The role of option compensation and other factors. *Journal of Financial Economics, 85*(3), 667–708.

Efon, B. (1981). Censored Data and the Bootstrap". *Journal of American Statistical Association, 76*(374), 312–319.

Efron, B., & Tibshirani, R. J. (1993). *An introduction to the bootstrap*. Chapman &: Hall.

Efron, B., & Tibshirani, R. (1986). Bootstrap Methods for Standard Erros, Confidnce Intervals, and Other Measures of Statistical Accuracy. *Statistical Science, 1*(1), 54–75.

Efron, B., Hastie, T., Johnstone, I., & Tibshirani, R. (2004). Least angle regression (with discussion). *Annals of Statistics, 32*, 409–499.

Efron, B., & Tibshirani, R. J. (1993). *Introduction to the bootstrap*. New York: Chapman and Hall.

Efron, B. (1979). Bootstrap methods: Another looks at the jackknife. *Annals of Statistics, 7*, 1–26.

Egozcue, M., & Wong, W. K. (2010). Gains from diversification on convex combinations: a majorization and stochastic dominance approach. *European Journal of Operational Research, 200*, 893–900.

Egozcue, M., Fuentes García, L., Wong, W. K., & Zitikis, R. (2011). Do investors like to diversify? A study of Markowitz preferences. *European Journal of Operational Research, 215*, 188–193.

Ehrbeck, T., & Waldmann, R. (1996). Why are professional forecasters biased? agency versus behavioral explanations. *The Quarterly Journal of Economics, 111*, 21–40.

Ehrenberg, R. G., & Bognanno, M. L. (1990). Do tournaments have incentive effects?". *Journal of Political Economy, 98*, 1307–1324.

Ehrhardt, M. C. (1994). *The Search for Value: Measuring the Company's Cost of Capital*. Boston: Harvard Business School Press.

Ehrmann, M. (1998). Will EMU generate asymmetry ? Comparing monetary policy transmissions across European countries. European University Institute Working Paper ECO 98/28.

Eichler, M., Grothe, O., Manner, H., & Tuerk, D. (2012). Modeling spike occurrences in electricity spot prices for forecasting. Working paper, RM/12/029, Maastricht University.

Eisenbeis, R. A., & Avery, R. B. (1972). *Discriminant Analysis and Classification Procedure*. Lexington, Massachusetts: Lexington Books.

El Karoui, N., & Quenez, M. C. (1995). Dynamic programming and pricing of contingent claims in an incomplete market. *SIAM Journal of Control Optimization, 33*, 29–66.

Elango, R. (2003). *Which fund yields more returns?* Mutual Funds: A Comparative analysis on the NAV performance of select public v/s private/foreign open-ended mutual fund schemes in India.

Eliasson, A. (2002). Sovereign Credit Ratings. Working Papers 02-1, Deutsche Bank

Eling, M., & Schuhmacher, F. (2007). Does the choice of performance measure influence the evaluation of hedge funds? *Journal of Banking and Finance, 31*, 2632–2647.

Elizalde, A. (2005). Do we need to worry about credit risk correlation?". *Journal ofFixed Income, 15*(3), 42–59.

Elliott, G., & Ito, T. (1999). Heterogeneous expectations and tests of efficiency in the yen/dollar forward exchange rate market. *Journal of Monetary Economics, 43*, 435–456.

Elliott, G., & Timmermann, A. (2002). *Optimal Forecast Combination under Regime Switching*. Manuscript: Department of Economics, University of California, San Diego.

Ellis, K., Michaely, R., & O'Hara, M. (2000). The accuracy of trade classification rules: Evidence from Nasdaq. *Journal of Financial and Quantitative Analysis, 35*, 529–551.

Elton, E. J., Gruber, M. J., Brown, S. J., & Goetzmann, W. N. (2006). *Modern Portfolio Theory and Investment Analysis*. New York: John Wiley & Sons.

Elton, E. J., & Gruber, M. J. (1995). *Modern Portfolio Theory and Investment Analysis* (5th ed.). Wiley &: Sons.

Elton, E. J., Gruber, M. J., & Blake, C. R. (1996). Market Timing Ability and Volatility Implied in investment Newsletters Asset Allocation Recommendations. *Journal of Financial Economics, 42*, 397–421

Elton, E. J., & Gruber, M. J. (1994). Introduction. Financial Markets, Institutions & Instruments. *Estimating Cost of Capital: Methods and Practice, 3*(3), 1–5, (August).

Elton, E. J., Gruber, M. J., & Mei, J. (1994). Cost of Capital Using Arbitrage Pricing Theory: A Case Study of Nine New York Utilities. Financial Markets, Institutions & Instruments. *Estimating Cost of Capital: Methods and Practice, 3*(3), 46–73, (August).

Elyasiani, E., & Iqbal, M. (1998). Sensitivity of the bank stock returns distribution to changes in the level and volatility of interest rate: A GARCH-M model. *Journal of Banking & Finance, 22*, 535–563.

Emanuel, D., & Mcbeth, J. (1982). Further Results on the Constant Elasticity of Variance Call Option Pricing Model. *Journal of Financial Quantitative Analysis, 17*, 533–554.

Embrechts, P., McNeil, A., & Straumann, D. (1999). *Correlation and dependence in risk management: Properties and pitfalls*. ETHZ Zentrum, Zurich: Mimeo.

Embrechts, P., Frey, R., & McNeil, A. (2005). *Quantitative Risk Management*. Princeton, NJ: Princeton University Press.

Embrechts, P., Lindskog, F., & McNeil, A. (2003). Modelling dependence with copulas and applications to risk management". In S. T. Rachev (Ed.), *Handbook of Heavy Tailed Distributions in Finance* (pp. 329–384). Amsterdam: Elsevier/North-Holland.

Enders, W. (2009). *Applied Econometric Time Series* (3rd ed.). New York: Wiley.

Engel, C. (1984). Testing for the absence of expected real profits from forward market speculation. *Journal of International Economics, 17*, 299–308

Engel, C., & Hamilton, J. D. (1990). Long swings in the dollar: Are they in the data and do markets know it?". *American Economic Review, 80*, 689–713.

Engle, R. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica, 50*, 987–1007.

Engle, R. F., & Sheppardy, K. (2008). "Evaluating the Specification of Covariance Models for Large Portfolios." New York University, Working paper

Engle, R. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance with estimates of the variance of United Kingdom inflation". *Econometrica, 50*, 987–1007.

Engle, R. (2002a). Dynamic conditional correlation: a simple class of multivariate GARCH models. *Journal of Business and Economic Statistics, 20*, 339–350.

Engle, R. (2002b). New frontiers for ARCH models. *Journal of Applied Econometrics, 17*, 425–446.

Engle, R. (2009). *Anticipating Correlations: A New Paradigm for Risk Management*, Princeton University Press

Engle, R., & Granger, C. W. S. (1987). Cointegration and error correction representation, estimation and testing. *Econometrica, 55*, 251–276.

Engle, R., & Gallo, G. M. (2006). A multiple indicators model for volatility using intra-daily data. *Journal of Econometrics, 131*, 3–27.

Engle, R., & Russell, J. R. (1998). Autoregressive conditional duration: A new model for irregularly spaced transaction data. *Econometrica, 66*, 1127–62.

Engle, R., & Manganelli, S. (2004). CAViaR: Conditional autoregressive value at risk by regression quantiles. *Journal of Business and Economic Statistics, 22*, 367–381.

Engle, R. F., & Rosenberg, J. (1994). Hedging options in a GARCH environment: Testing the term structure of stochastic volatility models, NBER Working Paper, No. 4958

Engle, R. F. (1982). Autoregressive conditional heteroskedasticity with estimates of variance of United Kingdom ination. *Econometrica, 50*, 987–1008.

Engle, R. F. (2002). "Dynamic conditional correlation: a simple class of multivariate generalized autoregressive conditional heteroskedasticity models''. *Journal of Business and Economic Statistics, 20*, 339–350.

Engle, R. F., & Rangel, J. G. (2008). The spline-GARCH models for low-frequency volatility and its global macroeconomic causes. *Review of Financial Studies, 21*, 1187–1222.

Engle, R. F., & Mezrich, J. (1996). GARCH for groups. *Risk August, 1996*, 36–40.

Engle, R. F., & Bollerslev, T. (1986). Modeling the persistence of conditional variances. *Econometric Reviews, 5*, 1–50.

Engle, R. F., & Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *Journal of Finance, 48*, 1749–78.

Engle, R. F. (1982). Autoregressive conditional heteroscedasticity with estimates of variance of United Kingdom inflation. *Econometrica, 50*, 987–1008.

Engle, R. F., & Manganelli, S. (1999). "CAViaR: Conditional autoregressive value at risk by regression quantiles," working paper, University of California, San Diego

Engle, R. F., & Ng, V. K. (1993). Measuring and testing the impact of news on volatility. *Journal of Finance, 48*, 1749–1778.

Engle, R. F., Lilien, D. M., & Robins, R. P. (1987). Estimating time varying risk premia in the term structure: The ARCH M model. *Econometrica, 55*, 391–407.

Engle, R. (1982). Autoregressive conditional heteroskedasticity with estimates of the variance of U.K. inflation". *Econometrica, 50*, 987–1008.

Engle, R. (2002). Dynamic conditional correlation: A simple class of multivariate GARCH models". *Journal of Business & Economic Statistics, 20*(3), 339–350.

Engle, R., & Sheppard, K. (2001). "Theoretical and empirical properties of dynamic conditional correlation multivariate GARCH," *NBER Working Paper* 8554.

Engle, R., Lilien, D., & Robins, R. (1987). Estimating Time-Varying Risk Premia in the Term Structure: The ARCH-M Model. *Econometrica, 55*, 391–408.

Engle, R., Gallo, G. M., & Velucchi, M. (2012). Volatility spillovers in east Asian financial markets: A MEM-based approach. *Review of Economics and Statistics, 94*, 222–223.

Engle, R., Granger, C., & Hallman, J. (1989). Merging Short and Long-Run Forecasts: An Application of Seasonal Cointegration to Monthly Electricity Sales Forecasting". *Journal of Econometrics, 50*, 297–327.

Engle, R., Focardi, S. M., & Fabozzi, F. J. (2008). ARCH/GARCH Models in applied financial econometrics. In F. J. Fabozzi (Ed.), *Handbook of finance, vol III*. Hoboken, New Jersey: John Wiley & Sons.

Engle, R., Ito, T., & Lin, W. L. (1994). Do bulls and bears move across borders? International transmission of stock returns and volatility". *Review of Financial Studies, 7*, 507–538.

Epple, D., Romano, R., & Sieg, H. (2006). Admission, Tuition, and Financial Aid Policies in the Market for Higher Education. *Econometrica, 74*(4), 885–928.

Epps, T. W., & Epps, L. M. (1976). The stochastic dependence of security price changes and transaction volumes: implications for the mixture-of-distribution hypothesis. *Econometrica, 44*, 305–321.

Epstein, L. G., & Zin, S. E. (1991). "Risk Aversion, and the Temporal Behavior of Consumption and Asset Returns: An Empirical Analysis," *Journal of Political Economy, 99*, 263–86.

Eraker, B., Johannes, M., & Polson, N. (2000). The impact of jumps in volatility and return. Working paper, University of Chicago.

Eraker, B. (2004). Do stock prices and volatility jump? Reconciling evidence from spot and option prices. *Journal of Finance, 59*, 1367–1404.

Eraker, B. (2003). Do Stock Prices and Volatility Jump? *Reconciling Evidence from Spot and Option Prices, Journal of Finance, 59*, 1367–1404.

Eraker, B., Johannes, M., & Polson, N. (2003). The impact of jumps in volatility and returns. *Journal of Finance, 53*, 1269–1300.

Erb, C. B., Harvey, C. R., & Viskanta, T. E. (1994). "Forecasting international equity correlations''. *Financial Analysts Journal, 50*, 32–45.

Ergungor, O. E. (2004). Market- vs. bank-based financial systems: do rights and regulations really matter? *Journal of Banking and Finance, 28*, 2869–2887.

Erickson, T., & Whited, T. (2000). Measurement Error and the Relationship between Investment and Q. *Journal of Political Economy, 108*, 1027–1057.

Erickson, T., & Whited, T. (2002). Two-Step GMM Estimation of the Errors-In-Variables Model Using High-Order Moments. *Econometric Theory, 18*, 776–799.

Errais, E., Giesecke, K., & Goldberg, L. R. (2010). Affine point processes and portfolio credit risk. *SIAM J. Financial Math., 1*, 642–665.

Errunza, V., & Losq, E. (1985). International Asset Pricing Under Mild Segmentation: Theory and Test. *Journal of Finance, 40*, 105–124.

Errunza, V., Senbet, L. W., & Hogan, K. (1998). The Pricing of Country Funds From Emerging Markets: Theory and Evidence. *International Journal of Theoretical and Applied Finance, 1*, 111–143.

Escribano, A. J., Peaea, J., & Villaplana, P. (2002). Modelling Electricity Prices: International Evidence. Working paper, Universidad Carlos III de Madrid

Esty, B. C. (1998). The impact of contingent liability on commercial bank risk taking. *Journal of Financial Economics, 47*, 189–218.

Eun, C. S., Kolodny, R., & Resnick, B. G. (1991). U.S. Based International Mutual Funds: A Performance Evaluation. *The Journal of Portfolio Management, 17*, Spring, 88–94.

Evans, G. W. (1991). Pitfalls in testing for explosive bubbles in asset prices. *American Economic Review, 81*, 922–930.

Evans, M. D. (1994). Expected returns, time-varying risk, and risk premia. *Journal of Finance, 49*, 655–679.

ExecuComp. (2008). Standard & Poor's ExecuComp website: http://www.compustat.com/products.aspx?id=2147492873

Fair, R. C., & Shiller, R. J. (1989). The informational content of ex ante forecasts. *Review of Economics and Statistics, 71*, 325–331.

Fair, R. C., & Shiller, R. J. (1990). Comparing information in forecasts from econometric models. *American Economic Review, 80*, 375–389.

Fair, R. C. (2007). Evaluating inflation targeting using a macroeconometric model. *Economics: The Open-Access, Open-Assessment E-Journal, Vol., 1.*

Fair, R. C., & Jaffee, D, M. (1972). Methods of Estimation of Market in Disequilibrium. *Econometrica, 40*, 497–514

Faires, J. D., & Burden, R. (2003). *Numerical methods*. (3rd ed.). Pacific Grove, CA : Tomson Learning.

Fairley, W. B. (1979). Investment Income and Profit Margins in Property-Liability Insurance: Theory and Empirical Results. *Bell Journal of Economics, 10*, (1), 191–210, (Spring).

Fama, E., & French, K. (1997). Industry Costs of Equity. *Journal of Financial Economics, 43*, 153–193.

Fama, E., & MacBeth, J. (1973). Risk, Return and Equilibrium: Empirical Tests. *Journal of Political Economy, 81*, 607–636.

Fama, E. F. (1972). Components of Investment Performance". *Journal of Finance, 27*, 551–567.

Fama, E. (1965). The Behavior of Stock-Market Prices. *The Journal of Business, 38*, 34–105.

Fama, E. (1981). Stock returns, real activity, inflation, and money. *The American Economic Review, 71*, 545–565.

Fama, E., & French, K. (1992). The cross-section of expected returns. *Journal of Finance, 47*, 427–466.

Fama, E., & French, K. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*, 3–56.

Fama, E., & French, K. (1996). Multifactor explanations of asset pricing anomalies. *Journal of Finance, 51*, 55–84.

Fama, E., & French, K. (1997). Industry costs of equity. *Journal of Financial Economics, 43*, 153–193.

Fama, E., & French, K. (1998). Market efficient, long-term returns, and behavioral finance. *Journal of Financial Economics, 49*, 283–306.

Fama, E., & Bliss, R. (1986). The information in long-maturity forward rates. *American Economic Review, 77*, 680–692.

Fama, E. F. (1971). Parameter estimates for symmetric stable distributions. *Journal of the American Statistical Association, 66*, 331–338.

Fama, E. F. (1984). Forward and spot exchange rates. *Journal of Monetary Economics, 14*, 319–338.

Fama, E. F. (1998). Market efficiency, long-term returns and behavioral finance. *Journal of Financial Economics, 49*, 283–306.

Fama, E. F., & French, K. R. (1988). Permanent and temporary components of stock prices. *Journal of Political Economy, 96*, 246–273.

Fama, E. F., & French, K. R. (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics, 25*, 23–49.

Fama, E. F., & French, K. R. (1992). The cross-section of expected stock returns. *Journal of Finance, 47*, 427–465.
Fama, E. F., & French, K. R. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*, 3–56.
Fama, E. F., & French, K. R. (2002). Testing tradeoff and pecking order predictions about dividends and debt. *Review of Financial Studies, 15*, 1–33.
Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. *Journal of Finance, 25*, 383–417.
Fama, E. F. (1991). Efficient markets II. *Journal of Finance, 46*, 1575–1617.
Fama, E. F., & MacBeth, J. D. (1973). Risk, return, and equilibrium: empirical tests. *Journal of Political Economy, 81*, 607–636.
Fama, E. F., & French, K. R. (1988). Dividend yields and expected stock returns. *Journal of Financial Economics, 22*, 3–25.
Fama, E. F., & French, K. R. (1989). Business conditions and expected returns on stocks and bonds. *Journal of Financial Economics, 25*, 23–49.
Fama, E. F., & French, K. R. (2001). Disappearing dividends: Changing firm characteristics or lower propensity to pay? *Journal of Financial Economics, 60*, 3–43.
Fama, E., & Babiak, H. (1968). Dividend policy: An empirical analysis. *Journal of the American Statistical Association, 63*, 1132–1161.
Fama, E., & Schwert, G. (1977). Asset returns and inflation. *Journal of Financial Economics, 5*, 115–146.
Fama, E., & French, K. R. (1992). The Cross-section of Expected Stock Return. *Journal of Finance, 47*, 427–465.
Fama, E., Fisher, L., Jensen, M., & Roll, R. (1969). The adjustment of stock prices to new information. *International Economic Review, 10*, 1–21.
Fama, E., & Jensen, M. (1983). Agency problems and residual claims. *Journal of Law and Economics, 26*, 327–349.
Fama, E., Fisher, K., Jensen, M., & Roll, R. (1969). The Adjustment of Stock Prices to New Information. *International Economic Review, 10*, 1–21.
Fama, E. F. (1970). Efficient Capital Markets: A review of theory and empirical work". *Journal of Finance, 25*(2), 383–417.
Fama, E. F. (1972). Components of investment performance. *Journal of Finance, 27*(3), 551–567.
Fama, E. F. (1991). Efficient Capital Markets: II". *Journal of Finance, 46*, 1575–1617.
Fama, E.F. (1998). "Market efficiency, long-term returns, and behavioral finance". Journal of Financial Economics" 49, 283–306.
Fama, E. F., & MacBeth, J. D. (1973). *Risk, return and equilibrium: empirical tests"* (pp. 607–636). May-June: Journal of Political Economy.
Fama, E. F., & French, K. (1993). Common risk factors in the returns on stocks and bonds. *Journal of Financial Economics, 33*(1), 3–56.
Fama, E. F., & French, K. R. (1995). Size and Book-to-Market factors in earnings and returns. *Journal of Finance, 50*, 131–55.
Fama, E. F., & French, K. R. (2008). Dissecting anomalies". *Journal of Finance, 63*(4), 1653–1678.
Fama, E. F., & French, K. R. (2010). Luck versus skill in the cross-section of mutual fund returns". *Journal of Finance, 65*(5), 1915–1947.
Fama, E. F. (1974). The empirical relationships between the dividend and investment decisions of firms. *The American Economic Review, 64*, 304–318.
Fama, E. F., & Gibbons, M. R. (1982). Inflation, real returns and capital investment. *Journal of Monetary Economics, 9*, 297–323.
Fama, E. F., & Schwert, W. G. (1977). Asset returns and Inflation. *Journal of Financial Economics, 4*, 115–146.
Fama, E. F., & French, K. R. (1992) The Cross-Section of Expected Stock Returns. *Journal of Finance, 47*, (2), 427–465, (June).

Fama, E. F., & French, K. R. (1989). Business Conditions and Expected Returns On Stocks and Bonds. *Journal of Financial Economics, 25*, 23–49.

Fama, E. F. (1965). The Behavior of Stock Prices. *Journal of Financial Economics, 47*, 244–280, (January).

Fama, E. F. (1973). A note on the market model and the two-parameter model. *Journal of Finance, 28*, 1181–1185.

Fama, E., & French, K. (1988). Dividend yields and expected stock returns. *Journal of Financial Economics, 22*, 3–26.

Fan, D., & Yeh, J. (2006). Analyst earnings forecasts for publicly traded insurance companies. *Review of Quantitative Finance and Accounting, 26*, 105–136.

Fan, J. (1997). Comments on 'Wavelets in Statistics: A Review' by A. Anto- niadis. *Journal of the Italian Statistical Association, 6*, 131–38.

Fan, J., & Li, R. (2006). Statistical challenges with high dimensionality: feature selection in knowledge discovery. In Proceedings of the International Congress of Mathematicians, European Mathematical Society, Zurich Vol. III, 595–622.

Fan, J., & Lv, J. (2008). Sure independence screening for ultrahigh dimensional feature space (with discussion). *Journal of the Royal Statistical Society, Series B, 70*, 849–911.

Fan, J., & Lv, J. (2010). A selective overview of variable selection in high dimensional feature space". *Statistica Sinica, 20*, 101–148.

Fan, J., & Yao, Q. (2003). *Nonlinear time series: Nonparametric and parametric methods*. Verlag: Springer.

Fan, J., & Song, R. (2010). Sure independence screening in generalized linear models with NPdimensionality. *The Annals of Statistics, 38*, 3567–3604.

Fan, J., & Wang, Y. (2007). Multi-scale jump and volatility analysis for high-frequency financial data. *Journal of the American Statistical Association, 102*(480), 1349–1362.

Fan, J., & Li, R. (2001). Variable selection via nonconcave penalized likelihood and it oracle properties. *Journal of the American Statistical Association, 96*, 1348–1360.

Fan, J., Feng, Y., & Song, R. (2011). "Nonparametric independence screening in sparse ultra-high dimensional additive models." *Journal of the American Statistical Association*, to appear.

Fan, J., Feng, Y., & Wu, Y. (2010). Ultrahigh dimensional variable selection for Cox's proportional hazards model. *IMS Collection, 6*, 70–86.

Fan, J., Hu, T.-C., & Truong, Y. K. (1994). Robust nonparametric function rstimation. *Scandinavian Journal of Statistics, 21*, 433–446.

Fan, J., Lv, J., & Qi, L. (2011). Sparse high-dimensional models in economics. *Annual Review of Economics., 3291–3317*.

Fan, J., Samworth, R., & Wu, Y. (2009). Ultrahigh dimensional feature selection: beyond the linear model. *Journal of Machine Learning Research, 10*, 1829–1853.

Fan, J., & Wong, T. J. (2002). Corporate Ownership Structure and the Informativeness of Accounting Earnings in East Asia.". *Journal of Accounting & Economics, 33*, 401–425.

Fan, J., & Wong, T. J. (2005). Do External Auditors Perform A Corporate Governance Role in Emerging Markets? *Journal of Accounting Research, 43*(1), 35–72.

Fan, J., Fan, Y., & Yazhen, W. (2007). Multiscale jump and volatility analysis for high-frequency financial data. *Journal of American Statistical Association, 102*, 618–631.

Fan, R., Gupta, A., & Ritchken, P. (2003). Hedging in the Possible Presence of Unspanned Stochastic Volatility: Evidence from Swaption Markets. *Journal of Finance, 58*, 2219–2248.

Fang, F., & Oosterlee, C. W. (2011). A Fourier-based valuation method for Bermudan and barrier options under Heston's model, SIAM J. *Financial Math, 2*, 439–463.

Fang, L. H. (2005). Investment bank reputation and the price and quality of underwriting services. *Journal of Finance, 60*, 2729–2761.

Faulkender, M., & Petersen, M. A. (2006). Does the source of capital affect capital structure? *Review of Financial Studies, 19*, 45–79.

Faure, H. (1982). Discrepance de Suites Associees a un Systeme de Numeration (en Dimension s). *Acta Arithmetica, 41*, 337–351.

Fazzari, S., Hubbard, R. G., & Petersen, B. (1988). Financing Constraints and Corporate Investment. *Brooking Papers on Economic Activity, 1*, 141–195.

Feenstra, R. C., & Knittel, C. R. (2004). Re-Assessing the U.S. Quality Adjustment to Computer Prices: The Role of Durability and Changing Software(October). NBER Working Paper No. W10857

Feltham, G., & Xie, J. (1994). Performance measure congruity and diversity in multi-task principal/agent relations. *The Accounting Review, 69*(3), 429–453.

Feltham, G. A., & Ohlson, J. A. (1995). Valuation and clean surplus accounting for operating and financial activities. *Contemporary Accounting Research, 11*(2), 689–731.

Feltham, G. A., & Ohlson, J. A. (1999). Residual earnings valuation with risk and stochastic interest rates. *Accounting Review, 74*(2), 165.

Feng, M., Li, C., & McVay, S. (2009). Internal control and management guidance. *Journal of Accounting and Economics, 48*, 190–209.

Fenn, G. W., & Liang, N. (2001). Corporate payout policy and managerial stock incentives. *Journal of Financial Economics, 60*, 45–72.

Fermanian, J. D., & Salanié, B. (2004). A Nonparametric Simulated Maximum Likelihood Estimation Method. *Econometric Theory, 20*, 701–734.

Fernandes, M., Mota, B., & Rocha, G. (2005). A multivariate conditional autoregressive range model. *Economics Letters, 86*, 435–440.

Fernández, V. (2011). Los fondos de pensiones como activo financiero familiar. Comparativa europea y perspectivas en España. *Ahorro familiar en España, 17*, 319–332.

Fernandez, V., & Lucey, B. (2007). Portfolio management under sudden changes in volatility and heterogeneous investment horizons. *Physica A: Statistical Mechanics and its Applications, 375*, 612–624.

Ferri, G., Liu, L. G., & Stiglitz, J. (1999). The Procyclical Role of Rating Agencies: Evidence from the East Asian Crisis. *Economic Notes, 3*, 335–355.

Ferruz, L., Gomez-Bezares, F., & Vargas, M. (2010). *"Portfolio theory, CAPM, and performance measures", in Handbook of Quantitative Finance and Risk Management* (pp. 267–283). C. F. Lee et al ed.: Springer.

Ferruz, L., Sarto, J. L., & Vicente, L. (2008a). Herding behavior in Spanish equity funds. *Applied economics letters, 15*(7), 573–576.

Ferruz, L., Sarto, J. L., & Vicente, L. (2008b). "Convergencia estratégica en la industria española de fondos de inversión." *El Trimestre Económico LXXV*(4), 1043–1060.

Ferson, W., & Harvey, C. R. (1991). The variation of economic risk premiums. *Journal of Political Economy, 99*, 285–315.

Ferson, W., & Harvey, C. R. (1993). The risk and predictability of international equity returns. *Review of Financial Studies, 6*, 527–566.

Ferson, W., & Korajczyk, R. A. (1995). Do arbitrage pricing models explain the predictability of stock returns?". *Journal of Business, 68*, 309–349.

Ferson, W. E., & Siegel, A. F. (2001). The efficient use of conditioning information in portfolios. *Journal of Finance, 56*, 967–982.

Ferson, W. E., & Siegel, A. F. (2003). Stochastic discount factor bounds with conditioning information. *Review of Financial Studies, 16*, 567–595.

Ferson, W. E., & Siegel, A. F. (2009). Testing portfolio efficiency with conditioning information. *Review of Financial Studies, 22*, 2735–2758.

Ferson, W. E., & Schadt, R. (1996). Measuring Fund Strategy and Performance in Changing Economic Conditions. *Journal of Finance, 51*(2), 425–461.

Ferson, W. E., Siegel, A. F., & Pisun, X. (2006). Mimicking portfolios with conditioning information. *Journal of Financial and Quantitative Analysis, 41*, 607–635.

Ferson, W. E., Sarkissian, S., & Simin, T. (2003). Spurious regressions in financial economics? *Journal of Finance, 58*, 1393–1414.

Ferson, W. (1995). Theory and Empirical Testing of Asset Pricing Models. In R. Jarrow, V. Maksimovic, & W. T. Ziemba (Eds.), *Handbooks in Operation Research and Management Science, vol 9*. Amsterdam: North-Holland.

Figlewski, S., & Gao, B. (1999). The Adaptive Mesh Model: a New Approach to Efficient Option Pricing. *Journal of Financial Economics, 53*, 313–351.

Figlewski, S., & Wachtel, P. (1983). Rational expectations, informational efficiency, and tests using survey data: A reply. *Review of Economics and Statistics, 65*, 529–531.

Figlewski, S. (1989). Option Arbitrage in Imperfect Markets, *Journal of Finance, 44*, 1289–1311.

Figlewski, S., Frydman, H., & Liang, W. (2012). Modeling the effects of macroeconomic factors on corporate default and credit rating transitions. *International Review of Economics and Finance, 21*(1), 87–105.

Figlewski, S. (1978). Market "efficiency" in a market with heterogeneous information. *Journal of Political Economy, 86*, 581–597.

Figlewski, S. (1982). Information diversity and market behavior. *Journal of Finance, 37*, 87–102.

Figlewski, S. (1984). Information diversity and market behavior: A reply. *Journal of Finance, 39*, 299–302.

Finger, C. (2000). A comparison of stochastic default rate models. *RiskMetrics Journal, 1*(2), 49–73.

Fishburn, P. C. (1977). Mean-risk analysis with risk associated with below-target returns. *American Economic Review, 67*, 116–126.

Fishburn, P. C. (1964). *Decision and Value Theory*. New York: John Wiley & Sons.

Fishburn, P. C. (1970). *Utility Theory for Decision Making*. New York: Wiley.

Fisher, J. (1992). Use of nonfinancial performance measures. *Journal of Cost Management, 31–38*.

Ratings, F. (2001). *Fitch Simplifies Bank Rating Scales*. Technical: Report.

Ratings, F. (2006). *The Role of Support and Joint Probability Analysis in Bank Ratings*. Fitch Special: Report.

Flannery, M. J., & Protopapadakis, A. A. (2002). Macroeconomic factors do influence aggregate stock returns. *Review of Financial Studies, 15*, 751–782.

Flannery, M. J., & Rangan, K. P. (2008). What caused the bank capital build-up of the 1990s?". *Review of Finance, 12*, 391–430.

Flannery, M. J., Hameed, A. S., & Harjes, R. H. (1997). Asset Pricing Time-varying Risk Premia and Interest Rate Risk". *Journal of Banking and Finance, 21*, 315–335.

Flannery, M. J., & James, C. M. (1984). The Effect of Interest Rate Changes on the Common Stock Returns of Financial Institutions. *Journal of Finance, 39*, 1141–1153.

Fledelius, P., Lando, D., & Nielsen, J. (2004). Nonparametric analysis of rating transition and default data. *Journal of Investment Management, 2*(2), 71–85.

Fleming, J., & Whaley, R. (1994). The value of wildcard options. *Journal of Finance, 49(1), pp. 215–36*(March).

Fligstein, N. (2001). *The architecture of markets: an economic sociology of twenty-first-century capitalist societies*. Princeton, N.J.: Princeton University Press.

Flood, R. P., & Hodrick, R. J. (1990). On testing for speculative bubbles". *Journal of Economic Perspectives, 4*, 85–102.

Flood, R., Hodrick, R., & Kaplan, P. (1986). An evaluation of recent evidence on stock market bubbles, NBER working paper

Floros, C. (2009). Modelling volatility using high, low, open and closing prices: evidence from four S&P indices. *International Research Journal of Finance and Economics, 28*, 198–206.

Fluck, Z., & Lynch, A. W. (1999). Why Do Firms Merge and Then Divest? A Theory of Financial Synergy. *Journal of Business, 72*, 319–346.

Fogler, H. R., John, K., & Tipton, J. (1981). Three factor interest rate differentials and stock groups. *Journal of Finance, 36*, 323–336.

Fok, R. C. W., Lee, C. F., & Cheng, D. C. (1996). Alternative specifications and estimation methods for determining random beta coefficients: Comparison and extensions. *Journal of Financial Studies, 4*(2), 61–88.

Follain, J. R., & Jimenez, E. (1985). Estimating the Demand for Housing Characteristics: A Survey and Critique". *Regional Science and Urban Economics, 15*(1), 77–107.

Follmer, H., & Leukert, P. (1999). Quantile hedging. *Finance and Stochastics, 3*, 251–273.

Follmer, H., & Leukert, P. (2000). Efficient hedging: cost versus shortfall risk. *Finance and Stochastics, 4*, 117–146.

Fong, G., & Vasicek, O. A. (1997). A Multidimensional Framework for Risk Analysis. *Financial Analysts Journal* (July-August), 51–57.

Fong, K., Gallagher, D. R., Gardner, P., & Swan, P. L. (2011). Follow the Leader: Fund Managers Trading in Signal-Strength Sequence. *Accounting & Finance, 51*(3), 684–710.

Fong, W. M., Lean, H. H., & Wong, W. K. (2008). Stochastic dominance and behavior towards risk: the market for internet stocks. *Journal of Economic Behavior and Organization, 68*, 194–208.

Fong, W. M., Wong, W. K., & Lean, H. H. (2005). International momentum strategies: a stochastic dominance approach. *Journal of Financial Markets, 8*, 89–109.

Fontes, D. B. M. M. (2008). Fixed versus Flexible Production Systems: A Real Option Analysis, European. *Journal of Operational Research, 188*, 169–184.

Forbes, S. W. (1971). Rates of Return in the Nonlife Insurance Industry. *Journal of Risk and Insurance, 38*, 409–422.

Forbes, S. W. (1972). The Cost of Capital of Insurance Companies: Comment. *Journal of Risk and Insurance, 39*(3), 491–492.

Forman, E. H., Saaty, T. L., Selly, M. A., & Waldron, R. (1985). *Expert choice*. Pittsburgh, PA: Decision Support Software Inc.

Fornari, F., & Mele, A. (1997). Sign- and volatility-switching ARCH models: Theory and applications to international stock markets. *Journal of Applied Econometrics, 12*, 49–65.

Forsythe, R., Palfrey, T., & Plott, C. (1982). Asset valuation in an experimental model". *Econometrica, 50*, 537–563.

Fortson, J. C. (1986). Factors affecting the discount rate for forestry investments. *Forest Products Journal, 36*, 67–72.

Foster, D. P., & George, E. I. (1994). The risk inflation criterion for multiple regression". *Annals of Statistics, 22*, 1947–1975.

Foster, F. J. (1995). Volume-volatility relationships for crude oil futures markets. *Journal of Futures Markets, 15*, 929–951.

Foster, F. D., & Viswanathan, S. (1993). Variations in Trading Volume. *Return Volatility, and Trading Costs: Evidence on Recent Price Formation Models, Journal of Finance, 48*, 187–211.

Foster, G. (1977). Valuation Parameters of Property-Liability Companies. *Journal of Finance, 32*(3), 823–835, (June).

Foster, I., & Kessleman, C. (2004). The Grid: Blueprint for a New Computing Infrastructure. (2nd edn.), Morgan Kaufmann

Foster, I., Zhao, Y., Raicu, I., & Lu, S. (2008). Cloud computing and grid computing 360-degree compared. Grid Computing Environments Workshop, 1-10.Galanti, S. and A. Jung. (1997). "Low-discrepancy sequences: Monte Carlo simulation of option prices", *Journal of Derivatives 5*(1), 63–83.

Fouque, J.-P., Papanicolaou, G., & Sircar, R. (2000). Derivatives in Financial Markets with Stochastic Volatility. Cambridge University Press

Frances, P. H., Marches, M., & Murray, A. (2005). A hybrid genetic-neural-architecture for stock index forecasting. *Information Science, 170*(1), 3–37.

Francioni, R., Hazarika, S., Schwartz, R., & Reck, M. (2010). *"Security Market Microstructure: the Analysis of a Non-Frictionless Market", Handbook of Quantitative Finance and Risk Management* (pp. 333–353). Editors, Springer: Cheng-Few Lee and Alice C. Lee.

Francis, J., Nanda, D., & Wang, X. (2006). Re-examining the effects of regulation fair disclosure using foreign listed firms to control for concurrent shocks. *Journal of Accounting and Economics, 41*, 271–292.

Francis, J., Lafond, R., Olsson, P., & Schipper, K. (2005). The Market Pricing of Accrual Quality. *Journal of Accounting & Economics, 39*, 295–327.

Francis, P., & Ghijsels, H. (1999). Additive outliers GARCH and forecasting volatility. *International Journal of Forecasting, 15*(1), 1–9.
Frank, I. E., & Friedman, J. H. (1993). A statistical view of some chemometrics regression tools". *Technometrics, 35*, 109–148.
Franke, J., Hardle, W., & Hafner, C. (2010). *Statistics of Financial Markets: an introduction*. New York, NY: Springer-Verlag.
Frankel, R., & Lee, C. (1998). *Accounting valuation, market expectation, and cross-*.
Frankel, J. A., & Froot, K. A. (1987). Using survey data to test standard propositions regarding exchange rate expectations. *American Economic Review, 77*, 133.
Frankel, J. A., & Froot, K. A. (1990). Chartists, fundamentalists, and trading in the foreign exchange market. *American Economic Review, 80*, 181–185.
Frankel, R., & Lee, C. M. C. (1998). Accounting valuation, market expectation, and cross-sectional stock returns. *Journal of Accounting and Economics, 25*(3), 283–319.
Frecka, T. J., & Lee, C. F. (1983). Generalized financial ratio adjustment processes and their implications. *Journal of Accounting Research, 21*, 308–316.
Freeman, R. (1989). Monthly Earnings Per Share-Time Series Models. *Omega, 17*, 37–44.
Frees, E. (2004). *Longitudinal and Panel Data Analysis and Applications in the Social Sciences,* New York, NY: Cambridge University Press.
French, K. R., Schwert, G. W., & Stambaugh, R. F. (1987). Expected stock returns and volatility. *Journal of Financial Economics, 19*, 13–29.
French, K. R., Ruback, R. C., & Schwert, G. W. (1983). Effects of Nominal Contracting Hypothesis on Stock Returns. *Journal of Political Economy, 91*, 70–96.
Frey, R., McNeil, A., & Nyfeler, M. (2001). Copulas and credit models. *RISK*, 111–113
Frey, R., & Sin, C. (1999). Bounds on European option prices under stochastic volatility". *Mathematical Finance, 9*, 97–116.
Friedman, B. M. (1992). Monetary economics and financial markets: retrospect and prospect:. *NBER Reporter, 1–5*.
Friedman, D. (1993). *How Trading Institutions Affect Financial Market Performance: Some Laboratory Evidence* (pp. 410–435). July: Economic Inquiry.
Friedman, J. (2008). Fast sparse regression and classification. Working paper, Stanford University
Friedman, M. (1957). *A theory of the consumption function*. Princeton, NJ: Princeton University Press.
Friend, I., & Blume, M. (1975). The Demand for Risky Assets. *American Economic Review, 65*, 900–927.
Friend, I., & Westerfield, R. (1980). Co-skewness and capital asset pricing. *Journal of Finance, 35*, 897–913.
Friend, I., Brown, F. E., Herman, E. S., & Vickers, D. (1962). *A Study of Mutual Funds*. Washington, D.C.: U.S. Government Printing Office.
Frieze A. M., Kannan. R., & Polson, N. (1194). Sampling from log-concave distributions. Ann App Probab4:812–837.
Frishling, V. (2002). A Discrete Question. *Risk, 15*, 115–116.
Fromlet, H. (2001). Behavioral Finance – Theory and Practical Application. *Business Economics, 36*, 63–69.
Froot, K. A., & Frankel, J. A. (1989). Forward discount bias: Is it an exchange risk premium? *Quarterly Journal of Economics, 104*, 139–161.
Froot, K. A., & Thaler, R. H. (1990). Anomalies foreign exchange. *Journal of Economic Perspectives, 4*(3), 179–192.
Froot, K. A., Scharfstein, D. S., & Stein, J. C. (1992). Herd on the street: Informational inefficiencies in a market with short-term speculation. *Journal of Finance, 47*, 1461–1484.
Froot, K. A., O'Connell, P. G. J., & Seasholes, M. S. (2001). The portfolio flows of international investors. *Journal of Financial Economics, 59*, 151–193.
Frost, P. A., & Savarino, J. E. (1988). For better performance: Constrain portfolio weights. *Journal of Portfolio Management, 15*, 29–34.
Fruhwirth, M., & Sogner, L. (2006). The Jarrow/Turnbull default risk model-evidence from the German market. *European Journal of Finance, 12*, 107–135.

Frydman, H., & Schuermann, T. (2008). Credit ratings dynamics and Markov mixture models. *Journal of Banking and Finance, 32*, 1062–1075.
Fuller, W. A. (1977). Some properties of a modification of the limited information estimator. *Econometrica, 45*, 939–953.
Fung, H. G., & Patterson, G. A. (1999). The dynamic relationship of volatility, volume, and market depth in currency futures markets. *Journal of International Financial Markets, Institutions and Money, 17*, 33–59.
Furfine, C., & Remolona, E. (2002). What's behind the liquidity spread? On-the-run and off-the-run US treasuries in autumn 1998. *BIS Quarterly Review, 51–58*.
Fusai, G., & Luciano, E. (2001). Dynamic value at risk under optimal and suboptimal portfolio policies. *European Journal of Operational Research, 135*, 249–269.
Gabaix, X., & Landier, A. (2008). Why has CEO pay increased so much? *Quarterly Journal of Economics, 123*, 49–100.
Gabudean, R. C. (2007). Strategic interaction and the co-determination of firms financial policies, Working Paper
Gagnon, L., Lypny, G. J., & McCurdy, T. H. (1998). Hedging foreign currency portfolios. *Journal of Empirical Finance, 5*, 197–220.
Gajendra Sidana. (2007). Classifying Mutual Funds in India: Some results from clustering, *Indian Journal of Economics and Business, II*(2).
Galai, D. (1983). The Components of the Return from Hedging Options against Stocks, *Journal of Business, 56*, 45–54.
Galai, D. (1983). *"A Survey of Empirical Tests of Option Pricing Models," in Option Pricing, ed: M* (pp. 45–80). Brenner, Lexington, MA: Heath.
Galindo, J., & Tamayo, P. (2000). Credit Risk Assessment Using Statistical and Machine Learning/Basic Methodology and Risk Modeling Applications. *Computational Economics, 15*, 107–143.
Gallant, A. R., & Tauchen, G. (1996). Which Moments to Match. *Econometric Theory, 12*, 657–681.
Gallant, A. R., & Tauchen, G. (1997). Estimation of Continuous Time Models for Stock Returns and Interest Rates. *Macroeconomic Dynamics, 1*, 135–168.
Gallant, A. R. (1987). *Nonlinear Statistical Models*. New York: John Wiley and Sons.
Gallant, A. R., & Tauchen, G. (1998). Reprojecting Partially Observed Systems with Applications to Interest Rate Diffusions. *Journal of American Statistical Association, 93*, 10–24.
Gallant, A. R., Hsu, C. T., & Tauchen, G. E. (1999). Using daily range data to calibrate volatility diffusions and extract the forward integrated variance. *Review of Economics and Statistics, 81*, 617–631.
Gallo, G. M., & Otranto, E. (2008). Volatility spillovers, interdependence and comovements: A Markov Switching approach. *Computational Statistics and Data Analysis, 52*, 3011–3026.
Gan, J. (2007). Financial Constraints and Corporate Investment: Evidence from an Exogenous Shock to Collateral. *Journal of Financial Economics, 85*, 709–734.
Gande, A., Puri, M., Saunders, A., & Walter, I. (1997). Bank underwriting of debt securities: Modern evidence. *Review of Financial Studies, 10*, 1175–1202.
Gande, A., Puri, M., & Saunders, A. (1999). Bank entry, competition, and the market for corporate securities underwriting. *Journal of Financial Economics, 54*, 165–195.
Gannon, G. L. (1994). Simultaneous Volatility Effects in Index Futures. *Review of Futures Markets, 13*, 1027–66.
Gannon, G. L. (2010). Simultaneous Volatility Transmissions and Spillovers: Theory and Evidence". *Review of Pacific Basin Financial Markets and Policies., 13*, 127–156.
Garbade, K. (1978). The Effect of Interdealer Brokerage on the Transactional Characteristics of Dealer Markets. *Journal of Business, 51*(3), 477–498.
Garber, P. (1990). Famous first bubbles". *Journal of Economic Perspectives, 4*, 35–54.
Garcia, P., Leuthold, M. R., & Zapata, H. (1986). Lead-lag relationships between trading volume and price variability: New evidence. *Journal of Futures Markets, 6*, 1–10.

Garcia, R., & Perron, P. (1996). An analysis of the real interest rate under regime shifts. *Review of Economics and Statistics, 78*, 111–125.

Gardner, D. C. (1996). Introduction to Mergers and Acquistions. Pitman Publishing London

Gargett, D. (1978). The link between stock prices and liquidity. *Financial Analysis Journal, 34*, 50–54.

Garmaise, M. J. (2008). Production in entrepreneurial firms: the effects of financial constraints on labor and capital. *Review of Financial Studies, 21*, 543–577.

Garman, M. (1976). Market Microstructure. *Journal of Financial Economics, 3*, 257–275.

Garman, M., & Klass, M. (1980). On the estimation of security price volatilities from historical data. *Journal of Business, 53*, 67–78.

Gasbarro, D., Wong, W. K., & Zumwalt, J. K. (2007). Stochastic dominance analysis of iShares. *European Journal of Finance, 13*, 89–101.

Gastaldi, T. (1993). A Kolmogorov-Smirnov Test Procedure Involving a Possibly Censored or Truncated Sample. *Communications in Statistics: Theory and Method, 22*, 31–39.

Gastwirth, J. L. (1971). A general definition of the Lorenz curve. *Econometrica, 39*, 1037–1039.

Gatev, E., Schuermann, T., & Strahan, P. E. (2009). Managing bank liquidity risk: How deposit-loan synergies vary with market conditions. *Review of Financial Studies, 22*, 995–1020.

Gatheral, J. (2006). The volatility surface. Wiely

Gaudenzi, M., & Zanette, A. (2009). Pricing American barrier options with discrete dividends by binomial trees. *Decisions in Economics and Finance, 32*, 129–148.

Gavin, W. T. (2003). FOMC forecasts: Is all the information in the central tendency? Economic Review, Federal Reserve Bank of St. Louis, 27–46.

Gay, G., & Manaster, S. (1984). The quality option implicit in futures contracts. *Journal of Financial Economics, 13(3), pp. 353–370*(September).

Gay, G., & Manaster, S. (1986). Implicit delivery options and optimal delivery strategies for financial futures contracts. *Journal of Financial Economics, 16(1), pp. 41–72*(May).

Ge, W., & McVay, S. (2005). On the disclosure of material weaknesses in internal control after the Sarbanes-Oxley act. *Accounting Horizon, 19*, 137–158.

Gebhardt, W., Lee, C., & Swaminathan, B. (2001). Toward an Implied Cost of Capital. *Journal of Accounting Research, 39*(1), 135–176.

Gebizlioglu, O. L., Şenoğlu, B., & Kantar, Y. M. (2011). Comparison of certain value-at-risk estimation methods for the two-parameter Weibull loss distribution". *Journal of Computational and Applied Mathematics, 235*, 3304–3314.

Gehrlein, W. V., & Lepelley, D. (1998). The Condorcet Efficiency of Approval Voting and the Probability of Electing the Condorcet Loser. *Journal of Mathematical Economics, 29*, 271–283.

Gelfand, A. E., & Smith, A. F. M. (1990). Sampling-based approaches to calculating marginal densities. *Journal of the American Statistical Association, 85*, 398–409.

Geman, H., Madan, D. B., & Yor, M. (2001). Times changes for Lévy processes. *Mathematical Finance, 11*, 79–96.

Gençay, R., & Gradojevic, N. (2011). Errors-in-variables estimation with wavelets. *Journal of Statistical Computation and Simulation, 81*, 1545–1564.

Gençay, R., & Fan, Y. (2010). Unit root tests with wavelets. *Econometric Theory, 26*, 1305–1331.

Gençay, R., Selçuk, F., & Ulugülyaˇgci, A. (2003). High volatility, thick tails and extreme value theory in value-at-risk estimation. *Insurance: Mathematics and Economics, 33*, 337–356.

Gençay, R., Selçuk, F., & Whitcher, B. (2002). *An introduction to wavelets and other filtering methods in finance and economics"*. San Diego: Academic Press.

Gençay, R., Selçuk, F., & Whitcher, B. (2003). Systematic risk and timescales. *Quantitative Finance, 3*, 108–116.

Gençay, R., Selçuk, F., & Whitcher, B. (2004). "Information Flow Between Volatilities Across Time Scales", Working Paper, Simon Fraser University

Gençay, R., Selçuk, F., & Whitcher, B. (2005). Multiscale systematic risk. *Journal of International Money and Finance, 24*, 55–70.

Genest, C., Ghoudi, K., & Rivest, L. (1995). A semiparametric estimation procedure of dependence parameters in multivariate families of distributions. *Biometrika, 82*, 543–52.

Gennotte, G., & Marsh, T. A. (1993). Variations in economic uncertainty and risk premiums on capital assets. *European Economic Review, 37*, 1021–1041.

Comer, G. (2006). 'Hybrid Mutual Funds and Market Timing Performance. *Journal of Business, 79*(2), 771–797.

George, T., & Longstaff, F. (1993). "Bid-Ask Spreads and Trading Activity in the S&P 100 Index Options Market. *Journal of Financial and Quantitative Analysis, 28*, 381–397.

Gervais, S., Heaton, J. B., & Odean, T. (2011). Overconfidence, Compensation Contracts, and Capital Budgeting. *Journal of Finance, 66*, 1735–1777.

Geske, R. (1979). A Note on an Analytical Valuation Formula for Unprotected American Call Options on Stocks with Known Dividends. *Journal of Financial Economics, 7*(4), 375–380.

Geske, R., & Roll, R. (1983). The fiscal and monetary linkage between stock returns and inflation. *Journal of Finance, 38*, 1–33.

Geweke, J. (1994a). Priors for macroeconomic time series and their application. *Econometric Theory, 10*, 609–632.

Geweke J. (1994b). Bayesian comparison of econometric models. Working Paper, Federal Reserve Bank of Minneapolis Research Department

Ghose, D., & Kroner, K. F. (1994). Common Persistence in Conditional Variances: Inplications for Optimal Hedging. Paper presented at the 1994 meeting of the Australasian Econometric Society.

Ghosh, A., & Lubberink, M. (2006). Timeliness and mandated disclosures on internal controls under section 404. Working Paper, City University of New York

Ghoshal, D., Canon, R. S., & Ramakrishnan, L. (2011). I/O performance of virtualized cloud environments. The Second International Workshop on Data Intensive Computing in the Clouds, 71–80.

Ghysel, E., Santa-Clara, P., & Valkanov, R. (2005). There is a risk-return trade-off after all. *Journal of Financial Economics, 76*, 509–548.

Gibbons, M. R., Ross, S. A., & Shanken, J. (1989). A test of the efficiency of a given portfolio. *Econometrica, 57*, 1121–1152.

Gijbels, I., Lambert, A., & Qiu, P. (2007). Jump-preserving regression and smoothing using local linear fitting: a compromise. *Annals of the Institute of Statistical Mathematics, 59*(2), 235–272.

Gilks, W. R., & Wild, P. (1992). Adaptive rejection sampling for Gibbs sampling. *Journal of the Royal Statistical Society*, Ser. C, *41*, 337–348.

Gillam, L., Ahmad, K., & Dear, G. (2005). Grid-enabling Social Scientists: The FINGRID infrastructure. Proceedings of the 1st International Conference on e-Social Science, Manchester, 22–24.

Gillette, A., Noe, T., & Rebello, M. (2003). Corporate board composition, protocols and voting behaviorL experimental evidence. *Journal of Finance, 58*, 1997–2032.

Gilroy, B. M., & Lukas, E. (2006). The Choice between Greenfield Investment and Cross-Border Acquisition: A Real Option Approach. *The Quarterly Review of Economics and Finance, 46*, 447–465.

Giot, P., & Laurent, S. (2003). Value-at-Risk for long and short trading positions. *Journal of Applied Econometrics, 18*, 641–664.

Giot, P., & Laurent, S. (2003). Market risk in commodity markets: a VaR approach. *Energy Economics, 25*, 435–457.

Giraitis, L., Robinson, P. M., & Surgailis, D. (1999). Variance-type estimation of long memory". *Stochastic Processes and their Applications, 80*, 1–24.

Givoly, D., & Lakonishok, J. (1979). The information content of financial analysts' forecasts of earnings: Some evidence on semi-strong efficiency. *Journal of Accounting and Economics, 1*, 165–185.

Glaser, M., & Weber, M. (2007). Overconfidence and trading volume. *The Geneva Risk and Insurance Review, 32*, 1–36.

Glasserman, P., & Kou, S. (2003). The Term Structure of Simple Forward Rates with Jump Risk. *Mathematical Finance, 13*(3), 383–410.
Glasserman, P. et al. (2000). "Efficient Monte Carlo Methods for Value-at-Risk." NSF Research Report, RCC21723 (97823).
Glasserman, P. (2003). Monte Carlo Methods in Financial Engineering, Springer Verlag
Glasserman, P., Heidelberger, P., & Shahabuddin, P. (2000). Efficient Monte Carlo Methods for Value-at-Risk. NSF Research Report RC 21723 (97823) 12.
Gleason, K. C., Mathur, I., & Peterson, M. A. (2004). Analysis of Intraday Herding Behavior among the Sector ETFs. *Journal of Empirical Finance, 11*, 681–694.
Gleason, K. C., Lee, C. I., & Mathur, I. (2003). Herding Behavior in European Futures Markets. *Finance Letters, 1*, 5–8.
Glosten, L. R., Jagannathan, R., & Runkle, D. E. (1993). On the relation between the expected value and the volatility of the nominal excess returns on stocks. *Journal of Finance, 48*, 1779–1801.
Glosten, L. R., & Harris, L. E. (1988). Estimating the Components of the Bid/Ask Spread. *Journal of Financial Economics, 21*, 123–142.
Gode, D., & Mohanram, P. (2003). Inferring the Cost of Capital Using the Ohlson-Juettner Model. *Review of Accounting Studies, 8*, 399–431.
Godfrey, L. G. (1978). Testing for higher order serial correlation in regression equations when the regressors include lagged dependent variables. *Econometrica, 46*, 1303–1310.
Goetzmann, W., & Peng, L. (2006). Estimating House Price Indexes in the Presence of Seller Reservation Prices. *Review of Economics and Statistics, 88*(1), 100–112.
Goicoechea, A., Hansen, D. R., & Duckstein, L. (1982). *Multi-objective decision analysis applications*. NewYork: John Wiley & Sons.
Goldberg, M. D., & Frydman, R. (1996). Imperfect knowledge and behaviour in the foreign exchange market. *Economic Journal, 106*, 869–893.
Goldberger, A. (1972a). Maximum-likelihood estimation of regressions containing unobservable independent variables. *International Economic Review, 13*, 1–15.
Goldberger, A. (1972b). Structural equation methods in the social sciences. *Econometrica, 40*, 979–1001.
Goldfeld, S. M., & Quandt, R. E. (1973). A Markov model for switching regressions. *Journal of Econometrics, 1*, 3–16.
Goldreich, D., Hanke, B., & Nath, P. (2005). The price of future liquidity: Time-varying liquidity in the US Treasury market. *Review of Finance, 9*, 1–32.
Goldstein, R. S. (2000). The Term Structure of Interest Rates as a Random Field. *Review of Financial Studies, 13*, 365–384.
Gollinger, T. L., & Morga, J. B. (1993). Calculation of an Efficient Frontier for a Commercial Loan Portfolio. *Journal of Portfolio Management, pages*, 39–46.
Gomes, J. (2001). Financing Investment. *American Economic Review, 91*, 1263–1285.
Gomez-Bezares, F., & Gomez-Bezares, F. R. (2006). On the use of hypothesis testing and other problems in financial research". *The Journal of Investing, 15*(4), 64–67.
Gomez-Bezares, F., Madariaga, J. A., & Santibáñez, J. (1996). Valuation and efficiency models: Does the CAPM beat the Market?". *Análisis Financiero, 68*, 72–96.
Gompers, P. A., & Metrick, A. (2001). Institutional Investors and Equity Prices. *Quarterly Journal of Economics, 116*(1), 229–259.
Gompers, P., & Lerner, J. (2000). Money chasing deals? the impact of fund inflows on private equity valuations. *Journal of Financial Economics, 55*, 281–325.
Goncalves, S., & White, H. (2002). The Bootstrap of the Mean for Dependent Heterogeneous Arrays. *Econometric Theory, 18*, 1367–1384.
Gonzalez, G., Reeves, R., & Franz, L. S. (1987). Capital budgeting decision making: an interactive multiple objective linear integer programming search procedure. *Advances in Mathematical Programming and Financial Planning, 1*, 21–44.
González, P., & García, J. C. (2002). Regulación del sector de seguros y planes de pensiones. *Revista ICE, Sistema Financiero: Novedades y Tendencias, 801*, 51–67.

Gonzalez-Rivera, G. (1998). Smooth transition GARCH models. *Studies in Nonlinear Dynamics and Econometrics, 3*, 61–78.
Goodfellow, C., Bohl, M. T., & Gebka, B. (2009). Together we invest? Individual and institutional investors' trading behavior in Poland. *International Review of Financial Analysis, 18*(4), 212–221.
Gopalakrishnan, S., Smith., M. Slott, J., & Murray, A. (2011). The value of disappearing beaches: A hedonic pricing model with endogenous beach width. *Journal of Environmental Economics and Management, 61*(3), 297–310.
Gordon, M. J. (1962). The savings investment and valuation of a corporation. *Review of Economics and Statistics, 44*, 37–51.
Gordon, M. J. (1967). Some Estimates of the Cost of Capital to the Electric Utility Industry, 1954–57: Comment. *American Economic Review*(1267–1268).
Gordon, M. J. (1974). The General Solution to the Buy or Lease Decision: A Pedagogical Note. *Journal of Finance, 29*(243–250).
Gordon, M. J. (1947a). The Cost of Capital to a Public Utility. East Lansing, Mich.: MSU Public Utilities Studies.
Gorton, G., & Pennacchi, G. (1992). Financial innovation and the provision of liquidity services. In J. R. Barth & R. Dan Brumbaugh (Eds.), *Reform of Federal Deposit Insurance*. New York: Harper Collins.
Goto, M., & Karolyi, G. A. (2004). Understanding Electricity Price Volatility within and Across Markets. Dice Centre Working Paper 2004-12, 1–41.
Goto, S., & Xu, Y. (2010). On mean variance portfolio optimization: improving performance through better use of hedging relations. Working Paper, University of Rhode Island.
Gotoh, J. Y., & Konno, H. (2002). Bounding option prices by semidefinite programming: a cutting plane algorithm". *Management Science, 48*, 665–678.
Gould, J. P. (1968). Adjustment costs in the theory of investment of the firm. *Review of Economic Studies, 35*, 47–55.
Gould, J. R. (1964). Internal pricing in corporations when there are costs of using an outside market. *The Journal of Business, 61–67*.
Gourieroux, C., & Monfort, A. (1997). *Time Series and Dynamic Models*. Cambridge, UK: Cambridge University Press.
Gow, O., & Taylor. (2010). Correcting for cross-sectional and time series dependence in accounting research. The Accounting Review 85(2), 483–512
Goyenko, R. Y., Holden, C. W., et al. (2009). Do liquidity measures measure liquidity? *Journal of Financial Economics, 92*(2), 153–181.
Grabowski, H. G., & Mueller, D. C. (1972). Managerial and stockholder welfare models of firm expenditures. *The Review of Economics and Statistics, 54*, 9–24.
Grammatikos, T., & Saunders, A. (1986). Futures price variability: a test of maturity and volume effects. *Journal of Business, 59*, 319–330.
Granger, C. W. J. (1989a). Combining forecasts twenty years later. *Journal of Forecasting, 8*, 167–173.
Granger, C. W. (1989b). Invited review: Combining forecasts-twenty years later. *Journal of Forecasting, 8*, 167–173.
Granger, C. (1992). Evaluating economic theory. *Journal of Econometrics, 51*, 3–5.
Granger, C. W. J., & Newbold, P. (1973). Some comments on the evaluation of economic forecasts. *Applied Economics, 5*, 35–47.
Granger, C. W. J., & Newbold, P. (1986). *"Forecasting economic time series", Second Edition*. San Diego: Academic Press.
Granger, C. W. J., & Joyeux, R. (1980). An introduction to long-memory time series and fractional differencing. *Journal of Time Series 1*, 15–29.
Granger, C. W. J., & Ramanathan, R. (1984). Improved methods of forecasting. *Journal of Forecasting, 3*, 197–204.
Granger, C. W. J., & Newbold, P. (1974). Spurious regressions in econometrics. *Journal of Econometrics, 2*, 111–120.

Granger, C. W. J. (1991). "Developments in the study of cointegrated economic variables." In R. F. Engle & C. W. J. Granger (Eds.), Long Run Economic Relationships, pp. 65–80. Oxford University Press

Granger, C. W. J. (1980). Long memory relationships and the aggregation of dynamic models. *Journal of Econometrics, 14*, 227–238.

Granger, C. W. J., & Ding, Z. (1996). Varieties of long memory models. *Journal of Econometrics, 73*, 61–77.

Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross spectral methods. *Econometrica, 37*, 424–438.

Grant, D. (1977). Portfolio performance and the cost of timing decisions. *Journal of Finance, 32*, 837–846.

Gray, S. F. (1996). Modeling the conditional distribution of interest rates as a regime-switching process. *Journal of Financial Economics, 42*, 27–62.

Green, P. E. (1988). Research for Marketing Decisions (Fifth Ed.), Prentice-Hall, Englewood Cliffs.

Green, P. J., & Silverman, B. W. (1994). *Nonparametric regression and generalized linear models: A roughness penalty approach*. Chapman &: Hall/CRC. ISBN 0412300400.

Green, R. (1984). Investment incentives, debt, and warrants. *Journal of Financial Economics, 13*, 115–36.

Greene, W. (2008). *Econometric Analysis* (6th ed.). Prentice Hall

Greene, W. (2008). *Econometric Analysis*. Pearson. Gruber

Greene, W. H. (2004). *Econometric Analysis* (4th ed.). New Jersey: Prentice Hall.

Greene, W.H. (2007). *Econometric Analysis*. Prentice Hall.

Greene, W. H. (2011). *Econometric Analysis* (11th ed.). New York: Macmillan Publishing Company.

Greene, W. H. (2012). *Econometric Analysis* (7th ed.). Prentice Hall.

Greene, W. H., & Analysis, E. (2003). *Prentice Hall*.

Greene, W. H., & Analysis, E. (1997). *3nd Edition*. New York: Macmillan Publishing Company.

Gregoriou, G. N., & Gueyie, J. P. (2003). Risk-adjusted performance of funds of hedge funds using a modified Sharpe ratio. *Journal of Wealth Management, 6*, 77–83.

Gregory, J., & Laurent, J. P. (2005). Basket Default Swaps, CDO's and Factor Copulas. *Journal of Risk, 7*(4), 103–22.

Grewal, M. S., Andrews, A. P., & Corporation, E. (2001). *Kalman Filtering: Theory and Practice Using MATLAB*. Wiley Online: Library. ISBN 0471392545.

Griffin, J. M., Harris, J. H., & Topaloglu, S. (2003). The dynamics of institutional and individual trading. *Journal of Finance, 58*, 2285–2320.

Griffiths, P., & Remenyi, D. (2003). Information Technology in Financial Services: A Model for Value Creation. *Electronic Journal of Information Systems Evaluation, 6*(2), 107–116.

Griliches, Z., & Hausman, J. A. (1986). Errors in Variables in Panel Data. *Journal of Econometrics, 31*, 93–118.

Grimshaw, A. S., & Wulf, W. A. (1997). The Legion vision of a worldwide virtual computer. *Communications of the ACM, 40*(1), 39–45.

Grinblatt, M., & Keloharju, M. (2000). The investment behavior and performance of various investor-types: A study of Finland's unique data set. *Journal of Financial Economics, 55*, 43–67.

Grinblatt, M., Titman, S., & Wermers, R. (1995). Momentum investment strategies, portfolio performance and herding: a study of mutual fund behavior. *American Economic Review, 85*, 1088–1105.

Grinold, R. C., & Kahn, R. N. (1992). *Active portfolio management* (2nd ed.). NY, NY: McGraw Hill.

Grinols, E. L. (1984). Production and Risk Leveling in the Intertemporal Capital Asset Pricing Model. *The Journal of Finance, 39*(5), 1571–1595.

Grootveld, H., & Hallerbach, W. (1999). Variance vs downside risk: Is there really that much difference?". *European Journal of Operational Research, 114*, 304–319.

Grossman, S., & Hart, O. (1982). Corporate financial structure and managerial incentives. In *The Economics of Information and Uncertainty*. Chicago: University of Chicago Press.

Grossman, S. J., & Shiller, R. J. (1981). The determinants of the variability of stock market prices. *American Economic Review, 71*, 222–227.

Grullon, G., & Michaely, R. (2002). Dividends, share repurchases and the substitution hypothesis. *Journal of Finance, 57*, 1649–1684.

Grullon, G., Michaely, R., & Swaminathan, B. (2002). Are dividend changes a sign of firm maturity? *Journal of Business, 75*, 387–424.

Grundy, B. (1991). Option prices and the underlying assets return distribution". *Journal of Finance, 46*, 1045–1069.

Grunert, J., Norden, L., & Weber, M. (2005). The role of non-financial factors in internal credit ratings. *Journal of Banking and Finance, 29*, 509–531.

Gu, F., & Wang, W. (2005). Intangible assets, information complexity, and analysts' earnings forecasts. *Journal of Business Finance and Accounting, 32*, 1673–1702.

Gu, K. (2002). The predictability of house prices. *Journal of Real Estate Research, 24*, 213–234.

Gu, Z., & Wu, J. S. (2003). Earnings skewness and analyst forecast bias. *Journal of Accounting and Economics, 35*(1), 5–29.

Gu, Z. (2003). Controlling for actual earnings: Does it mitigate or create spurious relations in analyst forecast efficiency? Working paper, Carnegie Mellon University

Gu, Z., & Wu, J. (2003). Earnings skewness and analyst forecast bias. *Journal of Accounting and Economics, 35*, 5–29.

Gu, Z., & Xue, J. (2007). Do analysts overreact to extreme good news in earnings?". *Review of Quantitative Finance and Accounting, 29*, 415–431.

Guay, W. (1999). The sensitivity of CEO wealth to equity risk: An analysis of the magnitude and determinants. *Journal of Financial Economics, 53*, 43–71.

Guay, W., & Harford, J. (2000). The cash-flow permanence and information content of dividend increases versus repurchases. *Journal of Financial Economics, 57*, 385–415.

Guidi, M., Russell, A., & Tarbert, H. (2007). The efficiency of international oil markets in incorporating US announcements during conflict and non-conflict periods". *Petroleum Accounting and Financial Management Journal, 26*(2), 67–86.

Guidolin, M., & Timmermann, A. (2005). "Economic implications of bull and bear regimes in UK stock and bond returns''. *Economic Journal, 115*, 111–143.

Guilbaud, O. (1998). Exact Kolmogorov-Type Test for Left-Truncated and/or Right-Censored Data. *Journal of American Statistical Association, 83*, 213–221.

Gujarati, D. N. (2009). *Basic Econometrics, 5th ed*. New York, NY: McGraw-Hill.

Gukhal, C. R. (2001). Analytical valuation of American options on jump-diffusion processes. *Mathematical Finance, 11*, 97–115.

Gultekin, N. (1983). Stock market returns and inflation: Evidence from other countries. *Journal of Finance, 38*, 49–65.

Gunny, K. (2010). The relation between earnings management using real activities manipulation and future performance: evidence from meeting earnings benchmarks. *Contemporary Accounting Research, 27*, 855–888.

Guo, H., & Wu, K. (2006). Liquidity premium analysis of Treasury bonds market in Shanghai Stock Exchange. *Finance and Economics (In Chinese), 217*, 23–29.

Guo, H., & Whitelaw, R. F. (2006). Uncovering the risk-return relation in the stock market. *Journal of Finance, 61*, 1433–1463.

Guo, H., Wu, G., & Zhijie Xiao. (2007). Estimating Value at Risk for Defaultable Bond Portfolios by Regression Quantile. *Journal of Risk Finance, 8*(2), 166–185.

Guo, J., & Hung, M. (2007). A note on the discontinuity problem in Heston's stochastic volatility model. *Applied Mathematical Finance, 14*(4), 339–345.

Gupta, O. P., & Gupta, A. (2004). *'Performance Evaluation of Select Indian Mutual Fund Schemes: An Empirical Study* (pp. 81–97). December: The ICFAI Journal of Applied Finance.

Gupta, A., & Subrahmanyam, M. (2001). An Examination of the Static and Dynamic Performance of Interest Rate Option Pricing Models in the Dollar Cap-Floor Markets, Working paper, Case Western Reserve University.

Gupta, M. C., & Walker, D. A. (1975). Dividend disbursal practices in commercial banking. *Journal of Financial & Quantitative Analysis, 10*, 515–529.

Gupta, M., Pevzner, M., & Seethamraju, C. (2010). The implications of absorption cost accounting and production decisions for future firm performance and valuation". *Contemporary Accounting Research, 27*, 889–922.

Gupton, G., Finger, C., & Bhatia, M. (1997). *CreditMetrics Technical Document*. New York, NY: Morgan Guaranty Trust Company.

Bakshi, G., Cao, C., & Chen, Z. (1997). Empirical performance of alternative option pricing models". *Journal of Finance, 52*, 2003–2049.

Gurkaynak, R. S. (2005). Econometric tests of asset price bubbles: Taking stock. REDS working paper No. 2005-04.

Gurun, U. G., Johnston, R., & Markov, S. (2011). Sell-side debt analysts and market efficiency, working paper.

Gustavson, S. G., & Lee, C. F. (1986). Risk-return Tradeoff, Income Measurement and Capital Asset Pricing for Life Insurers: An Empirical Investigation. *The Geneva Papers on Risk and Insurance, 11*(38), 23–43.

Guttler, A. (2011). Lead-lag Relationships and Rating Convergence among Credit Rating Agencies". *Journal of Credit Risk, 7*, 95–119.

Gwilym, O., MacMillan, D., & Speight, A. (1999). The intraday relationship between volume and volatility in LIFFE futures markets. *Applied Financial Economics, 9*, 593–604.

Gyetvan, F., & Shi, Y. (1992). Weak duality theorem and complementary slackness theorem for linear matrix programming problems. *Operations Research Letters, 11*, 244–252.

Haas, G. C. (1922). *Sales Prices as a Basis for Farm Land Appraisal, Technical Bulletin 9*. St. Paul: The University of Minnesota Agricultural Experiment Station.

Habib, A., & Hansen, J. (2008). Target shooting: Review of earnings management around earnings benchmarks. *Journal of Accounting Literature, 27*, 25–70.

Hachicha, A., & Chaabane, A. (2007). Monetary policy transmission mechanism in Tunisia. *Euro-Mediterranean Economics and Finance Review, 1*, 104–26.

Hadar, J., & Russell, W. (1969). Rules for ordering uncertain prospects. *The American Economic Review, 59*, 25–34.

Hadlock, C. (1998). Ownership, Liquidity, and Investment. *RAND Journal of Economics, 29*, 487–508.

Hafner, C. (2012). Cross-correlating wavelet coefficients with applications to high-frequency financial time series. *Journal of Applied Statistics, 39*, 1363–1379.

Hahn, J., Hausman, J., & Kuersteiner, G. (2004). Estimation with weak instruments: accuracy of higher-order bias and MSE Approximation. *Econometrics Journal, 7*, 272–306.

Hahn, W. F., & Mathews, K. H. (2007). Characteristics and hedonic pricing of differentiated beef demands. *Agricultural Economics, 36*(3), 377–393.

Hai, W., Mark, N., & Yangru, W. (1997). Understanding spot and forward exchange rate regressions. *Journal of Applied Econometrics, 12*, 715–734.

Haight, R. G., & Holmes, T. P. (1991). Stochastic price models and optimal tree cutting: Results for Loblolly pine. *Natural Resource Modeling, 5*, 424–443.

Hakansson, N., Beja, A., & Kale, J. (1985). On the Feasibility of Automated Market Making by a Programmed Specialist. *Journal of Finance*, 1–20.

Hakkio, C. S., & Rush, M. (1989). Market efficiency and cointegration: an application to the sterling deutschemark exchange markets. *Journal of International Money and Finance, 8*, 75–88.

Hall, B., & Murphy, K. (2003). The trouble with stock options. *Journal of Economic Perspectives, 17*(3), 49–70.

Hall, A. R., & Inoue, A. (2003). The Large Sample Behavior of the Generalized Method of Moments Estimator in Misspecified Models. *Journal of Econometrics*, 361–394.

Hall, B. (2003). Six challenges in designing equity-based pay. *Journal of Applied Corporate Finance, 15*, 21–33.

Hall, B., & Liebman, J. B. (1998). Are CEOs really paid like bureaucrats? Quarterly. *Journal of Economics, 113*, 653–691.

Hall, B. J., & Murphy, K. J. (2002). Stock options for undiversified executives. *Journal of Accounting and Economics, 33*, 3–42.

Hall, B., & Murphy, K. J. (2000). Optimal Exercise Prices for Executive Stock Options. *American Economic Review, 90*, 209–14.

Hall, P., & Miller, H. (2009). Using generalized correlation to effect variable selection in very high dimensional problems. *Journal of Computational and Graphical Statistics, 18*, 533–550.

Hall, P. (1992). *The Bootstrap and Edgeworth Expansion*. New York: Springer.

Hallerback, W. G. P. M. (2004). *An Improved Estimator for Black-Scholes-Merton Implied Volatility*. ERIM Report Series: Erasmus University.

Haltiwanger, J. C., & Waldman, M. (1989). Rational expectations in the aggregate. *Economic Inquiry, 27*, 619–636.

Halton, J. H. (1960). On the Efficiency of Certain Quasirandom Sequences of Points in Evaluating Multidimensional integrals. *Numerische Mathematik, 2*, 84–90.

Halverson, R., & Pollakowski, H. O. (1981). Choice of Functional Form Equations. *Journal of Urban Economics, 10*(1), 37–49.

Halvorsen, R., & Palmquist, R. (1980). The Interpretation of Dummy Variables in Semilogrithmic Regressions. *American Economic Review, 70*, 474–5.

Hamao, Y., & Mei, J. (2001). Living with the "enemy": An analysis of foreign investment in the Japanese equity market. *Journal of International Money and Finance, 20*, 715–735.

Hamburger, M., & Kochin, L. (1972). Money and stock prices: The channels of influence. *Journal of Finance, 27*, 231–249.

Hamiltion, J. (1994). *Time Series Analysis*. Princeton University: Press.

Hamilton, D., & Cantor, R. (2004). Rating transitions and default conditional on watchlist, outlook and rating history. Special Comment, Moody's Investor Service.

Hamilton, J. (1994). Time Series Analysis, Princeton

Hamilton, J. D. (1989). A new approach to the economic analysis of nonstationary time series and the business cycle. *Econometrica, 57*, 357–384.

Hamilton, J. D. (1990). Analysis of time series subject to changes in regimes. *Journal of Econometrics, 45*, 39–70.

Hamilton, J. D. (1994b). *Time Series Analysis*. New Jersey: Princeton University Press.

Hamilton, J. D., & Susmel, R. (1994a). A conditional heteroskedasticity and change in regime. *Journal of Econometrics, 64*, 307–333.

Hamilton, J. D., & Susmel, R. (1994b). Autoregressive conditional heteroskedasticity and changes in regime. *Journal of Econometrics, 64*, 307–333.

Hamilton, J. D. (1988). Rational-expectations econometric analysis of changes in regime: an investigation of the term structure of interest rates. *Journal of Economic Dynamics and Control, 12*, 385–423.

Hammer, P. L. (1986). *Partially Defined Boolean Functions and Cause-Effect Relationships. International Conference on Multi-Attribute Decision Making Via OR-Based Expert Systems*. Passau, Germany: University of Passau.

Hammer, P. L., Kogan, A., & Lejeune, M. A. (2006). Modeling Country Risk Ratings Using Partial Orders. *European Journal of Operational Research, 175*(2), 836–859.

Hammer, P. L., Kogan, A., & Lejeune, M. A. (2011). Reverse-Engineering Country Risk Ratings: Combinatorial Non-Recursive Model. *Annals of Operations Research, 188*, 185–213.

Hammer, P. L., Kogan, A., & Lejeune, M. A. (2012). A Logical Analysis of Banks' Financial Strength Ratings. *Expert Systems with Applications, 39*(9), 7808–7821.

Hamori, S., & Tokihisa, A. (1997). Testing for a unit root in the presence of a variance shift. *Economics Letters, 57*, 245–253.

Han, B., Manry, D., & Shaw, W. (2001). Improving the precision of analysts' earnings forecasts by adjusting for predictable bias. *Review of Quantitative Finance and Accounting, 17*, 81–98.

Han, B. (2007). Stochastic Volatilities and Correlations of Bond Yields. *Journal of Finance, 62*, 1491–1524.

Han, C.-H. (2012). Robust Robust Hedging Performance and Volatility Risk in Option Markets, Submitted

Han, C.-H., Liu, W.-H., & Chen, T.-Y. (2011). An Improved Procedure for VaR/CVaR Estimation under Stochastic Volatility Models, Submitted.

Hand, D., & Henley, W. (1997). Statistical classification methods in consumer credit scoring: A review, Series. *Journal of the Royal Statistical Society, 160*, 523–541.

Hanke, J., & Wichern, D. (2005). *Business Forecasting* (8th ed.). New York: Prentice Hall.

Hanley, K. W., Lee, C., & Seguin, P. (1996). The Marketing of Closed-end Fund IPOs. *Journal of Financial Intermediation, 5*, 127–159.

Hanlon, M., Rajgopal, S., & Shevlin, T. (2003). Are executive stock options associated with future earnings? *Journal of Accounting and Economics, 36*, 3–43.

Hanoch, G., & Levy, H. (1969). The efficiency analysis of choices involving risk. *Review of Economic Studies, 36*, 335–346.

Hansen, B. E. (1992). Tests for parameter instability in regressions with I(1) processes. *Journal of Business and Economic Statistics, 10*, 321–336.

Hansen, B. E. (1996). Inference when a Nuisance Parameter is Not Identified Under the Null Hypothesis. *Econometrica, 64*, 413–430.

Hansen, B. E. (1997). Approximate asymptotic p-values for structural change tests. *Journal of Business and Economic Statistics, 15*, 60–67.

Hansen, B. E. (1999). Threshold effects in non-dynamic panels: Estimation, testing, and inference. *Journal of Econometrics, 93*, 345–368.

Hansen, B. E. (2000). Sample splitting and threshold estimation. *Econometrica, 68*(3), 575–603.

Hansen, B. E. (2000). Testing for structural change in conditional models. *Journal of Econometrics, 97*, 93–115.

Hansen, B. E. (1994). Autoregressive Conditional Density Estimation. *International Economic Review, 35*(3), 705–730.

Hansen, C. B., McDonald, J. B., & Theodossiou, P. (2006). Some Flexible Parametric Models for Partially Adaptive Estimators of Econometric Models. Working Paper, Brigham Young University.

Hansen, L. P. (1982). Large sample properties of generalized method of moments estimators. *Econometrica, 50*(4), 1029–1054.

Hansen, L. P., & Hodrick, R. J. (1980). Forward exchange rates as optimal predictors of future spot rates: An econometric analysis. *Journal of Political Economy, 88*, 829–853.

Hansen, L. P., & Jagannathan, R. (1991). Implications of security market data for models of dynamic economies. *Journal of Political Economy, 99*, 225–262.

Hansen, L. P., & Richard, S. F. (1987). The role of conditioning information in deducing testable restrictions implied by dynamic asset pricing models. *Econometrica, 55*(3), 587–613.

Hansen, R. S., & Crutchley, C. (1990). Corporate earnings and financings: An empirical analysis. *Journal of Business, 63*, 347–371.

Hansen, R. S., & Chaplinsky, S. (1993). Partial anticipation, the flow of information and the economic impact of corporate debt sales. *Review of Financial Studies, 6*, 709–32.

Haque, N. U., Kumar, M. S., Mark, N., & Mathieson, D. (1996). The Economic Content of Indicators of Developing Country Creditworthiness. *International Monetary Fund Working Paper, 43*(4), 688–724.

Haque, N. U., Kumar, M. S., Mark, N., & Mathieson, D. (1998). The Relative Importance of Political and Economic Variables in Creditworthiness Ratings. *International Monetary Fund Working Paper, 46*, 1–13.

Härdle, W., & Song, S. (2010). Confidence bands in quantile regression. *Econometric Theory, 26*, 1180–1200.

Härdle, W., Liang, H., & Gao, J. (2000). *Partially Linear Models*. Heidelberg: Physica-Verlag.
Härdle, W., Müller, M., Sperlich, S., & Werwatz, A. (2004). *Nonparametric and Semiparametric Models*. Berlin: Springer-Verlag.
Härdle, W., Ritov, Y., & Song, S. (2011). *Partial linear quantile regression and bootstrap confidence bands*. Forthcoming: Journal of Multivariate Analysis.
Härdle, W., Spokoiny, V., & Wang, W. (2011). Local quantile regression, Discussion Paper 2011-005, CRC 649, Humboldt-Universit¨at zu Berlin. Submitted to Journal of Royal Statistical Society: Series B (Statistical Method) on Feb. 1, 2012.
Hardouvellis, G., Wizman, T., & La Porta, R. (1993). What Moves the Discount on Country Equity Funds? working paper 4571, NBER.
Hardy, C., & Reeve, R. (2000). A study of the internal control structure for electronic data interchange systems using the analytic hierarchy process. *Accounting and Finance, 40*, 191–210.
Hardy, G. H., Littlewood, J. E., & Polya, G. (1934). *Inequalities*. Cambridge, MA: Cambridge University Press.
Harker, P. T., & Vargas, L. G. (1987). The theory of ratio scale estimation: Saaty's analytic hierarchy process. *Management Science, 33*, 1383–1403.
Harlow, W. V. (1991). Asset allocation in a downside-risk framework. *Financial Analysts Journal, 47*, 28–40.
Harrington, S. E. (1983). The Relationship between Risk and Return: Evidence for Life Insurance Stocks. Journal for Risk and Insurance 50, 587–610 (December).
Harris, L. (2003). Trading and Exchanges: Market Microstructure for Practitioners, Oxford University Press.
Harris, R. D. F., & Yilmaz, F. (2010). Estimation of the conditional variance–covariance matrix of returns using the intraday range. *International Journal of Forecasting, 26*, 180–194.
Harris, R. D. F., Stoja, E., & Yilmaz, F. (2011). A cyclical model of exchange rate volatility. *Journal of Banking and Finance, 35*, 3055–3064.
Harris, T., Baldwin, S., & Siry, J. (2010). *United States Timberland Markets 2000 to mid-2010: Transactions*. Timber Mart-South, Athens, GA: Values & Market Research.
Harrison, M., & Kreps, D. (1978). Speculative investor behavior in a stock market with heterogeneous expectations. *Quarterly Journal of Economics, 92*, 323–336.
Harrison, M., & Kreps, D. (1979). Martingales and Arbitrage in Multiperiod Securities Markets. *Journal of Economic Theory, 20*, 381–408.
Harrison, P., & Zhang, H. (1999). An investigation of the risk and return relation at long horizons. *Review of Economics and Statistics, 81*, 399–408.
Hart, O. (1995). *Firms, contracts, and financial structure*. New York: Oxford University Press.
Hart, O. (2001). Financial contracting. *Journal of Economic Literature, 39*, 1079–1100.
Hart, O. D., & Kreps, D. M. (1986). Price destabilizing speculation. *Journal of Political Economy, 94*, 927–953.
Hartza, C., Mittnik, S., & Paolellad, M. (2006). Accurate value-at-risk forecasting based on the normal-GARCH model. *Computational Statistics & Data Analysis, 51*, 2295–2312.
Harvey, A. C., & Shephard, N. (1996). Estimation of an asymmetric stochastic volatility model for asset returns. *Journal of Business and Economic Statistics, 14*, 429–434.
Harvey, C. R., & Siddique, A. (1999). Autoregressive Conditional Skewness. *Journal of Financial and Quantitative Analysis, 34*, 465–487.
Harvey, A., & Koopman, S. J. (2002). Forecasting hourly electricity demand using time-varying splines. *Journal of the American Statistical Association, 88*, 424.
Harvey, A., & Trimbur, T. (2008). Trend estimation and the Hodrick-Prescott filter. *J. Japan Statist. Soc, 38*(1), 41–49.
Harvey, A. C. (1989). *Forecasting, Structural Time Series Models and the Kalman Filter*. Cambridge, UK: Cambridge University Press.
Harvey, A. C. (1993). *The Econometric Analysis of Time Series*. Cambridge, MA: The MIT Press.
Harvey, A., Ruiz, E., & Shephard, N. (1994). Multivariate stochastic variance models. *Review of Economic Studies, 61*, 247–264.

Harvey, A. C. (1998). Long-memory in stochastic volatility. In J. Knight & S. Satchell (Eds.), *Forecasting Volatility in Financial Markets*. London: Butterworth-Heinemann.
Harvey, C. (1991). The World Price of Covariance Risk. *Journal of Finance, 46*, 111–159.
Harvey, C. R. (1989c). Time-varying conditional covariances in tests of asset pricing models. *Journal of Financial Economics, 24*, 289–317.
Harvey, C. R. (2001a). The specification of conditional expectations. *Journal of Empirical Finance, 8*, 573–638.
Harvey, C. R. (2001b). The specification of conditional expectations. *Journal of Empirical Finance, 8*, 573–638.
Harvey, C. R., & Siddique, A. (2000). Conditional Skewness in Asset Pricing Tests. *Journal of Finance, 55*, 1263–1295.
Harvey, C. R., Liechty, J. C., Liechty, M. W., & Muller, P. (2003). Portfolio Selection with Higher Moments. Working Paper, Duke University.
Harvey, C. R., Lins, K. V., & Roper, A. H. (2004). The effect of capital structure when expected agency costs are extreme. *Journal of Financial Economics, 74*, 3–30.
Harvey, C., & Whaley, R. (1992a). Market Volatility and the Efficiency of the S&P 100 Index Option Market, *Journal of Financial Economics, 31*, 43–73.
Harvey, C., & Whaley, R. (1992b). Dividends and S&P 100 Index Option Valuation. *Journal of Futures Markets, 12*, 123–137.
Harvey, C. R. (1989). Time-varying conditional covariances in tests of asset pricing models. *Journal of Financial Economics, 24*, 289–317.
Hasbrouck, J. (2009). Trading costs and returns for US equities: Estimating effective costs from daily data. *Journal of Finance, 64*(3), 1445–1477.
Hasbrouck, J. (1991). Measuring the Information Content of Stock Trades. *Journal of Finance, 46*, 179–207.
Hasbrouck, J. (1993). Assessing the quality of a security market: a new approach to transaction-cost measurement. *Review of Financial Studies, 6*, 191–212.
Hastie T., Tibshirani R., & Friedman, J. (2009). The Elements of Statistical Learning: Data Mining, Inference, and Prediction (2nd. ed.), Springer-Verlag, New York.
Hastie, T. (1994). Principal curves and surfaces. DTIC Document.
Hastings, W. K. (1970). Monte Carlo sampling methods using Markov chains and their applications. *Biometrika, 57*, 97–109.
Haugen, R. A., Talmor, E., & Torous, W. N. (1991). The effect of volatility changes on the level of stock prices and subsequent expected returns. *Journal of Finance, 46*, 985–1007.
Haugen, R. A., & Kroncke, C. O. (1971). Rate Regulation and the Cost of Capital in the Insurance Industry. *Journal of Financial and Quantitative Analysis, 6*(5), 1283–1305.
Haugen, R. A., & Senbet, L. W. (1978). The insignificance of bankruptcy costs to the theory of optimal capital structure. *Journal of Finance, 33*, 383–93.
Haugen, R. A., & Senbet, L. W. (1981). Resolving the agency problems of external capital through options. *Journal of Finance, 26*, 629–647.
Haugen, R. A., & Senbet, L. W. (1988). Bankruptcy and agency costs: Their significance to the theory of optimal capital structure. *Journal of Financial and Quantitative Analysis, 23*, 27–38.
Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica, 46*, 1251–1271.
Hauswald, R., & Marquez, R. (2003). Information Technology and Financial Services Competition. *The Review of Financial Studies, 16*(3), 921–948.
Hautsch, N., Schaumburg, J., & Schienle, M. (2011). Financial network systemic risk contributions. Discussion Paper 2011-072, CRC 649, Humboldt-Universit¨at zu Berlin.
Haw, I., Hu, B., Hwang, L., & Wu, W. (2004). Ultimate ownership, income management, and legal and extra-legal institutions. *Journal of Accounting Research, 42*, 423–462.
Hawkes, A.G. (1971). Spectra of some self-exciting and mutually exciting point processes. Biometrika 58, 1, 83–90.

Hayashi, F. (1982). Tobin's Marginal q and Average q: A Neoclassical Interpretation. *Econometrica, 50*, 213–224.

Hayashi, F. (2000). *Econometrics*. Princeton, New Jersey: Princeton University Press.

He, A., & Wan, A. (2009). Predicting daily highs and lows of exchange rates: a cointegration analysis. *Journal of Applied Statistics, 36*, 1191–1204.

He, A., Kwok, J., & Wan, A. (2010). An empirical model of daily highs and lows of West Texas Intermediate crude oil prices. *Energy Economics, 32*, 1499–1506.

Healy, P., & Wahlen, J. (1999). A review of the earnings management literature and its implications for standard setting. *Accounting Horizons, 13*, 365–383.

Healy, P. M., & Palepu, K. G. (1988). Earnings information conveyed by dividend Initiations and omissions. *Journal of Financial Economics, 21*, 149–175.

Healy, P. M., Palepu, K. G., & Ruback, R. S. (1992). Does Corporate Performance Improved after Mergers? *Journal of Financial Economics*. 135–175.

Heath, D., & Jarrow, R. (1988). Ex-Dividend Stock Price Behavior and Arbitrage Opportunities. *Journal of Business, 61*, 95–108.

Heath, D., Jarrow, R., & Morton, A. (1992). Bond pricing and the term structure of interest rates: a new methodology for contingent claims valuation. *Econometrica, 60*, 77–105.

Heckman, J. J. (1979). Sample selection bias as a specification error. *Econometrica, 47*, 153–161.

Heckman, J., Matzkin, R., & Nesheim, L. (2010). Nonparametric Identification and Estimation of Nonadditive Hedonic Models. *Econometrica, 78*(5), 1569–1591.

Hedge, S. (1988). An empirical analysis of implicit delivery options in the Treasury bond futures contract. *Journal of Banking and Finance, 12*(2), 469–492.

Hedge, S. (1990). An ex post valuation of the quality option in Treasury bond futures contract. *Journal of Banking and Finance, 14*(4), 741–760.

Heffernan, J. E., & Tawn, J. A. (2004). A conditional approach for multivariate extreme values. *Journal of the Royal Statistical Society B, 66*, 497–546.

Heidari, M., & Wu, L. (2003). Are Interest Rate Derivatives Spanned by the Term Structure of Interest Rates? *Journal of Fixed Income, 13*, 75–86.

Heikkinen, V. P. (2002). Co-integration of timber and financial markets-Implications for portfolio selection. *Forest Science, 48*, 118–128.

Heinemann, F. (2003). Risk Aversion and Incentive Effects: Comment, Working paper (Ludwig-Maximilians-Universitat Munchen). Hirota, and Sunder, 2005, Price Bubbles sans Dividend Anchors: Evidence from Laboratory Stock Markets, working paper.

Hek, P. A. (1999). Endogenous Technological Change with Leisure-Dependent Utility. *Economic Theory, 14*, 669–84.

Helder, J., & Jin, H. (2007). Long memory in commodity futures volatility: a wavelet perspective. *The Journal of Futures Markets, 27*, 411–437.

Hemler, M. (1990). The quality delivery option in Treasury bond futures contracts. *Journal of Finance, 45*(5), 1565–1586.

Hemmer, T., Kim, O., & Verrechia, R. K. (2000). Introducing convexity into optimal compensation contracts. *Journal of Accounting and Economics, 28*, 307–327.

Hemmer, T., Matsunaga, S., & Shevlin, T. (1996). The influence of risk diversification on the early exercise of employee stock options by executive officers. *Journal of Accounting and Economics, 21*, 45–68.

Hendricks, D. (1996). Evaluation of value at risk models using historical data. *Federal Reserve Bank of New York Economic Policy Review, 2*, 39–69.

Hendricks, K. B., & Singhal, V. R. (2001). The long-run stock price performance of firms with effective TQM programs. *Management Science, 47*, 359–368.

Henker, J., Henker, T., & Mitsios, A. (2006). Do Investors Herd Intraday in Australian Equities? *International Journal of Managerial Finance, 2*(3), 196.

Hennessy, C. (2004). Tobin's Q, Debt Overhang, and Investment. *Journal of Finance, 59*, 1717–1742.

Henriksson, R. D. (1984). Market timing and mutual fund performance: An empirical investigation. *Journal of Business, 57*(1), 73–96.
Hermalin, B., & Weisbach, M. (1988). The determinants of board composition. *Rand Journal of Economics, 19*, 589–606.
Hermalin, B. E., & Weisbach, M. S. (2003). Boards of directors as an endogenously determined institution: a survey of the economic literature. *Federal Reserve Bank of New York Economic Policy Review, 9*, 7–26.
Heston, S. L. (1993). A closed-form solution for options with stochastic volatility with applications to bond and currency options. *Review of Financial Studies, 6*(2), 327–343.
Heung, C. J., & McDonald, J. (2005). Forecasting Asymmetries in Aggregate Stock Market Returns: Evidence from Conditional Skewness. *Journal of Empirical Finance, 12*, 666–685.
Hicks, A. (1994). Introduction to Pooling, Cambridge University Press.
Hidalgo, J., & Robinson, P. M. (1996). Testing for structural change in a longmemory environment. *Journal of Econometrics, 70*, 159–174.
Hietala, P. T. (1989). Asset Pricing in Partially Segmented Markets: Evidence from the Finnish Market. *Journal of Finance, 44*(3), 697–718.
Higgins, R. C. (1977). How much growth can a firm afford? *Financial Management, 6*, 7–16.
Higgins, R. C. (1972). The corporate dividend-saving decision. *Journal of Financial and Quantitative Analysis, 7*, 1527–1541.
Higgins, R. C., & Schall, L. D. (1975). Corporate Bankruptcy and Conglomerate Merger. *The Journal of Finance, 30*, 93–119.
Hill, P., Brooks, R., & Faff, R. (2010). Variations in sovereign credit quality assessments across rating agencies. *Journal of Banking and Finance, 34*(6), 1324–1343.
Hill, R. D., & Modigliani, F. (1987). The Massachusetts Model of Profit Regulation in Nonlife Insurance: An Appraisal and Extension. In J. D. Cummins & S. E. Harrington (Eds.), *Fair Rate of Return in Property-Liability Insurance*. Norwell, MA: Kluwer Academic Publishers.
Hill, R. D. (1979). Profit Regulation in Property-Liability Insurance. Bell Journal of Economics 10 (No. 1), 172–191, (Spring).
Hillier, F. S. (1963). The derivation of probabilities information for the evaluation of risky investments. Management Science, 443–457.
Hirschey, M., & Richardson, V. (2004). Are scientific indicators of patent quality useful to investors? *Journal of Empirical Finance, 11*, 91–107.
Hirshleifer, D., Subrahmanyam, A., & Titman, S. (1994). Security Analysis and Trading Patterns When Some Investors Receive Information before Others. *Journal of Finance, 49*(5), 1665–1698.
Hirshleifer, J. (1956). On the economies of transfer pricing. The Journal of Business, 172–184.
Hirst, E., Koonce, L., & Venkataraman, S. (2008). Management Earnings Forecasts: A Review and Framework. *Accounting Horizons, 22*, 315–338.
Hite, G., & Warga, A. (1997). The effects of bond rating changes on bond price performance. *Financial Analysts Journal*, 35–51.
Ho, L. C. J., & Tsay, J. (2004). Analysts' forecasts of Taiwanese firms' earnings: Some empirical evidence. *Review of Pacific Basin Financial Markets and Policies, 7*(4), 571–597.
Ho, T., Lee, S., & Son, K. (1992). Term structure estimation and pricing of callable Treasury bonds. *Review of Quantitative Finance and Accounting, 2*(2), 127–143.
Ho, J. (2004). Analysts' forecasts of Taiwanese firms' earnings: Some empirical evidence. *Review of Pacific Basin Financial Markets and Policies, 7*, 571–597.
Ho, T. S. Y., & Marcis, R. (1984). Dealer Bid-Ask Quotes and Transaction Prices: An Empirical Study of Some AMAX Options. *Journal of Finance, 39*, 23–45.
Ho, T. S. Y., & Lee, S. B. (2004a). *The Oxford Guide to Financial Modeling*. New York: Oxford University Press.
Ho, T. S. Y., & Lee, S. B. (2004b). Valuing High Yield Bonds: A Business Modeling Approach. *Journal of Investment Management, 2*, 1–12.
Ho, T., Schwartz, R., & Whitcomb, D. (1985). The Trading Decision and Market Clearing Under Transaction Price Uncertainty. *Journal of Finance*, 21–42.

Hobson, D., Laurence, P., & Wang, T. H. (2005). Static-arbitrage upper bounds for the prices of basket options. *Quantitative Finance, 5*, 329–342.

Hocking, R. R. (1976). The Analysis and Selection of Variables in Linear Regression. *Biometrics, 32*, 1–49.

Hodge, F. D., Rajgopal, S., & Shevlin, T. (2009). Do managers value stock options and restricted stock consistent with economic theory? *Contemporary Accounting Research, 26*, 899–932.

Hodrick, R., & Prescott, E. (1997). Postwar U.S. business cycles: an empirical investigation. *Journal of Money, Credit, and Banking, 29*, 1–16.

Hodrick, R. (1992). Dividend Yields and expected stock returns: alternative procedures for inference and measurement. *Review of Financial Studies, 5*, 357–386.

Hoerl, A., & Kennard, R. (1970). Ridge regression: biased estimation for nonorthogonal problems. *Technometrics, 12*, 55–67.

Hofstede, G. (2001). *Culture's consequences*. London: Sage Publications.

Hogg, R. V., & Klugman, S. A. (1983). On the Estimation of Long Tailed Skewed Distributions with Actuarial Applications. *Journal of Econometrics, 23*, 91–102.

Hogg, R. V., & Craig, A. T. (2004). *Introduction to mathematical statistics* (6th ed.). Englewood Cliffs, NJ: Prentice Hall.

Holden, C. W. (2009). New low-frequency spread measures. *Journal of Financial Markets, 12*(4), 778–813.

Holden, K., & Peel, D. A. (1990). On testing for unbiasedness and efficiency of forecasts. *Manchester School, 58*, 120–127.

Holland, L. C., & Elder, E. M. (2006). Employee stock options in compensation agreements: a financing explanation. *Journal of Corporate Finance, 12*, 367–379.

Hollander, M., & Wolfe, D. A. (1999). *Nonparametric statistical methods*. John Wiley & Sons.

Holmstrom, B., & Milgrom, P. (1991). Multi-task principal–agent analyses: incentive contracts, asset ownership, and job design. *Journal of Law, Economics, and Organization, 7*, 24–52.

Holmstrom, B. (1979). Moral Hazard and Observability. *Bell Journal of Economics, 10*, 74–91.

Holt, C. A., & Laury, S. K. (2002). Risk Aversion and Incentive Effects. *American Economic Review, 92*, 1644–1655.

Holthausen, R., & Watts, R. (2001). The relevance of the value relevance literature for financial accounting setting. *Journal of Accounting and Economics, 31*(September), 3–75.

Holthausen, R. W., & Leftwich, R. W. (1986). The effect of bond rating changes on common stock prices. *Journal of Financial Economics, 17*, 57–89.

Holtz-Eakin, D., Newey, W., & Rosen, H. S. (1988). Estimating Vector Autoregressions with Panel Data. *Econometrica, 56*, 1371–1395.

Homa, K., & Jaffee, D. (1971). The supply of money and common stock prices. *Journal of Finance, 26*, 1045–1066.

Hong, H., Kubik, J., & Salomon, A. (2000). Security analysts' career concerns and herding of earnings forecasts. *RAND Journal of Economics, 31*, 121–144.

Hong, H. G., Scheinkman, J. A., & Xiong, W. (2006). Asset Float and Speculative Bubbles. *Journal of Finance, 61*, 1073–1117.

Hong, Y. (2001). Evaluation of out-of-sample Probability Density Forecasts with Applications to S&P 500 Stock Prices. Working Paper, Cornell University.

Hong, Y. M., & Li, H. T. (2005). Nonparametric Specification Testing for Continuous Time Models with Applications to Term Structure Interest Rates. *Review of Financial Studies, 18*, 37–84.

Hong, Y. M., Li, H. T., & Zhao, F. (2007). Can the random walk model be beaten in out-of-sample density forecasts? Evidence from intraday foreign exchange rates. *Journal of Econometrics, 141*, 736–776.

Hong, Y., Lin, H., & Wang, S. (2010). Modeling the dynamics of Chinese spot interest rates. *Journal of Banking and Finance, 34*, 1047–1061.

Hoppe, R. (1999). It's time we buried Value-at-Risk. Risk Professional, 1(July/August).

Horvath, L., & Kokoszka, P. S. (2003). A bootstrap approximation to a unit root test statistic for heavy tailed observations. *Statistics and Probability Letters, 62*, 163–173.

Hoshi, T., Kashyap, A., & Scharfstein, D. (1991). Corporate Structure, Liquidity, and Investment: Evidence from Japanese Industrial Groups. *Quarterly Journal of Economics, 106*, 33–60.

Hoshino, T., & Kuriyama, K. (2010). Measuring the Benefits of Neighbourhood Park Amenities: Application and Comparison of Spatial Hedonic Approaches. *Environmental and Resource Economics, 45*(3), 429–444.

Hosking, J. R. M. (1981). Fractional differencing. Biometrika 68, 165–176.

Hosking, J. R. M. (1994). The Four-parameter Kappa Distribution. *IBM Journal of Research and Development, 38*, 251–258.

Hosmer, D., Lemeshow, S., & May, S. (2008). *Applied survival analysi*s (2nd ed.). Wiley Series in Probability and Statistics.

Hotchkiss, E. S., & Ronen, T. (2002). The informational efficiency of the corporate bond market: an intraday analysis. *Review of Financial Studies, 15*, 1325–1354.

Hovakimian, A., & Kane, E. J. (2000). Effectiveness of capital regulation at U.S. commercial banks: 1985 to 1994. *Journal of Finance, 55*, 451–468.

Howard, C. T., & D'Antonio, L. J. (1984). A risk-return measure of hedging effectiveness. *Journal of Financial and Quantitative Analysis, 19*, 101–112.

Howe, K. M., & Patterson, I. H. (1985). Capital investment decisions under economies of scale in flotation costs. *Financial Management, 14*, 61–69.

Howe, K. M., He, J., & Kao, G. W. (1992). One-time cash flow cash and free cash-flow Theory: Share Repurchases and special dividends. *Journal of Finance, 47*, 1963–1975.

Howell, R. A., & Schwartz, W. A. (1994). *Asset deployment and investment justification, in Handbook of cost management*. Boston, MA: Warren Gorham Lamont.

Hribar, P., & Nichols, C. (2007). The use of unsigned earnings quality measures in tests of earnings management. *Journal of Accounting Research, 45*, 1017–1053.

Hsiao, C., & Robinson, P. M. (1978). Efficient Estimation for a Dynamic Error-Shock Model. *International Economic Review, 19*, 467–480.

Hsiao, C. (2003). Analysis of Panel Data. 2nd Edition, Cambridge University Press.

Hsieh, D. A. (1991). Chaos and nonlinear dynamics: application to financial markets. *Journal of Finance, 46*, 1839–1877.

Hsieh, D. A. (1993). Implications of nonlinear dynamics for financial risk management. *Journal of Financial and Quantitative Analysis, 28*, 41–64.

Hsieh, J., & Walkling, R. A. (2005). Determinants and implications of arbitrage holdings in acquisitions. *Journal of Financial Economics, 77*, 605–648.

Hsieh, S., Yang, T.-Y., & Tai, Y. (2010). Positive Trading Effects and Herding Behavior in Asian Markets: Evidence from Mutual Funds. *The International Journal of Business and Finance Research, 4*(2), 177–188.

Hsu, D., & Chiao, C. H. (2011). Relative accuracy of analysts' earnings forecasts over time: a Markov chain analysis. *Review of Quantitative Finance and Accounting, 37*(4), 477–507.

Hu, Y.-T., Kiesel, R., & Perraudin, W. (2002). The Estimation of Transition Matrices for Sovereign Credit Ratings. *Journal of Banking and Finance, 26*(7), 1383–1406.

Hu, L. (2006). Dependence patterns across financial markets: a mixed copula approach. *Applied Financial Economics, 16*, 717–729.

Hu, N., Liu, L., Tripathy, A., & Yao, L. J. (2011). Value relevance of blog visibility. *Journal of Business Research, 64*, 1361–1368.

Hua, J. (2010). Option implied volatilities and corporate bond yields: a dynamic factor approach. Working Paper, Baruch College, City University of New York.

Huang, Z., Chen, H., Hsu, C.-J., Chen, W.-H., & Wu, S. (2004). Credit Rating Analysis with Support Vector Machines and Neural Networks: A Market Comparative Study. *Decision Support Systems, 37*, 543–558.

Huang, C. Y., & Lin, J. B. (2004a). Value-at-risk analysis for Taiwan stock index futures: fat tails and conditional asymmetries in return innovations. *Review of Quantitative Finance and Accounting, 22*, 79–95.

Huang, Y. C., & Lin, B. J. (2004b). Value-at-Risk analysis for Taiwan stock index futures: fat tails and conditional asymmetries in return innovations. *Review of Quantitative Finance and Accounting, 22*, 79–95.

Huang, C.-H., & Hsu, Y. Y. (1999). The impact of foreign investors on the Taiwan stock exchange. *Taipei Bank Monthly, 29*(4), 58–77.

Huang, H., Willis, R., & Zang, A. (2005). Bold security analysts' earnings forecasts and managers' Information Flow. Working Paper, Duke University.

Huang, J. (2004). Option pricing bounds and the elasticity of the pricing kernel". *Review of Derivatives Research, 7*, 25–51.

Huang, J., Ma, S. G., & Zhang, C. H. (2008). Adaptive Lasso for sparse high-dimensional regression models. *Statistica Sinica, 18*, 1603–1618.

Huang, J.-Z., & Shi, Z. (2010). Determinants of bond risk premia. American Finance Association 2011. Denver Meetings Paper. Univeristy Penn State.

Huang, R. D., & Stoll, H. R. (1997). The components of the bid-ask spread: A general approach. *Review of Financial Studies, 10*(4), 995–1034.

Huang, S., Hung, W., Yen, D. C., Chang, I., & Jiang, D. (2011a). Building the evaluation model of the IT general control for CPAs under enterprise risk management. *Decision Support Systems, 50*, 692.

Huang, X. (2008). Mean-variance model for fuzzy capital budgeting. Computers & Industrial Engineering 55, 34.

Huang, X., Zhou, H., & Zhu, H. (2011). Systemic risk contributions, Staff working papers 2011-08, The Federal Reserve Board.

Hubbard, R. G. (1998). Capital Market Imperfections and Investment. *Journal of Economic Literature, 36*, 193–227.

Huddart, S., & Lang, M. (1996). Employee stock option exercises: an empirical analysis. *Journal of Accounting and Economics, 21*, 5–43.

Hueng, C. J., & Brooks, R. (2002). Forecasting Asymmetries in Aggregate Stock Market Returns: Evidence from Higher Moments and Conditional Densities. Working Paper, WP02-08-01, Western Michigan University.

Hughes, J. P., Lang, W. W., Mester, L. J., Moon, C., & Pagano, M. (2003). Do bankers sacrifice value to build empires? Managerial incentives, industry consolidation, and financial performance. *Journal of Banking and Finance, 27*, 417–447.

Hughes, J. P., Mester, L. J., & Moon, C. (2001). Are scale economies in banking elusive or illusive? Incorporating capital structure and risk into models of bank production. *Journal of Banking and Finance, 25*, 2169–2208.

Huisman, R., & Huurman, C. (2003). Fat Tails in Power Prices. ERIM Report Series Reference No. ERS-2003-059-F&A. Available at SSRN: http://ssrn.com/abstract=450992

Hull, J. (2000). *Options, Futures and Other Derivative Securities* (4th ed.). New Jersey: Prentice-Hall.

Hull, J. (2003). Options, Futures, and Other Derivatives. Prentice Hall.

Hull, J. (2008). Options, Futures, and Other Derivatives. 7th edition: Prentice Hall.

Hull, J., & White, A. (1987a). The Pricing of Options on Assets with Stochastic Volatilities. *Journal of Finance, 42*, 281–300.

Hull, J., & White, A. (2000). Forward Rate Volatilities, Swap Rate Volatilities, and the Implementation of the LIBOR Market Models. *Journal of Fixed Income, 10*, 46–62.

Hull, J., & White, A. (2001). Valuing Credit Default Swaps II: Modeling Default Correlations. *Journal of Derivatives, 8*(3), 12–22.

Hull, J., & White, A. (1990). Pricing interest-rate-derivative securities. *Review of Financial Studies, 3*, 573–592.

Hull, J., & White, A. (2004a). How to value employee stock options. *Financial Analysts Journal, 60*, 114–119.

Hull, J., & White, A. (2004b). Valuation of a CDO and an nth to Default CDS without Monte Carlo Simulation. *Journal of Derivatives, 12*(2), 8–23.
Hull, J. C. (2011). *Options, Futures, and Other Derivatives* (8th ed.). Englewood Cliffs, NJ: Prentice-Hall.
Hull, J. C. (2003). *Options, Futures and other Derivatives* (5th ed.). Upper Saddle River, New Jersey: Prentice Hall.
Hull, J., & White, A. (1987b). The Pricing of Options with Stochastic Volatilities, *Journal of Finance, 42*, 281–300.
Hull, J., & White, A. (1987c). Hedging the Risks from Writing Foreign Currency Options, *Journal of International Money and Finance, 6*, 131–152.
Hull, J. C. (2003). *Options, Futures, and Other Derivatives* (5th ed.). Prentice Hall, pp. 287–288.
Hull, J., & White, A. (1998b). Incorporating Volatility Updating into the Historical Simulation Method for Value-at- Risk. *Journal of Risk, 1*(Fall), 5–19.
Hulton, C., & Wycoff, F. (1981). The Measurement Of Economic Depretiation, In: HULTEN, Charles (Ed.). Depretiation, Inflation, And The Taxation Of Income From Capital, Washington: Urban Institute Book.
Hurvich, C. M., Simonoff, J. S., & Tsai, C. L. (1998). Smoothing parameter selection in nonparametric regression using an improved Akaike information criterion. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 60*(2), 271–293.
Huston, T. L., Ruggiero, M., Conner, R., & Geis, G. (1981). Bystander intervention into crime: A study based on naturally-occurring episodes. *Social Psychology Quarterly, 44*, 14–23.
Hutchinson, J. M., Lo, A. W., & Poggio, T. (1994). A nonparametric approach to pricing and hedging derivative securities via learning networks. *Journal of Finance, 49*(3), 851–889.
Hwang, S., & Salmon, M. (2004). Market stress and herding. *Journal of Empirical Finance, 11*, 585–616.
Hwang, R. C., Wei, H. C., Lee, J. C., & Lee, C. F. (2008). On prediction of financial distress using the discrete-time survival model. *Journal of Financial Studies, 16*, 99–129.
Hwang, R. C., Chung, H., & Chu, C. K. (2010). Predicting issuer credit ratings using a semiparametric method. *Journal of Empirical Finance, 17*(1), 120–137.
Hwang, R. C., Cheng, K. F., & Lee, C. F. (2009). On multiple-class prediction of issuer crediting ratings. *Journal of Applied Stochastic Models in Business and Industry, 25*, 535–550.
Hwang, R. C., Cheng, K. F., & Lee, J. C. (2007). A semiparametric method for predicting bankruptcy. *Journal of Forecasting, 26*, 317–342.
Hwang, S., & Salmon, M. (2004). Market Stress and Herding. *Journal of Empirical Finance, 11*, 585–616.
Iacobucci, A. (2003). Spectral Analysis for Economic Time Series, Documents de Travail de l'OFCE.
Ibbotson Associates. (1994). *Chicago, Stocks, Bonds, Bills and Inflation 1994 Yearbook*. Chicago: Ibbotson Associates, Inc.
Ibbotson, R. (1975). Price performance of common stock new issues. *Journal of Financial Economics, 2*, 235–72.
Ibbotson, R. G., Diermeier, J. J., & Siegel, L. B. (2006). The demand for capital market returns: a new equilibrium theory. In W. N. Goetzmann & R. G. Ibbotson (Eds.), *The equity risk premium*. Oxford: Oxford University Press.
Ichiue, H., & Koyama, K. (2008). Regime Switches Exchange Rate Volatility and Uncovered Interest Parity. Working Paper.
Ignizio, J. P., & Romero, C. (2003). Goal programming. *Encyclopedia of Information Systems, 2*, 489–500.
Ignizio, J. P. (1976). An approach to the capital budgeting problem with multiple objectives. The Engineering Economist, 259–272.
Ikenberry, D., & Ramnath, S. (2002). Underreaction to self-selected news events: The case of stock splits. *Review of Financial Studies, 15*, 489–526.
Ikenberry, D., Lakonishok, J., & Vermaelen, T. (1995). Market underreaction to open market share repurchases. *Journal of Financial Economics, 39*, 181–208.

Ikenberry, D., Lakonishok, J., & Vermaelen, T. (2000). Stock repurchases in Canada: performance and strategic trading. *Journal of Finance, 55*, 2373–2397.

IMF. (2005). Does Inflation Targeting Work in Emerging Markets? *World Economic Outlook, Chapter, 4*, 161–178.

Imhoff, E., & Lobo, G. (1984). Information content of analysts' composite forecast revisions. *Journal of Accounting Research, 22*(2), 541–554.

Imhoff, E. A., & Pare, P. (1982). Analysis and comparison of earnings forecasts agents, Journal of Accounting Research, Fall. 429–439.

Ingersoll, J. E. (2006). The subjective and objective evaluation of incentive stock options. *The Journal of Business, 79*, 453–487.

Inoue. (2001). Testing for Distributional Change in Time Series. Econometric Theory 17, 156–187.

Intriligator, M. D. (2002). *Mathematical 0ptimization and Economic Theory*. Philadelphia, PA: Society for Industrial and Applied Mathematics.

Irani, A. (2004). The effect of Regulation Fair Disclosure on the relevance of conference calls to financial analysts. *Review of Quantitative Finance and Accounting, 22*, 15–28.

Irizarry, R. A. (2004). Choosing smoothness parameters for smoothing splines by minimizing an estimate of risk. Johns Hopkins University, Dept. of Biostatistics Working Papers, p. 30.

Ishizaka, A., Balkenborg, D., & Kaplan, T. (2011). Influence of aggregation and measurement scale on ranking a compromise alternative in AHP. *The Journal of the Operational Research Society, 62*, 700–711.

Ito, T. (1990). Foreign exchange rate expectations: Micro survey data. *American Economic Review, 80*, 434–449.

Ito, T. (1994). Short-run and long-run expectations of the yen/dollar exchange rate. *Journal of the Japanese and International Economies, 8*, 119–143.

Ittner, C., Lambert, R., & Larcker, D. (2003). The structure and performance consequences of equity grants to employees of new economy firms. *Journal of Accounting and Economics, 34* (1–3), 89–127.

Ittner, C. D., Larcker, D. F., & Rajan, M. V. (1997). The choice of performance measures in annual bonus contracts. *Accounting Review, 72*(2), 231–255.

Ivrendi, M., & Guloglu, B. (2010). Monetary shocks, exchange rates and trade balances: Evidence from inflation targeting countries. Economic Modelling, volume 27, Issue 5, September 2010, 1144–1155.

Jach, A., & Kokoszka, P. S. (2004). Subsampling unit root tests for heavy-tailed observations. Methodology and Computing in Applied Probability 6, 73–97.

Jackson, D. (1976). *Jackson Personality Inventory Manual*. Goshen, NY: Research Psychologists Press.

Jackson, D. (1977). Reliability of the Jackson personality inventory. *Psychological Reports, 40*, 613–614.

Jackson, D., Hourany, D., & Vidmar, N. (1972). A four dimensional interpretation of risk taking. *Journal of Personality, 40*, 433–501.

Jackwerth, J. C. (2000). Recovering risk aversion from options prices and realized returns. *Review of Financial Studies, 13*, 433–451.

Jacquier, E., Polson, N., & Rossi, P. (1994). Bayesian analysis of stochastic volatility models (with discussion). *Journal of Business and Economic Statistics, 12*(4), 371–417.

Jaffee, D. (2003). The interest rate risk of Fannie Mae and Freddie Mac. *Journal of Financial Services Research, 24*(1), 5–29.

Jagannathan, R., Kaplin, A., & Sun, S. (2003). An evaluation of multi-factor CIR models using LIBOR, swap rates, and cap and swaption prices. *Journal of Econometrics, 116*, 113–146.

Jagannathan, R., & Wang, Z. Y. (1996). The Conditional CAPM and the Corss-Section of Expected Returns. *Journal of Finance, 51*(1), 3–53.

Jagannathan, M., Stephens, C. P., & Weisbach, M. S. (2000). Financial flexibility and the choice between dividends and stock repurchases. *Journal of Financial Economics, 57*, 355–384.

Jagannathan, R., Kaplin, A., & Sun, S. (2003b). An Evaluation of Multi-factor CIR Models Using LIBOR, Swap Rates, and Cap and Swaption Prices. *Journal of Econometrics, 116*, 113–146.

Jagannathan, R., & Ma, T. (2003). Risk reduction in large portfolios: Why imposing the wrong constraints helps. *Journal of Finance, 58*, 1651–1684.

Jaggi, B. (1980). Further evidence on the accuracy of management forecasts vis-à-vis analyst' forecasts. The Accounting Review, January, 96–101.

Jain, K., Duin, R., & Mayo, J. (2000). Statistical Pattern Recognition: A Review. *IEEE Transactions on Pattern Analysis and Machine Intelligence, 22*, 4–37.

Jain, A. K., Murty, M. N., & Flynn, P. J. (1999). Data clustering: a review. *ACM Computing Surveys, 31*, 264–323.

Jame, R. E. (2011). Pension Fund Herding and Stock Return. Working paper. Available at: http://ssrn.com [Accessed 21 December 2011].

Jamshidian F. (1987). Pricing of contingent claims in the one-factor term structure model. Working Paper, Merrill Lynch Capital Markets.

Jamshidian, F. (1989). An exact bond option formula. *Journal of Finance, 4*(1), 205–209.

Jarque, C. M., & Bera, A. K. (1987). A test for normality of observations and regression residuals. *International Statistics Review, 55*, 163–172.

Jarrow, R., & Yu, F. (2001). Counterparty risk and the pricing of defaultable securities. *The Journal of Finance, 56*, 1765–1799.

Jarrow, R., & Turnbull, S. (1995). Pricing derivatives on financial securities subjectto credit risk. *Journal of Finance, 50*, 53–86.

Jarrow, R., Li, H., & Zhao, F. (2007). Interest Rate Caps. Smile. Too! But Can the LIBOR Market Models Capture It? *Journal of Finance, 62*, 345–382.

Jarrow, R., & TurnBull, S. (1996). *Derivative Securities*. Cincinnati: South-Western College Publishing.

Jayadeve, M. (1996). Mutual Fund Performance: An Analysis of Monthly Returns. *Finance India, X*(1), 73–84.

Jegadeesh, N., Kim, J., Krische, S., & Lee, C. (2004). Analyzing the analysts: When do recommendations add value? *Journal of Finance, 59*, 1083–24.

Jegadeesh, N. (1990). Evidence of predictable behavior of security returns. *Journal of Finance, 45*, 881–898.

Jegadeesh, N. (2000). Long-term performance of seasoned equity offerings: benchmark errors and biases in expectations. *Financial Management, 29*, 5–30.

Jegadeesh, N., & Karceski, J. (2009). Long-run performance evaluation: correlation and heteroskedasticity-consistent tests. *Journal of Empirical Finance, 16*, 101–111.

Jegadeesh, N., & Titman, S. (1993). Returns to buying winners and selling losers: implications for stock market efficiency. *Journal of Finance, 48*, 65–91.

Jegadeesh, N., & Titman, S. (2001). Profitability of momentum strategies: an evaluation of alternative explanations. *Journal of Finance, 56*, 699–720.

Jensen, M. C. (1993). The modern industrial revolution, exit, and the failure of internal control systems. *The Journal of Finance, 48*(3), 831–880.

Jensen, G. R., Solberg, D. P., & Zorn, T. S. (1992). Simultaneous determination of insider ownership, debt, and dividend policies. *Journal of Financial & Quantitative Analysis, 27*, 247–263.

Jensen, M. C. (1968). The performance of mutual funds in the period 1945-1964. *The Journal of Finance, 23*, 389–416.

Jensen, M. C., & Meckling, W. H. (1976). Theory of the firm: managerial behavior, agency costs and ownership structure. *Journal of Financial Economics, 3*, 305–360.

Jensen, M. C. (1986). Agency costs of free cash flow, corporate finance, and takeovers. *American Economic Review, 76*, 323–329.

Jensen, M. C. (1969). Risk, the pricing of capital assets, and the evaluation of investment portfolios. *Journal of Business, 42*(2), 167–247.

Jensen, M. C. (1972). Optimal utilization of market forecasts and the evaluation of investment performance. In G. P. Szego & K. Shell (Eds.), *Mathematical Methods in Investment and Finance*. Amsterdam: Elsevier.

Jensen, P. H., & Webster, E. (2007). Labelling Characteristics And Demand For Retail Grocery Products In Australia. *Australian Economic Papers, 47*(2), 129–140.

Ji, P., & Jin, J. (2010). UPS delivers optimal phase diagram in high dimensional variable selection. Preprint.

Jiang, G. J. (1998). Nonparametric Modelling of US Term Structure of Interest Rates and Implications on the Prices of Derivative Securities. *Journal of Financial and Quantitative Analysis, 33*, 465–497.

Jiang, J. (1993). Factors associated with analyst's earnings forecast accuracy underdifferent assumptions in Taiwan. Unpublished thesis, School of Business, Taiwan University.

Jiao, L., & Liu, Y. (2010). Geographic Field Model based hedonic valuation of urban open spaces in Wuhan, China. *Landscape and Urban Planning, 98*(1), 47–55.

Jobson, J., & Korkie, B. (1981). Performance hypothesis testing with the Sharpe and Treynor measures. *Journal of Finance, 36*, 889–908.

Jobson, J. D., & Korkie, B. (1982). Potential performance and tests of portfolio efficiency. *Journal of Financial Economics, 10*, 433–466.

Joe, H. (1997a). *Multivariate Models and Dependence Concepts*. New York, NY: Chapman & Hall/CRC.

Joe, H. (1997b). *Multivariate Models and Dependence Concepts*. London: Chapman and Hall.

Joe, H., & Xu, J. J. (1996). The estimation method of inference functions for margins for multivariate models, Dept. of Statistics, University of British Columbia, Tech. Rept. 166.

Joe, S., & Kuo, F. Y. (2003). Remark on Algorithm 659: Implementing Sobol's Quasirandom Sequence Generator. *ACM Transactions on Mathematical Software, 29*(1), 49–57.

Johannes, M. (2004). The Statistical and Economic Role of Jumps in Interest Rates. *Journal of Finance, 59*, 227–260.

Johansen, S. (1988). Statistical analysis of cointegrating vectors. *Journal of Economics and Dynamic Control, 12*, 231–254.

Johansen, S. (1995). *Likelihood-based Inference in Cointegrated Vector Autoregressive Models*. Oxford: Oxford University Press.

John, K., & Williams, J. (1985). Dividends, dilution, and taxes: A signaling equilibrium. *Journal of Finance, 40*, 1053–1070.

John, K., Saunders, A., & Senbet, L. (2000). A theory of bank regulation and management compensation. *Review of Financial Studies, 13*, 95–125.

Johnson and Barz. (1999). Selecting Stochastic Processes For Modelling Electricity Prices. Energy Modelling and the Management of Uncertainty, Risk Publications.

Johnson, H., & Schill, M. J. (2006). Asset Pricing When Returns Are Nonnormal: Fama-French Factors versus Higher Order Systematic Comoments. *Journal of Business, 79*, 923–940.

Johnson, L. L. (1960). The theory of hedging and speculation in commodity futures. *Review of Economic Studies, 27*, 139–151.

Johnson, L., Reither, C., & Swieringa, R. (1995). Toward reporting comprehensive income. *Accounting Horizons, 9*(December), 128–137.

Johnson, N., & Kotz, S. (1969). *Discrete Distributions*. Boston: Houghton Mifflin.

Johnson, N. J. (1978). Modified t tests and confidence intervals for asymmetrical populations. *Journal of the American Statistical Association, 73*, 536–544.

Johnson, N. L. (1949). Systems of Frequency Curves Generated by Methods of Translation. *Biometrica, 36*, 149–176.

Johnson, N. L., Kotz, S., & Balakrishnan, N. (1994). *Continuous Univariate Distributions* (2nd ed., Vol. 1). New York: John Wiley & Sons.

Johnson, R. (2004). Rating agency actions around the investment grade boundary. *Journal of Fixed Income, 13*(4), 25–37.

Johnson, R., & Wichern, D. (2007). Applied Multivariate Statistical Analysis, Sixth Edition. Pearson Education.

Johnson, S. A. (2003). Debt maturity and the effects of growth opportunities and liquidity risk on leverage. *Review of Financial Studies, 16*, 209–236.

Johnson, S., Boone, P., Breach, A., & Friedman, E. (2000). Corporate governance in the Asian financial crisis. *Journal of Financial Economics, 58*, 141–186.

Johnson, W., & Schwartz, W. Jr. (2001). Are investors misled by "Pro Forma" earnings? Working Paper, University of Iowa and University of Arizona.

Johnson, W. B. (1979). The cross-sectional stability of financial patterns. *Journal of Financial and Quantitative Analysis, 14*, 1035–1048.

Johnston, J., & Dinardo, J. (1996). Econometrics Methods (Fourth ed.), McGraw-Hill, New York.

Johnston, J., & DiNardo, J. (1997). *Econometrics Method*. Mc-Graw Hill.

Johnstone, D. (2002). More on the relative efficiency of probability scoring rules". *Journal of Statistical Computation and Simulation, 72*, 11–14.

Jondeau, E., & Rockinger, M. (2006). The Copula-GARCH model of conditional dependencies: An international stock market application. *Journal of International Money and Finance, 25*, 827–853.

Jondeau, E., & Rockinger, M. (2002). Conditional Dependency of Financial Series: The Copula-GARCH Model. FAME Working Paper.

Jondrow, J., Lovell, C. A., Materov, I. S., & Schmidt, P. (1982). On the estimation of technical inefficiency in the stochastic frontier production function model. *Journal of Econometrics, 19*, 233–238.

Jones, C. M., Kaul, G., & Lipson, M. L. (1994). Transactions, volume, and volatility. *The Review of Financial studies, 7*(4), 631–651.

Jones, J. Earnings management during import relief investigations. *Journal of Accounting Research, 29*, 193–228.

Jones, R., & Wu, Y. (2010). Executive compensation, earnings management and share holder litigation. *Review of Quantitative Finance and Accounting, 35*, 1–20.

Joo, J. H., & Qiu, P. (2009). Jump detection in a regression curve and its derivative. *Technometrics, 51*(3), 289–305.

Joreskog, K. G. (1977). Structural equation models in the social sciences: specification estimation and testing. In P. R. Krishnaiah (ed.).

Joreskog, K. G., & Goldberger, A. S. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of American Statistical Association, 70*, 631–639.

Joreskog, K. G., & Sorbom, D. (1981a). *LISREL V, analysis of linear structural relationships by the method of maximum likelihood*. Chicago: National Educational Resources.

Joreskog, K. G., & Sorbom, D. (1989). *LISREL 7: A guide to the program and applications*, 2nd ed. SPSS.

Joreskog, K., & Goldberger, A. (1975). Estimation of a model with multiple indicators and multiple causes of a single latent variable. *Journal of the American Statistical Association, 70*, 631–639.

Joreskog, K., & Sorbom, D. (1981b). *LISREL V: Analysis of Linear Structural Relationships by the Method of Maximizing Likelihood, User's Guide*. IL: National Educational Resources, Inc.

Joreskog, K. G. (1967). Some contributions to maximum-likelihood factor analysis. *Psychometrika, 32*, 443–82.

Jorge, S., Pilar, G., & Luis, M. D. (2006). Mutual fund performance and benchmark choice: The Spanish case. *Applied Financial Economics Letters, 2*(5), 317–321.

Jorion, P. (1996a). Risk2: Measuring the Risk in Value at Risk. *Financial Analysts Journal* (November/December), 47–56.

Jorion, P. (1997). In Defense of VAR. *Derivatives Strategy, 2*(April), 20–23.

Jorion, P. (2000a). *Value at Risk: the new benchmark for managing financial risk*. New York: McGraw-Hill.

Jorion, P. (2001). *Value at Risk* (2nd ed.). New York, N.Y.: McGraw Hill.

Jorion, P., & Zhang, G. (2007). Information effects of bond rating changes: The role of the rating prior to the announcement. *Journal of Fixed Income, 16*, 45–59.

Jorion, P. H. (1996b). *Value at Risk: A new benchmark for Measuring Derivative Risk*. New York: Irwin Professional Publishers.

Jorion, P. (1988). On jump processes in the foreign exchange and stock markets. *The Review of Financial Studies, 1*, 427–445.

Jorion, P. (2000b). Risk Management Lessons from Long Term Capital Management. *European Financial Management, 6*, 277–300.

Jorion, P. (1992). Portfolio optimization in practice. *Financial Analysts Journal, 48*, 68–74.

Jorion, P. (2003). Portfolio optimization with tracking error constraints. *Financial Analysts Journal, 59*, 70–82.

Joshi, M. S., & Kainth, D. (2004). Rapid and Accurate Development of Prices and Greeks For Nth to Default Credit Swaps in the Li Model. *Quantitative Finance, 4*, 266–275.

Chang, J.-R., Hung, M.-W., & Lee, C.-f. (2003). An Intemporal CAPM approach to Evaluate Mutual Fund Performance. *Review of Quantitative Finance and Accounting, 20*, 425–433.

Jow-Ran C., Mao-Wei, H., & Cheng-Few, L. (2003). An Intemporal CAPM approach to evaluate mutual fund performance. Review of Quantitiative Finance and Accounting, 20, pp. 414–433.

Judge, G., Hill, R., Griffiths, W., Lutkepohl, H., & Lee, T. C. (1988). *Introduction to the Theory and Practice of Econometrics* (2nd ed.). New York: John Wiley.

Judge, G., Hill, R., Griffiths, W., Lutkepohl, H., & Lee, T. C. (1980). *The Theory and Practice of Econometrics*. New York: John Wiley.

Julier, S. J., & Uhlmann, J. K. (1997). A new extension of the Kalman filter to nonlinear systems. In Int. Symp. Aerospace/Defense Sensing, Simul. and Controls, volume 3, page 26. Citeseer.

Julio, B., & Ikenberry, D. (2004). Reappearing dividends. *Journal of Applied Corporate Finance, 16*, 89–100.

Jung, K., Kim, Y., & Stulz, R. M. (1996). Timing, investment opportunities, managerial discretion, and the security issue decision. *Journal of Financial Economics, 42*, 159–85.

Jurczenko, E., & Maillet, B. (2002). The Four Moment Capital Asset Pricing Model: Some Basic Results. Working Paper, Ecole supérieure de commerce de Paris, Paris, France.

Kücher, U., Neumann, K., Sørensen, M., & Streller, A. (1999). Stock returns and hyperbolic distributions. Mathematical and Computer Modelling, 29, 1–15.

Kadam, A., & Lenk, P. (2008). Bayesian inference for issuer heterogeneity in credit rating migration. *Journal of Banking and Finance, 32*, 2267–2274.

Kahl, C., & Jäckel, P. (2005). Not-so-complex logarithm in the Heston model. *Wilmott Magazine, 19*, 94–103.

Kahle, K. (2002). When a buyback isn't a buyback: Open market repurchases and employee options. *Journal of Financial Economics, 63*, 235–261.

Kahnemann, D., & Tversky, A. (1979). Prospect theory; an analysis of decisions under risk. *Econometrica, 47*, 262–291.

Kahya, E., Maggina, A., & Theodossiou, P. (1993). The Information Content of Accounting Earnings in Greece: An Empirical Investigation. *Journal of Multinational Financial Management, 3*, 143–56.

Kajshmi, N., Malabika, D., & Murugesan, B. (2008). Performance of the Indian Mutual Funds; A Study with special reference to growth schemes. Asia-Pacific Business Review, Vol. Iv, No.3, July-September, pp. 75–81.

Kakes, J. I., & Sturm, J. E. (2002). Monetary policy and bank lending: evidence from German banking groups. *Journal of Banking and Finance, 26*, 2077–2092.

Kalay, A., & Michaely, R. (1993). Dividends and taxes: A reexamination, Working Paper.

Kalay, A., & Loewenstein, U. (1986). The informational content of the timing of dividend announcements. *Journal of Financial Economics, 16*, 373–388.

Kalbfleisch, J. D., & Prentice, R. L. (1980). *The Statistical Analysis of Failure Time Data*. New York: John Wiley and Sons, Inc.

Kalev, P. S., & Duong, H. N. (2008). A test of the Samuelson Hypothesis using realized range. *Journal of Futures Markets, 28*, 680–696.

Kallsen, J., & Taqqu, M. (1998). Option pricing in ARCH-type models. *Mathematical Finance, 8*, 13–26.

Kalman, R. E. (1960). A new approach to linear filtering and prediction problems. *Journal of basic Engineering, 82*(1), 35–45.

Kamara, A. (1994). Liquidity, taxes, and short-term Treasury yields. *Journal of Financial and Quantitative Analysis, 29*, 403–417.

Kamath, R. R., & Khaksari, S. Z. (1991). Real estate investment: an application of the analytic hierarchy process. The Journal of Financial and Strategic Decisions, 73–100.

Kamesaka, A., Nofsinger, J. R., & Kawakita, H. (2003). Investment patterns and performance of investor groups in Japan. *Pacific-Basin Finance Journal, 11*, 1–22.

Kaminski, V. (1997). The Challenge of Pricing and Risk Managing Electricity Derivatives. The U.S. Power Market, 149–171.

Kaminsky, G., & Schmukler, S. L. (2002). Emerging Market Instability: Do Sovereign Ratings Affect Country Risk and Stock Returns? *World Bank Economic Review, 16*, 171–195.

Kamstra, M., & Kennedy, P. (1998). Combining qualitative forecasts using logit. *International Journal of Forecasting, 14*, 83–93.

Kanagaretnam, K., Lobo, G., & Mathieu, A. (2012). CEO compensation mix and analysts forecast accuracy and bias. The Review of Quantitative Finance and Accounting, forthcoming.

Kandel, S. A., & Stambaugh, R. F. (1987). On correlations and inferences about mean-variance efficiency. *Journal of Financial Economics, 18*, 61–90.

Kandel, S. A., & Stambaugh, R. F. (1989). A mean-variance framework for tests of asset pricing models. *Review of Financial Studies, 2*, 125–156.

Kandel, S. A. (1984). The likelihood ratio test statistic of mean variance efficiency without a riskless asset. *Journal of Financial Economics, 13*, 575–592.

Kane, A., & Marcus, A. (1984). Conversion factor risk and hedging in the Treasury bond futures market. *Journal of Futures Markets, 4*(1), 55–64.

Kane, A., & Marcus, A. (1986a). The quality option in the Treasury bond futures market: An empirical assessment. *Journal of Futures Markets, 6*(2), 231–248.

Kane, A., & Marcus, A. (1986b). Valuation and optimal exercise of the wild card option in the Treasury bond futures market. *Journal of Finance, 41*(1), 195–207.

Kao, C., & Wu, C. (1994). Tests of dividend signaling using the March-Merton model: A generalized friction approach. *Journal of Business, 67*, 45–68.

Kao, L. J., & Lee, C. F. (2011). Alternative statistical method to forecast bond rating. Working paper, Rutgers University.

Kaplan, R. S. (1986). Accounting lag: the obsolescence of cost accounting systems. *California Management Review, 28*, 175–199.

Kaplan, R. S., & Norton, D. P. (1992). The balanced scorecard-measures that drive performance. Harvard Business Review, 71–79.

Kaplan, R. S., & Urwitz, G. (1979). Statistical models of bond ratings: A methodological inquiry. *Journal of Business, 52*, 231–261.

Kaplan, S., & Zingales, L. (1997). Do Financing Constraints Explain Why Investment Is Correlated with Cash Flow? *Quarterly Journal of Economics, 112*, 169–215.

Karali, B., & Thurman, W. N. (2010). Components of grain futures price volatility. *Journal of Agricultural and Resource Economics, 35*, 167–182.

Karanasos, M., & Kartsaklas, A. (2009). Dual long-memory, structural breaks and the link between turnover and the range-based volatility. *Journal of Empirical Finance, 16*, 838–851.

Karanovic, G., Baresa, S., & Bogdan, S. (2010). Techniques for managing projects risk in capital budgeting process. *UTMS Journal of Economics, 1*, 55–66.

Karlin, S., & Taylor, H. M. (1981). *A Second Course in Stochastic Processes*. San Diego: Academic Press.

Karolyi, G. A. (2002). Did the Asian financial crisis scare foreign investors out of Japan? *Pacific-Basin Finance Journal, 10*, 411–442.

Karpoff, J. M. (1987). The relation between price changes and trading volume: a survey. *Journal of Financial and Quantitative Analysis, 22*, 109–126.

Kasch, M., & Caporin, M. (2012). Volatility threshold dynamic conditional correlations: An international analysis. Available at SSRN: http://ssrn.com/abstract=968233.

Kassie, G., Abdulai, A., & Wollny, C. (2011). Heteroscedastic hedonic price model for cattle in the rural markets of central Ethiopia. *Applied Economics, 43*(24), 3459–3464.

Kau, J. B., Lee, C. F., & Sirmans, C. F. (1986). Urban econometrics: Model developments and empirical results. Research in Urban Economics, Vol. 6 JAI Press.

Kau, J. B., & Lee, C. F. (1976). Functional form, density gradient and the price elasticity of demand for housing. *Urban Studies, 13*(2), 181–192.

Kaufman, L., & Rousseeuw, P. J. (1990). Finding Groups in Data: An Introduction to Cluster Analysis. 9th Edition, Wiley-Interscience.

Kaufmann, D., Kraay, A., & Zoido-Lobaton, P. (1999a). Aggregating Governance Indicators. World Bank Policy Research Department Working Paper 2195.

Kaufmann, D., Kraay, A., & Zoido-Lobaton, P. (1999b). Governance Matters. World Bank Policy Research Department Working Paper 2196.

Kaul, A., & Sapp, S. (2006). Y2K fears and safe haven trading of the US dollar. *Journal of International Money and Finance, 25*, 760–79.

Kaur Dhanda, S., Batra, G. S., & Anjum, B. (2012). Performance Evaluation Of Selected Open Ended Mutual Funds In India, International Journal Of Marketing, Financial Services & Management Research, Vol.1 No. 1, January 2012, ISSN 2277 3622 Online Available at indianresearchjournals.com.

Kavajecz, K. A., & Odders-White, E. R. (2001). An examination of changes in specialists' posted price schedules. *Review of Financial Studies, 14*, 681–704.

Kavussanos, M. G., & Visvikis, I. D. (2000). Futures hedging when the structure of the underlying asset changes: the case of BIFFEX contracts. *Journal of Futures Markets, 20*, 775–801.

Kavussanos, M. G., & Visvikis, I. D. (2008). Hedging effectiveness of the Athens stock index futures contracts. *European Journal of Finance, 14*, 243–170.

Kayhan, A., & Titman, S. (2007). Firms' histories and their capital structures. *Journal of Financial Economics, 83*, 1–32.

Kazakevicius, V., Leipus, R., & Viano, M. (2004). Stability of random coefficient ARCH models and aggregation schemes. *Journal of Econometrics, 120*, 139–158.

Ke, B., & Yu, Y. (2006). The effect of issuing biased earnings forecasts on analysts' access to management and survival. *Journal of Accounting Research, 44*(5), 965–999.

Kealhofer, S. (1996). Managing default risk in portfolios of derivatives. In Derivative CreditRisk, London: Risk Publications.

Keane, M. P., & Runkle, D. E. (1990). Testing the rationality of price forecasts: New evidence from panel data. *American Economic Review, 80*, 714–735.

Keeley, M. C. (1990). Deposit insurance, risk, and market power in banking. *American Economic Review, 80*, 1183–1200.

Keeley, M. C. (1992). Bank capital regulation in the 1980s: Effective or ineffective? In Anthony Saunders, Gregory F. Udell and Lawrence J. White, eds.: Bank Management and Regulation: A Book of Readings. Mayfield Publishing Company: Mountain View, California.

Keim, D. B., & Stambaugh, R. F. (1986). Predicting returns in the stock and bond markets. *Journal of Financial Economics, 17*, 357–390.

Kennedy, P. (2008). A Guide to Econometrics. 6th Edition, Wiley-Blackwell.

Kenneth R. French's website, http://mba.tuck.dartmouth.edu/pages/faculty/ken.french/data_library.html.

Kenney, J. F., & Keeping, E. S. (1962). Skewness. §7.10 in Mathematics of Statistics, Princeton, NJ: Van Nostrand, 100–101.

Keran, M. (1971). Expectations, Money, and Stock Market. Review, Federal Reserve Bank of St. Louis, Jan., 16–31.

Khindanova, I., Rachev, S., & Schwartz, E. (2000). Stable modeling of value at risk. working paper, University of California, Santa Barbara.

Khurshid, S. M. Z., Rohit & Sing, G. P. (2009). Level and trends of competition among the mutual funds in India. Research Journal of Business Management, 3(2), 47–67.

Kiefer, N., & Larson, C. (2007). A simulation estimator for testing the time homogeneity of credit rating transitions. *Journal of Empirical Finance, 14*, 818–835.

Kiker, B. F., & Mendes-de-Oliveira, M. (1990). Estimation and Valuation of Non-leisure Time. *Oxford Bulletin of Economics and Statistics, 52*, 115–141.

Kilcollin, T. (1982). Difference systems in financial futures markets. *Journal of Finance, 37*(5), 1883–1897.

Kim, S. N., Shephard, N., & Chib, S. (1998a). Stochastic volatility: likelihood inference and comparison with ARCH models. *Review of Economic Studies, 65*, 365–393.

Kim, C., Nelson, C., & Startz, R. (1998b). Testing for mean reversion in Heteroskedastic data based on Gibbs-sampling-augmented randomization. *Journal of Empirical Finance, 5*, 131–154.

Kim, C., & Nelson, C. (1999). Has the U.S. economy become more stable? A Bayesian approach based on a Markov-switching model of the business cycle. *Review of Economics and Statistics, 81*(4), 608–616.

Kim, C., Morley, J. C., & Nelson, C. R. (2001). Does an intertemporal tradeoff between risk and return explain mean reversion in stock price? *Journal of Empirical Finance, 8*, 403–426.

Kim, C. S., & Marron, J. S. (2006). Sizer for jump detection. *Journal of nonparametric statistics, 18*(1), 13–20.

Kim, D. (1995). The errors in the variables problem in the cross-section of expected stock returns. *Journal of Finance, 50*(5), 1605–1634.

Kim, D. (1997). A reexamination of firm size, book-to-market, and earnings price in the cross-section of expected stock returns. *Journal of Financial and Quantitative Analysis, 32*(4), 463–489.

Kim, D. (2010). Issues related to the errors-in-variables problems in asset pricing tests. Handbook of Quantitative Finance and Risk Management, Part V, Chapter 70, 1091–1108.

Kim, D., & Santomero, A. M. (1988). Risk in banking and capital regulation. *Journal of Finance, 42*, 1219–1233.

Kim, E. H. (1978). A mean-variance theory of optimal capital structure and corporate debt capacity. *Journal of Finance, 33*, 45–63.

Kim, I., Ramaswamy, K., & Sundaresan, S. (1993). Does Default Risk in Coupons Affect the Valuation of Corporate Bonds? A Contingent Claims Model. *Financial Management, 22*, 117–131.

Kim, I.-J. (1990). The Analytical Valuation of American Options. *Review of Financial Studies, 3*(4), 547–572.

Kim, J. B., Song, B.Y., & Zhang, L. (2009). Internal control quality and analyst forecast behavior: Evidence from SOX Section 404 disclosures. Working Paper, City University of Hong Kong.

Kim, K., & Wei, S. (2002). Offshore investment funds: monsters in emerging markets? *Journal of Development Economics, 68*(1), 205–224.

Kim, S. H., & Farragher, E. I. (1981). Current capital budgeting practices. *Management Accounting, 6*, 26–30.

Kim, S. G., Cho, S. H., Lambert, D., & Roberts, R. (2010). Measuring the value of air quality: application of the spatial hedonic model. *Air Quality, Atmosphere & Health, 3*(1), 41–51.

Kim, T. H., Leybourne, S., & Newbold, P. (2002). Unit root tests with a break in innovation variance. *Journal of Econometrics, 109*, 365–387.

Kim, Y. S., Rachev, S. T., Chung, D. M., & Bianchi, M. L. (2009b). The modified tempered stable distribution, GARCH-models and option pricing. *Probability and Mathematical Statistics, 29*, 91–117.

Kim, Y. S., Rachev, S. T., Bianchi, M. L., & Fabozzi, F. J. (2010b). Tempered stable and tempered infinitely divisible GARCH models. *Journal of Banking and Finance, 34*, 2096–2109.

Kim, Y. S., Rachev, S. T., Bianchi, M. L., Mitov, I., & Fabozzi, F. J. (2011). Time series analysis for financial market meltdowns. *Journal of Banking and Finance, 35*, 1879–1891.

Kim, Y., Choi, H., & Oh, H. S. (2008). Smoothly clipped absolute deviation on high dimensions. *Journal of the American Statistical Association, 103*, 1665–1673.

Kindleberger, C. P. (2005). Manias, Panics, and Crashes: A History of Financial Crises.

King, B. F. (1966). Market and industry factors in stock price behavior. *The Journal of Business, 39*, 139–190.

King, R. R., Smith, V. L., Williams, A., & Van Boening, M. (1990). The robustness of bubbles and crashes. Working paper, University of Washington, Seattle, WA.

King, R., Plosser, C., Stock, J., & Watson, M. (1991). Stochastic trends and economic fluctuations. *The American Economic Review, 81*(4), 819–840.

Kirman, A., & Teyssiere, G. (2005). Testing for bubbles and change-points. *Journal of Economic Dynamic Control, 29*, 765–799.

Kirman, A. P. (1992). Whom or what does the representative individual represent? *Journal of Economic Perspectives, 6*, 117–136.

Kiviet, J. F. (1987). Testing Linear Econometric Models. University of Amsterdam.

Klassen, K., & Mawani, A. (2000). The impact of financial and tax reporting incentives on option grants to Canadian CEOs. *Contemporary Accounting Research, 17*(2), 227–262.

Kleibergen, F. (2002). Pivotal statistics for testing structural parameters in instrumental variables regression. *Econometrica, 70*, 1781–1803.

Klein, A. (2002). Audit committee, board of director characteristics, and earnings management. *Journal of Accounting and Economics, 33*(3), 375–400.

Kloeden, P. E., & Platen, E. (1999). *Numerical Solution of Stochastic Differential Equations*. New York: Springer and Verlag.

Knittel, C. R., & Roberts, M. (2001). An Empirical Examination of Deregulated Electricity Prices. Working paper, PWP-087, University of California Energy Institute.

Koenker, R. (2005). *Quantile Regression, Econometric Society Monographs 38*. New York: Cambridge University Press.

Koenker, R. (2009). Quantile regression in R: a vignette.' Mimeo.

Koenker, R., & Bassett, G. S. (1978a). Regression quantiles. *Econometrica, 46*, 33–50.

Koenker, R., & Xiao, Z. (2002). Inference on the Quantile Regression Processes. *Econometrica, 70*, 1583–1612.

Koenker, R., & Xiao, Z. (2004). Unit Root Quantile Regression. *Journal of the American Statistical Association, 99*(467), 775–787.

Koenker, R., & Xiao, Z. (2006). Quantile Autoregression. *Journal of the American Statistical Association, 101*(475), 980–1006.

Koenker, R., & Bassett, G. (1978b). Regression quantiles. *Econometrica, 84*, 33–50.

Koenker, R., & Zhao, Q. (1996). Conditional quantile estimation and inference for ARCH models. *Econometric Theory, 12*, 793–813.

Koenker, R., & d'Orey, V. (1987). Computing regression quantiles. *Applied Statistics, 36*, 383–393.

Kohli, R., & Devaraj, S. (2004). Realizing the business value of information technology investments: An organizational process. MIS Quarterly Executive. 3.

Kokoszka, P. S., & Parfionovas, A. (2004). Bootstrap unit root tests for heavy tailed time series. In: Rachev, S.T. (Ed.), Handbook of Computational and Numerical Methods in Finance, 175–197.

Kolari, J. W., & Pynnonen, S. (2010). Event study testing with cross-sectional correlation of abnormal returns. *Review of Financial Studies, 23*, 3996–4025.

Kolbe, A. L., Read, J. A., & Hall, G. R. (1984). *The Cost of Capital: Estimating the Rate of Return for Public Utilities*. Cambridge, Mass.: MTT Press.

Kole, E., Koedijk, K., & Verbeek, M. (2007). Selecting copulas for risk management. *Journal of Banking and Finance, 31*, 2405–2423.

Kolmogorov A. N. (1933). Sulla Determinazione Empirica di una Legge di Distribuzione. Giornale dell.Istitutodegli Attuari 4, 83–91.

Konno, H., & Yamazaki, H. (1991). Mean-absolute deviation portfolio optimization model and its application to Tokyo stock market. *Management Science, 37*, 519–531.

Koopmans, T. C. (1950). *Statistical inference in dynamic economic models*. New York: John Wiley.

Kopa, M., & Post, T. (2009). A portfolio optimality test based on the first-order stochastic dominance criterion. *Journal of Financial & Quantitative Analysis, 44*(5), 1103–1124.

Kormendi, R., & Lipe, R. (1987). Earnings Innovations, Earnings Persistence and Stock Returns. *Journal of Business, 60*, 323–45.

Korpela, E., et al. (2001). SETI@home-Massively Distributed Computing for SETI. *Computing in Science and Engineering, 3*(1), 78–83.

Koschnick, T. (2008). Cluster & Security: Separating Application Environments. Int'l Supercomputing Conference and Exhibition, Dresden, 17–20.

Kothari, S. P., & Warner, J. B. (1997). Measuring long-horizon security price performance. *Journal of Financial Economics, 43*, 301–340.

Kothari, S. P., & Warner, J. B. (2007). Econometrics of event studies. In B. E. Eckbo (Ed.), *Handbooks of Corporate Finance: Empirical Corporate Finance*. Amsterdam: Elsevier/North-Holland.

Kothari, S. P., & Zimmerman, J. L. (1995). Price and Return Models. *Journal of Accounting and Economics, 20*, 155–92.

Kothari, S. P., Leone, A. J., & Wasley, C. E. (2005). Performance matched discretionary accruals. *Journal of Accounting and Economics, 39*, 163–197.

Koyuoglu, H. U., & Hickman, A. (1998). Reconcilable differences. *Risk, 11*(10), 56–62.

Kraus, A., & Ross, S. A. (1982). The Determination of Fair Profits for the Property-Liability Insurance Firm. *Journal of Finance, 37*(4), 1015–1026.

Kreyszig, E. (2010). *Advanced engineering mathematics* (10th ed.). New York, NY: Wiley.

Krishnamurthy, A. (2002). The bond/old-bond spread. *Journal of Financial Economics, 66*, 463–506.

Krishnan, C. N. V., Ritchken, P., & Thomson, J. (2010). Predicting credit spreads. *Journal of Financial Intermediation, 19*(4), 529–563.

Krishnan, J., & Krishnan, J. (1997). Litigation risk and auditor resignations. *The Accounting Review., 72*, 539–560.

Krishnaswami, S., & Subramaniam, V. (1998). Information asymmetry, valuation, and the corporate spin-off decision. *Journal of Financial Economics, 53*, 73–112.

Kristensen, D., & Shin, Y. (2008). Estimation of Dynamic Models with Nonparametric Simulated Maximum Likelihood. CREATES Research Paper 2008-58, University of Aarhus and Columbia University.

Kroner, K. E., & Ng, V. (1998). Modeling asymmetric comovements of asset returns. *Review of Financial Studies, 11*, 817–844.

Kroner, K. F., & Sultan, J. (1993). Time-varying distributions and dynamic hedging with foreign currency futures. *Journal of Financial and Quantitative Analysis, 28*, 535–551.

Kronmal, R. (1993). Spurious correlation and the fallacy of the ratio standard revisited. *Journal of the Royal Statistical Society A, 156*, 379–392.

Kross, W., Ro, B., & Schroeder, D. (1990). Earnings Expectations: The Analysts Information Advantage. *The Accounting Review, 65*, 461–76.

Kruse, S., & Nögel, U. (2005). On the pricing of forward starting options in Heston's model on stochastic volatility. *Finance and Stochastics, 9*, 223–250.

Kuan, C., & Hsu, C. (1998). Change-point estimation of fractionally integrated processes. *Journal of Time Series Analysis, 19*, 693–708.

Kuan, C. M., & Hornik, K. (1995). The generalized fluctuation test: A unifying view. *Econometric Reviews, 14*, 135–161.

Kuan, C.-M., Yeh, J.-H., & Hsu, Y.-C. (2009). Assessing value at risk with CARE, the Conditional Autoregressive Expectile models. *Journal of Econometrics, 150*, 261–270.

Kuipers, L., & Niederreiter, H. (1974). *Uniform Distribution of Sequence*. New York: John Wiley & Sons.

Kumar, K., Rajan, R. G., & Zingales, L. (2002). What determines firm size? Working paper, University of Chicago.

Kunczik, M. (2000). Globalization: News Media, Images of Nations and the Flow of International Capital with Special Reference to the Role of Rating Agencies. Paper presented at the IAMCR Conference, Singapore, 1–49.

Kunitomo, N. (1992). Improving the Parkinson method of estimating security price volatilities. *Journal of Business, 65*, 295–302.

Künsch, H. R. (1989). The Jackknife and the Bootstrap for General Stationary Observations. *Annals of Statistics, 17*, 1217–1241.

Kuo, C. (1996). Autocorrelation and seasonality in the real estate market. *Journal of Real Estate Finance and Economics, 12*, 139–162.

Kuosmanen, T. (2004). Efficient diversification according to stochastic dominance criteria. *Management Science, 50*(10), 1390–1406.

Kupiec, P. (1995). Techniques for verifying the accuracy of risk management models. *The Journal of Derivatives, 3*, 73–84.

Kutner, M. H., Nachtsheim, C. J., Neter, J., & Li, W. (2005). Applied Linear Statistical Models. Fifth Edition, McGraw-Hill/Irwin.

Kwan, C. C. Y. (1997). Portfolio selection under in-stitutional procedures for short selling: Normative and market- equilibrium considerations. *Journal of Banking & Finance, 21*, 369–391.

Kwan, S. H. (1996). Firm-specific information and the correlation between individual stocks and bonds. *Journal of Financial Economics, 40*, 63–80.

Kwapien, J., Oswiecimka, P., & Drozdz, S. (2007). Empirics versus RMT in financial cross-Correlations. *Acta Physica Polonica B, 38*, 4027–4039.

Kwok, C. C., & Tadesse, S. (2006). National culture and financial systems. *Journal of International Business Studies, 37*, 227–247.

Kwon, S. (2002). Financial analysts' forecast accuracy and dispersion: High-tech versus low-tech stocks. *Review of Quantitative Finance and Accounting, 19*, 65–91.

La Porta, R., Lakonishok, J., Shleifer, A., & Vishny, R. (1997). Good news for value stocks: further evidence on market efficiency. *Journal of Finance, 52*(2), 859–874.

La Porta, R., Lopez-De-Silanes, F., & Shleifer, A. (1999). Corporate ownership around the world. *Journal of Finance, 54*, 471–518.

La Porta, R., Lopez-de-Silanes, F., Shleifer, A., & Vishny, R. (1998). Law and Finance. *Journal of Political Economy, 106*, 1113–1155.

Ladd, G. W. (1966). Linear probability functions and discriminant functions. *Econometrica, 34*, 873–85.

Laeven, L. A., & Levine, R. (2009). Bank governance, regulation and risk taking. *Journal of Financial Economics, 93*, 259–275.

Lafuente, J., & Novales, A. (2003). Optimal hedging under departures from the cost-of-carry valuation: Evidence from the Spanish stock index futures markets. *Journal of Banking and Finance, 27*, 1053–1078.

Lahiri, K. (1976). Inflationary expectations: Their formation and interest rate effects. *The American Economic Review, 66*, 124–131.

Lai, K. (1990). An evaluation of survey exchange rate forecasts. *Economics Letters, 32*, 61–65.

Lai, T. L., & Xing, H. (2008). *Statistical Models and Methods for Financial Markets*. New York: Springer.

Lai, T. L., & Xing, H. (2010). Time series modeling and forecasting of volatilities of asset returns. Handbook of Quantitative Finance and Risk Management (C.F. Lee, A.C. Lee and J. Lee, eds.), 1417-1426. Springer-Verlag, New York.

Lai, T. L., & Xing, H. (2011). A simple Bayesian approach to multiple change-points. *Statistica Sinica, 21*, 539–569.

Lai, T. L., & Xing, H. (2012). Stochastic change-point ARX-GARCH models and their applications to econometric time series. Statistica Sinica 23, to appear.

Lai, T. L., Liu, H., & Xing, H. (2005). Autoregressive models with piecewise constant volatility and regression parameters. *Statistica Sinica, 15*, 279–301.

Lai, T. Y., Lee, C. F., & Tucker, A. L. (1992). An Alternative Method for Obtaining the Implied Standard Deviation. *Journal of Financial Engineering, 1*, 369–375.

Lakonishok, J., Shleifer, A., & Vishny, R. (1992). The impact of institutional trading on stock prices. *Journal of Financial Economics, 32*, 23–43.

Lakonishok, J., Shleifer, A., & Vishny, R. (1992b). The Structure and Performance of the Money Management Industry. Brookings Papers: Microeconomics 339–391.

Lakonishok, J., Shleifer, A., & Vishny, R. (1994). Contrarian investment, extrapolation, and risk. *Journal of Finance, 49*, 1541–1578.

Lakonishok, J., Shleifer, A., & Vishny, R. (1992). The impact of institutional trading on stock price. *Journal of Financial Economics, 32*, 23–44.

Lam, S. S., & Chng, B. F. (2006). Do executive stock option grants have value implications for firm performance? *Review of Quantitative Finance and Accounting, 26*, 249–274.

Lam, K., Liu, T., & Wong, W. K. (2010). A pseudo-Bayesian model in financial decision making with implications to market volatility, under- and overreaction. *European Journal of Operational Research, 203*, 166–175.

Lam, K., Liu, T., & Wong, W. K. (2012). A new pseudo Bayesian model with implications to financial anomalies and investors' behaviors. *Journal of Behavioral Finance, 13*(2), 93–107.

Lambert, R. (2001). Contracting theory and accounting. *Journal of Accounting and Economics, 32*(3), 3–87.

Lambert, R. A., Larcker, D. F., & Verrecchia, R. E. (1991). Portfolio considerations in valuing executive compensation. *Journal of Accounting Research, 29*, 129–149.

Lambert, R., & Larcker, D. (1987). An analysis of the use of accounting and market measures of performance in executive compensation contracts. *Journal of Accounting Research, 25*, 85–125.

Lamont, O. (1997). Cash Flow and Investment: Evidence from Internal Capital Markets. *Journal of Finance, 52*, 83–110.

Lamont, O. (2002). Macroeconomic forecass and microeconomic forecasters. *Journal of Economic Behavior and Organization, 48*, 265–280.

Lamoureux, C. G., & Lastrapes, W. D. (1990). Persistence in variance, structural change and the GARCH model. *Journal of Business and Economic Statistics, 8*, 225–234.

Lancaster, K. J. (1966). A New Approach to Consumer Theory. *Journal of Political Economy, 74*, 132–157.

Lando, D. (2004). *Credit Risk Modeling: Theory and Applications*. Princeton, New Jersey: Princeton University Press.

Lando, D., & Nielsen, M. S. (2010). Correlation in corporate default: Contagion or conditional independence? *Journal of Financial Intermediation, 19*(3), 355–372.

Lando, D., & Skodeberg, T. (2002). Analyzing ratings transitions and rating drift with continuous observations. *Journal of Banking and Finance, 26*, 423–444.

Lando, D. (1998). On Cox Processes and Credit Risky Securities. *Review of Derivative Research, 2*, 99–120.

Landsman, W. R., & Magliolo, J. (1988). Cross-Sectional Capital Market Research and Model Specification. *The Accounting Review, 64*, 586–604.

Lang, M., & Lundholm, R. (1996a). Corporate disclosure policy and analyst behavior. *The Accounting Review, 71*(4), 467–493.

Lang, L. H. P., & Litzenberger, R. H. (1989). Dividend announcements: Cash flow signaling vs. free cash flow hypothesis? *Journal of Financial Economics, 24*, 181–191.

Lang, M., Lins, K., & Miller, D. (2004). Concentrated Control, Analyst Following, and Valuation: Do Analysts Matter Most When Investors Are Protected Least?

Lang, M. H., & Lundholm, R. J. (1996b). Corporate disclosure policy and analyst behavior. *The Accounting Review, 71*(4), 467–492.

Langetieg, T. (1980). A multivariate model of the term structure. *Journal of Finance, 35*(1), 71–97.

Lanza, A., Manera, M., & McAleer, M. (2006). Modeling dynamic conditional correlations in WTI oil forward and futures returns. *Finance Research Letters, 3*, 114–132.

Larcker, D. F., & Rusticus, T. O. (2010). On the use of instrumental variables in accounting research. *Journal of Accounting and Economics, 49*, 186–205.

Larcker, D. F., Richardson, S. A., & Tuna, A. I. (2007). Corporate governance, accounting outcomes, and organizational performance. *The Accounting Review, 82*, 963–1008.

Larrain, G., Reisen, H., & von Maltzan, J. (1997). Emerging Market Risk and Sovereign Credit Ratings. OECD Development Center, 1–30.

Laster, D., Bennett, P., & Geoum, I. S. (1999). Rational bias in macroeconomic forecasts. Quarterly Journal of Economics, 293–318.

Latane, H. A., & Rendleman, R. J. (1976). Standard Deviation of Stock Price Ratios Implied by Option Premia. *Journal of Finance, 31*, 369–382.

Launie, J. J. (1971). The Cost of Capital of Insurance Companies. *Journal of Risk and Insurance, 38*(2), 263–268.

Launie, J. J. (1972). The Cost of Capital of Insurance Companies: Reply. *Journal of Risk and Insurance, 39*(3), 292–495.

Laurent, J. P., & Gregory, J. (2005). Basket Default Swaps, CDOs and Factor Copulas. *Journal of Risk, 7*(4), 103–122.

Lauret, J., Walker, M., Goasguen, S., & Hajdu, L. (2010). From Grid to cloud, the STAR experience. SciDAC 2010 Proceedings.

Lawless, J. F. (2003). *Statistical Models and Methods For Lifetime Data*. New York: John Wiley and Sons.

Lawley, D. N. (1940). The Estimation of Factor Loadings by the Method of Maximum Likelihood. *Proceedings of the Royal Society of Edinburg, Series A, 62*, 64–82.

Lawley, D. N., & Maxwell, A. E. (1963). *Factor Analysis as a Statistical Method*. London: Butterworth.

Lawrence, E. C., & Arshadi, N. (1995). A multinomial logit analysis of problem loan resolution choices in banking. *Journal of Money, Credit and Banking, 27*, 202–216.

Lawrence, K. D., & Reeves, G. (1982). Combining multiple forecasts given multiple objectives. *Journal of Forecasting, 1*, 271–279.

Lawrence, K. D., Klimberg, R. K., & Lawrence, S. M. (2009). *Fundamentals of Forecasting Using Excel*. New York: Industrial Press.

Lawrence, K. D., & Reeves, G. (1981). Consensus time series forecasting, Lecture Notes in Economics and Mathematical Systems, Edited by J. Morse, 190, Springer-Verlag, Berlin, 199–204.

Leamer, E. E. (1978). Specification Searches: Ad Hoc Inference With Nonexperimental Data. Wiley.

Lean, H. H., Phoon, K. F., & Wong, W. K. (2012). Stochastic Dominance Analysis of CTA Funds. *Review of Quantitative Finance and Accounting*. doi:10.1007/s11156-012-0284-1.

LeBaron, B. (2000). Technical trading profitability in foreign exchange markets in the 1990's. Technical report, Brandeis U.

Lecraw, D. J. (1985). Some evidence on transfer pricing by multinational corporations. In A. M. Rugman & L. Eden (Eds.), *Multinationals and transfer pricing*. New York: St. Martin's Press.

Ledoit, O., & Wolf, M. (2008). Robust performance hypothesis testing with the Sharpe ratio. *Journal of Empirical Finance, 15*, 850–859.

Lee, C., Lee, A., & Lee, J. (2000a). Statistics for Business and Financial Economics, World Scientific Publishers, Inc., Second edition.

Lee, A. (1996). Cost of capital and equity offerings in the insurance industry. Ph.D. Dissertation, The University of Pennsylvania in Partial.

Lee, A. C., & Cummins, J. D. (1998). Alternative models for estimating the cost of capital for property/casualty insurers. *Review of Quantitative Finance and Accounting, 10*(3), 235–267.

Lee, B.-S. (1996b). Time-series implications of aggregate dividend behavior. *Review of Financial Studies, 9*, 589–618.

Lee, B.-S. (1998). Permanent, temporary and nonfundamental components of stock prices. *Journal of Financial and Quantitative Analysis, 33*, 1–32.

Lee, C., & Ready, M. (1991). Inferring trade direction from intraday data. *Journal of Finance, 46*, 733–747.

Lee, C. F. (1973). Errors-in-variables estimation procedures with applications to a capital asset pricing model. Ph.D. Dissertation, The State University of New York at Buffalo.

Lee, C. F. (1976a). A note on the interdependent structure of security returns. *Journal of Financial and Quantitative Analysis, 11*, 73–86.

Lee, C. F. (1976b). Functional form and the dividend effect of the electric utility industry. *Journal of Finance, 31*(5), 1481–1486.

Lee, C. F. (1977a). Functional form, skewness effect and the risk-return relationship. *Journal of Financial and Quantitative Analysis, 12*, 55–72.

Lee, C. F. (1977b). Performance measure, systematic risk and errors-in-variable estimation method. Journal of Economics and Business, 122–127.

Lee, C. F. (1993). *Statistics for business and financial economics*. Lexington, MA: Heath.

Lee, C. F., & Su, J. B. (2011). Alternative statistical distributions for estimating Value-at-Risk: theory and evidence. *Review of Quantitative Finance and Accounting*. doi:http://dx.doi.org/10.1007/s11156-011-0256-x.

Lee, C. F., & Rahman, S. (1990). Market timing, selectivity, and mutual fund performance: An empirical investigation. *Journal of Business, 63*(2), 261–278.

Lee, C. F., & Gweon, S. C. (1986). Rational Expectation, Supply Effect and Stock Price Adjustment. Paper presented at 1986, Econometrica Society annual meeting.

Lee, C. F., & Wu, C. C. (1985). The impacts of kurtosis on risk stationarity: some empirical evidence. *Financial Review, 20*(4), 263–269.

Lee, C. F., & Jen, F. C. (1978). Effects of measurement errors on systematic risk and performance measure of a portfolio. *Journal of Financial and Quantitative Analysis, 13*(2), 299–312.

Lee, C. F., & Chen, H. Y. (2012). Alternative errors-in-variable models and their application in finance research. Working paper, Rutgers University.

Lee, C. F., & Zumwalt, J. K. (1981). Associations between alternative accounting profitability measures and security returns. *Journal of Financial and Quantitative Analysis, 16*, 71–93.

Lee, C. F., Chien, C. C., Lin, H. C., Lin, Y. C., & Lin, F. C. (2012). Tradeoff between reputation concerns and economic dependence for auditors- Threshold regression approach. Working paper, Rutgers University.

Lee, C. F., Wu, C. C., & Wei, K. C. J. (1990a). Heterogeneous investment horizon and capital asset pricing model: Theory and implications. *Journal of Financial and Quantitative Analysis, 25*, 361–376.

Lee, C. F., Tsai, C. M., & Lee, A. C. (2011). Asset pricing with disequilibrium price adjustment: Theory and empirical evidence. Quantitative Finance, forthcoming.

Lee, C. F., Wei, K. C. J., & Bubnys, E. L. (1989). The APT versus the multi-factor CAPM: Empirical evidence. Quarterly Review of Economics and Business, Vol. 29.

Lee, C. F., Gupta, M. C., Chen, H. Y., & Lee, A. C. (2011b). Optimal payout ratio under uncertainty and the flexibility hypothesis: Theory and empirical evidence. *Journal of Corporate Finance, 17*(3), 483–501.

Lee, C. F., Newbold, P., Finnerty, J. E., & Chu, C. C. (1986a). On accounting-based, market-based and composite-based beta predictions: Method and implications. *Financial Review, 21*, 51–68.

Lee, C. M. C., Myers, J., & Swaminathan, B. (1999a). What is the intrinsic value of the Dow? *Journal of Finance, 54*(5), 1693–1741.

Lee, C., Wu, C., & Djarraya, M. (1987a). A further empirical investigation of the dividend adjustment process. *Journal of Econometrics, 35*, 267–285.

Lee, C. F., & Wu, C. (1982). The Impacts of Skewness and Kurtosis on Risk Estimation and Determination.

Lee, C. F., & Wu, C. (1985b). The Impacts of Kurtosis on Risk Stationarity: Some Empirical Evidence. *Financial Review, 20*, 263–269.

Lee, C. F., Lee, J. C., & Lee, A. C. (2000b). *Statistics for Business and Financial Economics*. New Jersey: World Scientific.

Lee, C. F., Lee, K. W., & Yeo, H. H. (2009). Investor Protection and Convertible Debt Design. *Journal Of Banking and Finance, 33*(6), 985–995.

Lee, C. F., & Chen, H. Y. (2012). Alternative Errors-in-Variable Models and Their Applications in Finance Research. Rutgers working paper.

Lee, C. M. C., Shleifer, A., & Thaler, R. (1991). Investor Sentiment and the Closed-End Fund Puzzle. *Journal of Finance, 46*(1), 75–109.

Lee, C. F. (2009). *Handbook of Quantitative Finance*. New York: Springer.

Lee, C. F., & Forbes, S. W (1980). Dividend Policy, Equity Value, and Cost of Capital Estimates for the Property and Liability Insurance Industry. *Journal of Risk and Insurance* 47, 205–222 (June).

Lee, C. F., & Lee, A. C. (2006). *Encyclopedia of Finance*. New York: Springer.

Lee, Cheng F., & Shi, J. (2010). Application of alternative ODE in finance and econometric research, in Cheng F. Lee, Alice. C. Lee, and John Lee, ed.: Handbook of Quantitative Finance and Risk Management (Springer), 1293–1300.

Lee, C. F., Wu, C., & Djarraya, M. (1987). A further empirical investigation of the dividend adjustment process. *Journal of Econometrics, 35*, 267–285.

Lee, C. F., Newbold, P., Finnerty, J. E., & Chu, C.-C. (1986). On Accounting-Based, Market Based and Composite-Based Beta Predictions: Method and Implications. *Financial Review, 21*(1), 51–68.

Lee, Cheng Few., & Kau, James B. (1976). The Functional Form in Estimating the Density Gradient: An Empirical Investigation. *Journal of American Statistical Association, June*.

Lee, Cheng Few. (1976). Functional Form and the Dividend Effect of the Electric Utility Industry. *Journal of Finance*, December.

Lee, Cheng Few., Fabozzi Frank J., & Francis, Jack C. (1980). Generalized Functional Form for Mutual Fund Returns. *Journal of Financial and Quantitative Analysis*, December.

Lee, E. S., & Li, R. J. (1993). Fuzzy multiple objective programming and compromise programming with Pareto optimum. *Fuzzy Sets and Systems, 53*, 275–288.

Lee, E. T. (1980). *Statistical Methods for Survival Data*. Belmont, California: Wadsworth, Inc.

Lee, H. (1993b). A structured methodology for software development effort prediction using the analytic hierarchy process. *Journal of Systems and Software, 21*, 179–186.

Lee, H., Nazem, S. M., & Shi, Y. (1995). Designing rural area telecommunication networks via hub cities. *Omega: International Journal of Management Science, 22*(3), 305–314.

Lee, H.-T., & Yoder, J. K. (2007a). A bivariate Markov regime switching GARCH approach to estimate time varying minimum variance hedge ratios. *Applied Economics, 39*, 1253–1265.

Lee, H.-T., & Yoder, J. K. (2007b). Optimal hedging with a Markov regime switching time-varying correlation. *Journal of Futures Markets, 27*, 495–516.

Lee, H.-T. (2010). Regime switching correlation hedging. *Journal of Banking and Finance, 34*, 2728–2741.

Lee, I. (1997). Do firms knowingly sell overvalued equity? *Journal of Finance, 52*, 1439–1466.

Lee, Jack C., Lee, C. F., Wang, R. S., & Lin, T. I. (2004). On the Limit Properties of Binomial and Multinomial Option Pricing Models: Review and Integration. *Advances in Quantitative Analysis of Finance and Accounting New Series volume 1*. Singapore: World Scientific Vol. 1.

Lee, John C. (2001). Using Microsoft Excel and Decision trees to Demonstrate the Binomial Option Pricing Model. *Advances in Investment Analysis and Portfolio Management, 8*, 303–329.

Lee, K. W., & Lee, C. F. (2012). Multiple directorships, corporate ownership, firm performance. Working Paper, Rutgers University.

Lee, K. W. (2007). Corporate Voluntary Disclosure and the Separation of Cash Flow Rights from Control Rights. *Review of Quantitative Finance and Accounting, 28*, 393–416.

Lee, K. W., Lev, B., & Yeo, H. H. (2007). Organizational Structure and Earnings Management. *Journal of Accounting, Auditing and Finance, 22*(2), 293–391.

Lee, K. W., Lev, B., & Yeo, H. H. (2008a). Executive Pay Dispersion, Corporate Governance and Firm Performance. *Review of Quantitative Finance and Accounting, 30*, 315–338.

Lee, L.-F. (1978). Unionism and wage rates: a simultaneous equations model with qualitative and limited dependent variables. *International Economic Review, 19*, 415–433.

Lee, M. C., Su, J. B., & Liu, H. C. (2008b). Value-at-risk in US stock indices with skewed generalized error distribution. *Applied Financial Economics Letters, 4*, 425–431.

Lee, O., & Shin, D. W. (2008). Geometric ergodicity and β-mixing property for a multivariate CARR model. *Economics Letters, 100*, 111–114.

Lee, P. M., & Wahal, S. (2004). Grandstanding, certification and the underpricing of venture capital backed IPOs. *Journal of Financial Economics, 73*, 375–407.

Lee, R. (2005). Option pricing by transform methods: extensions, unification, and error control. *Journal of Computational Finance, 7*(3), 51–86.

Lee, S. M., & Lerro, A. (1974). Capital budgeting for multiple objectives. *Financial Management, 3*, 51–53.

Lee, T. (2003). Smoothing parameter selection for smoothing splines: a simulation study. *Computational statistics & Data analysis, 42*(1–2), 139–148.

Lee, T. S., & Oh, Y. L. (1995). Foreign investment, stock market volatility and macro variables, (in Chinese). *Fundamental Financial Institutions, 31*, 45–47.

Lee, T.-H., & Long, X. (2009). Copula-based Multivariate GARCH Model with Uncorrelated Dependent Errors. *Journal of Econometrics, 150*, 207–218.

Lee, Y. R., Shi, Y., & Yu, P. L. (1990). Linear optimal designs and optimal contingency plans. *Management Science, 36*, 1106–1119.

Lee, Y.-T., Lin, J. C., & Liu, Y. J. (1999). Trading patterns of big versus small players in an emerging market: An empirical analysis. *Journal of Banking and Finance, 23*, 701–725.

Lehman, B. (1993). Earnings, Dividend Policy, and Present Value Relations: Building Blocks of Dividend Policy Invariant Cash Flows. *Review of Quantitative Finance and Accounting, 3*, 263–282.

Lehmann, B. (1990). Fads, martingales, and market efficiency. *Quarterly Journal of Economics, 105*, 1–28.

Lehmann, B. N., & Modest, D. M. (1987). Mutual fund performance evaluation: A comparison of benchmarks and benchmark comparisons. *Journal of Finance, 42*(2), 233–265.

Lehmann, E. (1955). Ordered families of distributions. *Annals of Mathematical Statistics, 26*, 399–419.

Lehn, K., & Poulsen, J. (1989). Free Cash Flow and Stockholders Gain in Going Private Transactions. *Journal of Finance, 44*, 771–88.

Lei, Q., & Wu, G. (2005). Time-varying informed and uninformed trading activities. *Journal of Financial Markets, 8*, 153–181.

Lei, V., Noussair, C., & Plott, C. (2001). Nonspeculative Bubbles in Experimental Asset Markets: Lack of Common Knowledge of Rationality vs. Actual Irrationality. *Econometrica, 69*(4), 831–859.

Leippold, M., & Wu, L. (2002). Asset Pricing Under the Quadratic Class. *Journal of Financial and Quantitative Analysis, 37*, 271–295.

Leitch, G., & Tanner, J. E. (1991). Economic forecast evaluation: Profits versus the conventional error measures. *The American Economic Review, 81*, 580–590.

Leland, H. E. (1994). Corporate Debt Value.

Leland, H. E., & Toft, K. (1996). Optimal Capital Structure, Endogeneous Bankruptcy, and the Term Structure of Credit Spreads. *Journal of Finance, 51*, 987–1019.

Lendasse, A., DeBodt, E., Wentz, & Verlegsen, M. (2000). Nonlinear financial time series forecasting-application to the Belgian 20 stock market index. *European Journal of Economical and Social Systems 14*(1), 81–91

Leon, A., & Vaello-Sebastia, A. (2009). American GARCH employee stock option valuation. *Journal of Banking and Finance, 33*(6), 1129–1143.

Leone, A., & Wu, J. (2002). What does it take to become a superstar? Evidence from institutional investor rankings of financial analysts. Working Paper, University of Rochester.

Lerner, E. M., & Carleton, W. T. (1966a). Financing decisions of the firm. *Journal of Finance, 21*, 202–214.

Lerner, E. M., & Carleton, W. T. (1966b). *A Theory of Financial Analysis*. New York, NY: Harcourt, Brace & World.

LeRoy, S. F., & Porter, R. D. (1981). The present value relation: Tests based on implied variance bounds. *Econometrica, 49*, 555–574.

Leser, C. E. V. (1961). A simple method of trend construction. *Journal of the Royal Statistical Society. Series B (Methodological), 23*(1), 91–107.

Lesmond, D. A., Ogden, J. P., et al. (1999). A new estimate of transaction costs. *Review of Financial Studies, 12*(5), 1113–1141.

Leto, A., et al. (2005). EGRID Project: Experience Report on the Implementation of a Grid Infrastructure for the Analysis of Financial Data. In Proceedings of the International Conference on Information Technology: Coding and Computing.

Lettau, M., & Ludvigson, S. C. (2004). Measuring and modeling variation in the risk-return tradeoff. In Y. Aït-Sahalia, & L. P. Hansen (Eds.). Amsterdam: North Holland.

Leung, P. L., & Wong, W. K. (2008). On testing the equality of the multiple Sharpe Ratios, with application on the evaluation of iShares. *Journal of Risk, 10*, 1–16.

Leuz, C., Nanda, D., & Wysocki, P. (2003). Earnings management and investor protection: An international comparison. *Journal of Financial Economics, 69*, 505–527.

Lev, B. (2001). *Intangibles*. Washington, D.C.: Brookings Institution Press.

Lev, B., & Ohlson, J.A. (1982). Market Based Empirical Research in Accounting: A Review, Interpretations, and Extensions. *Supplement to Journal of Accounting Research 20*, 249–322.

Lev, B. (1989). On the Usefulness of Earnings, Lessons and Directions from Two Decades of Empirical Research. *Journal of Accounting Research 27*, 153–92.

Levy, H. (1985). Upper and lower bounds of put and call option value: stochastic dominance approach. *Journal of Finance, 40*, 1197–1217.

Levy, H. (2010). The CAPM is alive and well: A review and synthesis. *European Financial Management, 16*(1), 43–71.

Levy, H. (1994). Absolute and Relative Risk Aversion: An Experimental Study. *Journal of Risk and Uncertainty, 8*, 289–307.

Levy, H., & Sarnat, M. (1970). Diversification, Portfolio Analysis and the Uneasy Case for Conglomerate Mergers. *The Journal of Finance, 25*(1970), 795–802.

Levy, H. (2006). *Stochastic Dominance: Investment Decision Making Under Uncertainty* (2nd ed.). New York: Springer.

Lewellen, W. G. (1991). A Pure Financial Rationable for conglomerate Merger. *The Journal of Finance, 27*, 521–537.

Lewis, A. (2000). Option valuation under stochastic volatility. Finance Press.

Lewis, C., Rogalski, R., & Seward, J. K. (1998). Industry conditions, growth opportunities and the market reactions to convertible debt financing decisions. Unpublished mimeo, Vanderbilt University.

Leybourne, S., Taylor, R., & Kim, T. (2006). Cusum of squares-based tests for a change in persistence. *Journal of Time Series Analysis, 28*, 408–433.

Li, C. K., & Wong, W. K. (1999). Extension of stochastic dominance theory to random variables. *RAIRO Recherche Opérationnelle, 33*, 509–524.

Li, D., & He, Y. (2008). Study on the price gap between on-the-run and off-the-run Treasury securities of Shanghai Stock Exchange. *Journal of Shanxi Finance and Economics University (In Chinese), 30*, 80–85.

Li, D. X. (1999). The Valuation of Basket Credit Derivatives. CreditMetrics Monitor April, 34–50.

Li, D. X. (2000a). On default correlation: A copula function approach. *Journal of Fixed Income, 9*, 91–54.

Li, D. X. (2000b). On Default Correlation: A Copula Function Approach. *Journal of Fixed Income, 9*, 43–54.

Li, D. X. (1999). On Default Correlation: A copula function approach. Working Paper 99-07, The Risk Metrics Group.

Li, H., & Hong, Y. (2011). Financial volatility forecasting with range-based autoregressive volatility model. *Finance Research Letters, 8*, 69–76.

Li, H., & Zhao, F. (2006). Unspanned stochastic volatility: Evidence from hedging interest rate derivatives. *Journal of Finance, 61*, 341–378.

Li, H., & Zhao, F. (2009). Nonparametric Estimation of State-Price Densities Implicit in Interest Rate Cap Prices. *Review of Financial Studies, 22*(11), 4335–4376.

Li, K., & Prabhala, N. (2007). Self-selection models in corporate finance, Robert H. Smith School Research Paper No. RHS 06-020.

Li, K. C. (1991). Sliced inverse regression for dimension reduction. *Journal of the American Statistical Association, 86*, 285–299.

Li, L. (2007). Sparse sufficient dimension reduction. *Biometrika, 94*, 603–613.

Li, R., Zhong, W., & Zhu, L. P. (2012). Feature screening via distance correlation learning. *Journal of the American Statistical Association*. forthcoming.

Li, S. (2005). A New Formula for Computing Implied Volatility. *Applied Mathematics and Computation, 170*, 611–625.

Li, Y., Huang, M., Chin, K., Luo, X., & Han, Y. (2011). Integrating preference analysis and balanced scorecard to product planning house of quality. *Computers & Industrial Engineering, 60*, 256–268.

Liao, F. Y., & Lyuu, Y. D. (2009). An expanded model for the valuation of employee stock options. *Journal of Futures Markets, 29*, 713–735.

Liao, X., Zhang, Y., & Sun, C. (2009). Investments in timberland and softwood timber as parts of portfolio selection in the United States: A cointegration analysis and capital asset pricing model. *Forest Science, 55*, 471–479.

Libby, R., Hunton, J., Tan, H., & Seybert, N. (2008). Relationship incentives and the optimistic/pessimistic pattern in analysts' forecasts. *Journal of Accounting Research, 46*, 173–198.

Liberatore, M. J., Mohnhan, T. F., & Stout, D. E. (1992). A framework for integrating capital budgeting analysis with strategy. *The Engineering Economist, 38*, 31–43.

Lie, E. (2000). Excess funds and agency problems: An empirical study of incremental cash disbursements. *Review of Financial Studies, 13*, 219–248.

Lien, D. (2010a). An note on the relationship between the variability of the hedge ratio and hedging performance. *Journal of Futures Market, 30*(11), 1100–1104.

Lien, D., & Shrestha, K. (2007). An empirical analysis of the relationship between hedge ratio and hedging horizon using wavelet analysis. *Journal of Futures Market, 27*(2), 127–150.

Lien, D. (1996). The effect of cointegration relationship on futures hedging: a note. *Journal of Futures Markets, 16*, 773–780.

Lien, D. (2004). Cointegration and the optimal hedge ratio: the general case. *Quarterly Review of Economics and Finance, 44*, 654–658.

Lien, D. (2005a). A note on the superiority of the OLS hedge ratio. *Journal of Futures Markets, 25*, 1121–1126.

Lien, D. (2005b). The use and abuse of the hedging effectiveness measure. *International Review of Financial Analysis, 14*, 277–282.

Lien, D. (2006). Estimation bias of futures hedging performance: a note. *Journal of Futures Markets, 26*, 835–841.

Lien, D. (2008). A note on estimating the benefit of a composite hedge. *Journal of Futures Markets, 28*, 711–716.

Lien, D. (2009). A note on the hedging effectiveness of GARCH models. *International Review of Economics and Finance, 18*, 110–112.

Lien, D. (2010b). A note on the relationship between the variability of the hedge ratio and hedging performance. *Journal of Futures Markets, 30*, 1100–1104.

Lien, D. (2012). A note on the performance of regime switching hedge strategy. *Journal of Futures Markets, 32*, 389–396.

Lien, D., & Tse, Y. K. (1998). Hedging time-varying downside risk. *Journal of Futures Markets, 18*, 705–722.

Lien, D., & Tse, Y. K. (2000). Hedging downside risk with futures contracts. *Applied Financial Economics, 10*, 163–170.

Lien, D., & Tse, Y. K. (2001). Hedging downside risk: futures vs. options. *International Review of Economics and Finance, 10*, 159–169.

Lillo, F., Moro, E., Vaglica, G., & Mantegna, R. (2008). Specialization of strategies and herding behavior of trading firms in a financial market. *New Journal of Physics 10*, 043019 (15pp).

Lim, T. (2001). Rationality and analysts' forecast bias. *Journal of Finance, 56*, 369–385.

Lin, A. Y., & Swanson, P. E. (2003). The Behavior and Performance Foreign Investors in Emerging Equity Markets: Evidence from Taiwan. *International Review of Finance, 4*(3–4), 189–210.

Lin, B., & Yang, R. (2010). Does Regulation Fair Disclosure affect analysts' forecast performance? The case of restructuring firms. *Review of Quantitative Finance and Accounting, 38*, 495–517.

Lin, E., Lee, C. F., & Wang, K. (2012). Futures mispricing, order imbalance, and short-selling constraints. *International Review of Economics and Finance*, forthcoming.

Lin, E. M. H., Chen, C. W. S., & Gerlach, R. (2012). Forecasting volatility with asymmetric smooth transition dynamic range models. *International Journal of Forecasting, 28*, 384–399.

Lin, H., Liu, S., & Wu, C. (2011). Dissecting corporate bond and CDS spreads. *The Journal of Fixed Income, 20*, 7–39.

Lin, J., & Rozeff, M. S. (1994). Variance, return, and high-low price spreads. *Journal of Financial Research, 17*, 301–319.

Lin, L., Lefebvre, C., & Kantor, J. (1993). Economic determinants of international transfer pricing and the related accounting issues, with particular reference to Asian Pacific countries. *The International Journal of Accounting*, 49–70.

Lin, S. (2006). Testing the Information Set Perspective of UK Financial Reporting Standard No. 3: Reporting Financial Performance, *Journal of Business Finance & Accounting, 33*(7) & (8), September/October, 1110–1141.

Lin, T. W. (1993). Multiple-criteria capital budgeting under risk. *Advances in Mathematical Programming and Financial Planning, 3*, 231–239.

Linsmeier, T., Gribble, J., Jennings, R., Lang, M., Penman, S., Petroni, K., Shores, D., Smith, J., & Warfield, T. (1997). An issues paper on comprehensive income. *Accounting Horizons, 11*, 120–126.

Linsmeier, T., Gribble, J., Jennings, R., Lang, M., Penman, S., Petroni, K., Shores, D., Smith, J., & Warfield, T. (1997a). Response to FASB exposure draft, "Proposed statement of financial accounting standards – Reporting comprehensive income". *Accounting Horizons, 11*, 117–119.

Lintner, J. (1956). Distributions of incomes of corporations among dividends, retained earnings, and taxes. *American Economic Review, 46*, 97–113.

Lintner, J. (1965a). The valuation of risk assets and the selection of risky investments in stock portfolios and capital budgets. *The Review of Economics and Statistics, 47*, 13–37.

Lintner, J. (1962). Dividends, earnings, leverage, stock prices and the supply of capital to corporations. *Review of Economics and Statistics, 44*, 243–269.

Lintner, J. (1963). The cost of capital and optimal financing of corporate growth. *Journal of Finance, 18*, 292–310.

Lintner, J. (1964). Optimal dividends and corporate growth under uncertainty. *Quarterly Journal of Economics, 78*, 49–95.

Lintner, J. (1965b). Security prices, risk, and maximal gains from diversification. *Journal of Finance, 20*, 587–615.

Lipe, R. (1986). The Information Contained in the Components of Earnings. *Journal of Accounting Research, 24*, 37–64.

Lipe, R. (1990). The Relation Between Stock Returns and Accounting Earnings Given Alternative Information. *The Accounting Review, 65*, 49–71.

Lipson, M. L. (2003). Market microstructure and corporate finance. *Journal of Corporate Finance, 9*(4), 377–384.

Lipson, M. L., & Mortal, S. (2007). Liquidity and firm characteristics: evidence from mergers and acquisitions. *Journal of Financial Markets, 10*(4), 342–361.

Lipson, M. L., & Mortal, S. (2009). Liquidity and capital structure. *Journal of Financial Markets, 12*(4), 611–644.

Litterman, R., & Scheinkman, J. (1991). Common Factors Affecting Bond Returns. *Journal of Fixed Income, 1*, 62–74.

Litzenberger, R. H., & Ramaswamy, K. (1979a). The effects of personal taxes and dividends on capital asset prices: Theory and empirical evidence. *Journal of Financial Economics, 7*, 163–195.

Litzenberger, R. H., & Ramaswamy, K. (1979b). The Effects of Personal Taxes and Dividends on Capital Asset Prices: Theory and Empirical Evidence. *Journal of Financial Economics, 7*, 163–196.

Litzkow, M. J., Livny, M., & Mutka, M. W. (1988). Condor-a hunter of idle workstations. 8th International Conference on Distributed Computing Systems, Proceeding, 104–111.

Liu, B. (2006). Two essays in financial econometrics: I: Functional forms and pricing of country funds. II: The term structure model of inflation risk premia. Ph.D. Dissertation, Rutgers University.

Liu, C., Yao, L. J., Hu, N., & Liu, L. (2011). The impact of IFRS on accounting quality in a regulated market: an empirical study of China. *Journal of Accounting, Auditing and Finance, 26*, 659–676.

Liu, C. (2011). IFRS and US-GAAP comparability before release No. 33-8879: some evidence from US-listed Chinese companies. *International Journal of Accounting and Information Management, 19*, 24–33.

Liu, C., Yao, L. J., & Yao, M. (2012a). Value relevance change under international accounting standards: an empirical study of Peru. *Review of Pacific Basin Financial Markets and Policies, 15*, 1–17.

Liu, H., & Hung, J. (2010). Forecasting volatility and capturing downside risk of the Taiwanese futures markets under the financial tsunami. *Managerial Finance, 36*, 860–875.

Liu, H., Chiang, S., & Cheng, N. (2012b). Forecasting the volatility of S&P depositary receipts using GARCH-type models under intraday range-based and return-based proxy measures. *International Review of Economics and Finance, 22*, 78–91.

Liu, J., & Ohlson, J. A. (2000). The Feltham-Ohlson (1995) Model: Empirical Implications. *Journal of Accounting, Auditing and Finance, 15*, 321–331.

Liu, J., Nissim, D., & Thomas, J. (2002). Equity Valuation using Multiples. *Journal of Accounting Research, 40*(1), 135–172.

Liu, P., & Maddala, G. S. (1992). Rationality of survey data and tests for market efficiency in the foreign exchange markets. *Journal of International Money and Finance, 11*, 366–381.

Liu, R. Y. (1988). Bootstrap procedures under some non-IID models. *Annals of Statistics, 16*, 1696–1708.

Liu, S. M., & Brorsen, B. (1995). Maximum Likelihood Estimation of a GARCH – STABLE Model. *Journal of Applied Econometrics, 10*, 273–285.

Liu, W. (2006b). A liquidity-augmented capital asset pricing model". *Journal of Financial Economics, 82*, 631–671.

Livingston, M. (1987). The delivery option on forward contracts. *Journal of Financial and Quantitative Analysis, 22*(01), 79–87.

Ljung, G., & Box, G. (1978). On a measure of lack of fit in time series models. *Biometrica, 66*, 265–270.

Ljungqvist, A., Marston, F., & Wilhelm, W. J. (2006). Competing for securities underwriting mandates: banking relationships and analyst recommendations. *Journal of Finance, 61*, 301–340.

Llyod, W. P., & Shick, R. (1977). A Test of Stone's Two Index Models of Returns. *Journal of Financial and Quantitative Analysis, 12*, 363–376.

Lo, A. (1987a). Semi-parametric upper bounds for option prices and expected payoff". *Journal of Financial Economics, 19*, 373–388.

Lo, A. (2002). The statistics of Sharpe ratios. *Financial Analysis Journal, 58*, 36–52.

Lo, K., & Lys, T. (2000). The Ohlson Model: Contribution to valuation theory, limitations, and empirical applications. *Journal of Accounting, Auditing and Finance, 15*(3), 337–367.

Lo, A., & Mackinlay, A. C. (1990). "An Econometric Analysis of Nonsynchronous Trading." *Journal of Econometrics 45*, 181–211.

Lo, A. W., & MacKinlay, A. C. (1990a). When are contrarian profits due to stock market overreaction? *Review of Financial Studies, 3*, 175–205.

Lo, A., & MacKinlay, A. C. (1990b). Data-Snooping Biases in Tests of Financial Asset Pricing Models. *Review of Financial Studies, 3*, 431–467.

Lo, A. W., & Wang, J. (2000). Trading Volume: Definition, Data Analysis, and Implications of Portfolio Theory. *Review of Financial Studies, 13*, 257–300.

Lo, A. W. (1987). Semi-parametric upper bounds for options prices and expected payoffs. *Journal of Financial Economics, 19*(2), 373–387.

Lobao, J., Serra, A., & Serra, P. (2007). Herding Behavior evidence from Portuguese Mutual Funds. In G. Gregoriou (Ed.), *Diversification and portfolio management of mutual funds* (pp. 167–197). New York: Palgrave MacMillan.

Lobato, I., Nankervis, J. C., & Savin, N. (2001). Testing for Autocorrelation Using a Modified Box-Pierce Q Test. *International Economic Review, 42*(1), 187–205.

Lobo, G., Song, M., & Stanford, M. (2012). Accruals quality and analyst coverage. *Journal of Banking and Finance, 36*, 497–508.

Locke, E. A., Shaw, K. N., & Lathom, G. P. (1981). Goal setting and task performance: 1969–1980. *Psychological Bulletin*, 125–152.

Loffler, G. (2012 April.). Can rating agencies look through the cycle? *Review of Quantitative Finance and Accounting*, 1–24. http://dx.doi.org/10.1007/s11156-012-0289-9.

Loh, R., & Mian, G. (2006). Do accurate earnings forecasts facilitate superior investment recommendations? *Journal of Financial Economics, 80*, 455–483.

Longin, F. (1996). The Asymptotic Distribution of Extreme Stock Market Returns. *Journal of Business, 63*, 383–408.

Longin, F., & Solnik, B. (1995). Is the correlation in international equity returns constant: 1960–1990? *Journal of International Money and Finance, 14*, 3–26.

Longin, F., & Solnik, B. (2001). Extreme correlations in international equity markets. *Journal of Finance, 56*, 649–676.

Longstaff, F., & Schwartz, E. (1992). Interest rate volatility and the term structure: A two-factor general equilibrium model. *Journal of Finance, 47*(4), 1259–1289.

Longstaff, F. (1995). Option Pricing and the Martingale Restriction. *Review of Financial Studies, 8*(4), 1091–1124.

Longstaff, F., Santa-Clara, P., & Schwartz, E. (2001). The Relative Valuation of Caps and Swaptions: Theory and Evidence. *Journal of Finance, 56*, 2067–2109.

Longstaff, F., Mithal, S., & Neis, E. (2005). Corporate yield spreads: Default risk or liquidity? New evidence from the credit default swap market. *The Journal of Finance, 60*, 2213–2253.

Lopez, J. A. (1999). Methods for evaluating value-at-risk estimates. *Federal Reserve Bank of San Francisco Economic Review, 2*, 3–17.

Lopez, J. (1998). Methods for evaluating value-at-risk estimates. *Economic Policy Review, 4*. Federal Reserve Bank of New York.

Lord, R., & Kahl, C. (2010). Complex logarithms in Heston-like models. *Mathematical Finance, 20*(4), 671–694.

Lorenz, M. O. (1905). Methods of measuring concentration of wealth. *Journal of the American Statistical Association, 9*, 209–219.

Lorie, J. H., & Savage, L. I. (1955). Three problems in rationing capital. *Journal of Business, 28*, 229–239.

Loughran, T., & Ritter, J. (2004). Why has IPO underpricing changed over time? *Financial Management, 33*(3), 5–37.

Loughran, T., & Ritter, J. R. (1995). The new issues puzzle. *Journal of Finance, 50*, 23–52.

Loughran, T., & Ritter, J. R. (2000). Uniformly least powerful tests of the market efficiency. *Journal of Financial Economics, 55*, 361–389.

Louzis, D. P., Xanthopoulos-Sisinis, S., & Refenes, A. N. (2012). The role of high frequency intra-daily data, daily range and implied volatility in multi-period value-at-risk forecasting, Working paper.

Lowry, M., & Shu, S. (2002). Litigation risk and IPO underpricing. *Journal of Financial Economics, 65*, 309–335.

Lu, C., Wu, S. C., & Ho, L. C. (2009). Applying VaR to REITs: A comparison of alternative methods. *Review of Financial Economics, 18*, 97–102.

Lucas, A., & Klaassen, P. (1998). Extreme Returns, Downside Risk, and Optimal Asset Allocation. *Journal of Portfolio Management, 25*(1), 71–79.

Lucas, D., & Lonski, J. (1992). Changes in corporate credit quality 1970–1990. *Journal of Fixed Income, 1*, 7–14.

Lucas, D. (1995). Default correlation and credit analysis. *Journal of Fixed Income, 9*(4), 76–87.

Lucas, R. E., Jr. (1978). Asset prices in an exchange economy. *Econometrica, 46*, 1429–1445.

Lucia, J. J., & Schwartz, E. S. (2002). Electricity prices and power derivatives: Evidence from the nordic power exchange. *Review of Derivatives Research, 5*(1), 5–50.

Lucke, B. (2010). Identification and overidentification in SVECMs. *Economics Letters, 108*(3), 318–321.

Ludvigson, S. C., & Ng, S. (2007). The empirical risk-return relation: A factor analysis approach. *Journal of Financial Economics, 83*, 171–222.

Lundblad, C. (2007). The risk return tradeoff in the long run: 1836–2003. *Journal of Financial Economics, 85*, 123–150.

Lundholm, R., & Myers, L. (2002). Bringing the Future Forward: The Effect of Disclosure on the Returns-Earnings Relation. *Journal of Accounting Research, 40*, 809–839.

Lutkepohl, H. (1989). Prediction tests for structural stability of multiple time series. *Journal of Business Economic Statistics, 7*, 129–135.

Lutkepohl, H. (1991). *Introduction to Multiple Time Series Analysis*. Berlin: Springer.

Lutkepohl, H., (2005). Structural Vector Autoregressive Analysis for Cointegrated Variables. EUI Working paper ECO.

Lütkepohl, H., Benkwitz, A., & Wolters, J. (1998). A money demand system for German M3. *Empirical Economics, 23*, 382–384.

Lütkepohl, H., Benkwitz, A., & Wolters, J. (2001). Comparison of bootstrap confidence intervals for impulse responses of German monetary systems. *Macroeconomic Dynamics, 5*, 81–100.

Lütkepol, H. (1993). *Introduction to Multiple Time Series Analysis*. New York: Springer.

Lyandres, E. (2007). External Financing Costs, Investment Timing, and Investment-Cash Flow Sensitivity. *Journal of Corporate Finance, 13*, 959–980.

Lynch, P., & Zumbach, G. (2003). Market heterogeneities and the causal structure of volatility. *Quantitative Finance, 3*, 320–331.

Lynge, M., & Zumwalt, J. (1980). An Empirical Study of Interest Rate Sensitivity of Commercial Bank Returns: A Multi Index Approach. *Journal of Financial and Quantitative Analysis, 15*, 731–742.

Lyon, J. D., Barber, B. M., & Tsai, C.-L. (1999). Improved methods for tests of long-run abnormal stock returns. *Journal of Finance, 54*, 165–201.

Lyons, R. K. (2002). Foreign exchange: Macro puzzles, micro tools. *Economic Review*, Federal Reserve Bank of San Francisco.

Lys, T., & Soo, L. (1995). Analysts' forecast precision as a response to competition. *Journal of Accounting, Auditing and Finance, 10*, 751–765.

Ma, C., & Wong, W. K. (2010). Stochastic dominance and risk measure: a decision theoretic foundation for VaR and C-VaR. *European Journal of Operational Research, 207*, 927–935.

Ma, L., & Koenker, R. (2006). Quantile regression methods for recursive structural equation models. *Journal of Econometrics, 134*, 471–506.

Macbeth, J. D., & Merville, L. J. (1979). An empirical examination of the Black-Scholes call option pricing model. *Journal of Finance, 34*(5), 1173–1186.

MacDonald, R. (1992). Exchange rate survey data: A disaggregated G-7 perspective. The Manchester School vol LX Supplement, 47–62.

Mackie-Mason, J. K. (1990). Do taxes affect corporate financing decisions? *Journal of Finance, 45*, 1471–1493.

MacKinlay, A. C. (1987). On multivariate tests of the CAPM. *Journal of Financial Economics, 18*, 341–371.

MacKinlay, A. C. (1995). Multifactor models do not explain deviations from the CAPM. *Journal of Financial Economics, 38*, 3–28.

Mackinlay, C., & Richardson, M. (1991). Using Generalized Method of Moments to Test Mean-Variance Efficiency. *Journal of Finance, 46*(2), 511–527.

Maclennan, D. (1977). Some Thoughts on the Nature and Purpose of Hedonic Price Functions. *Urban Studies, 14*, 59–71.

Maclennan, D. (1982). *Housing Economics*. London: Longman.

Madan, D. B., & Seneta, E. (1990). The variance gamma (V.G.) model for share market returns. *Journal of Business, 63*(4), 511–524.

Madan, D. B., & Milne, F. (1991). Option Pricing with VG Martingale Components. *Mathematical Finance, 1*(4), 39–55.

Madan, D. B., Carr, P., & Chang, E. C. (1998). The variance gamma process and option pricing. *European Finance Review, 2*, 79–105.

Maddala, G. S. (1986). *Limited-dependent and qualitative variables in econometrics*. Cambridge/New York: Cambridge University Press.

Maddala, G. S., & Nimalendran, M. (1996). Error-in-variables problems in financial models. *Handbook of Statistics 14*, Elsevier Science Publishers B. V., 507–528.

Maddala, G. S., Rao, C. R., & Maddala, G. S. (1996). Handbook of Statistics 14: Statistical Methods in Finance. Elsevier Science & Technology.

Madhavan, A., & Smidt, S. (1991). A Bayesian Model of intraday specialist pricing. *Journal of Financial Economics, 30*, 99–134.

Mah, S., & Verde, M. (2004). Rating path dependency: An analysis of corporate and structured finance rating momentum. Fitch Ratings Report (March). New York.

Mahajan, S. (2002). Role of Out-of-Money Options in Executive Compensation. NYU, Stern School of Business, Working paper.

Maheu, J. M., & McCurdy, T. H. (2004). News arrival, jump dynamics and volatility components for individual stock returns. *Journal of Finance, 59*(2), 755–793.

Mahmoud, E. (1984). Accuracy in forecasting: A survey. *Journal of Forecasting, 3*, 139–159.

Mahoud, E. (1989). Combining forecasts: some managerial issues. *International Journal of Forecasting, 5*(4), 599–600.

Malaty, R., Toppinen, A., & Viitanen, J. (2007). Modeling and forecasting Finnish pine sawlog stumpage prices using alternative time-series methods. *Canadian Journal of Forest Research, 37*, 178–187.

Malevergne, Y., & Sornette, D. (2006). Higher Moment Portfolio Theory: Capitalizing on Behavioral Anomalies of Stock Markets. Working Paper, ISFA Graduate School of Actuarial Science.

Malkiel, B. (1987). Efficient Market Hypothesis. In The New Palgrave: A Dictionary of Economics, Macmillan Press Ltd., London, pp. 127–134.

Malkiel, B. G. (2003). The efficient Market hypothesis and its critics. *Journal of Economic Perspectives, 17*(1), 59–82.

Mallat, S. G., & Hwang, W. L. (1992). Singularity detection and processing with wavelets. *IEEE transactions on information theory, 38*(2), 617–643.
Malliavin, P., & Mancino, M. E. (2009). A Fourier transform method for nonparametric estimation of multivariate volatility. *Annals of Statistics, 37*(4), 1983–2010.
Mallows, C. L. (1973). Some comments on Cp. *Technometrics, 15*, 661–675.
Malmendier, U., & Tate, G. (2005a). Does Overconfidence Affect Corporate Investment? CEO Overconfidence Measures Revisited. *European Financial Management, 11*, 649–659.
Malmendier, U., & Tate, G. (2005b). CEO Overconfidence and Corporate Investment. *Journal of Finance, 60*, 2661–2700.
Malmendier, U., & Tate, G. (2008). Who Makes Acquisitions? CEO Overconfidence and the Market's Reaction. *Journal of Financial Economics, 89*, 20–43.
Malpezzi, S. (1998). A Simple Error Correction Model of House Prices. Wisconsin-Madison CULER working papers 98-11, University of Wisconsin Center for Urban Land Economic Research.
Mammen, E. (1993). Bootstrap and wild bootstrap for high dimensional linear models. *Annals of Statistics, 21*, 255–285.
Manaster, S., & Koehler, G. (1982). The Calculation of Implied Variances from the BlackScholes Model: A Note. *The Journal of Finance, 37*, 227–230.
Manaster, S., & Rendelman, R. J. (1982). Option prices as predictors of equilibrium stock prices. *Journal of Finance, 37*, 1043–1057.
Manchaldore, J., Palit, I., & Soloviev, O. (2010). Wavelet decomposition for intra-day volume dynamics. *Quantitative Finance, 10*, 917–930.
Mandelbrot, B. (1971). When can price be arbitraged efficiently? A limit to validity of the random walk and martingale models. *Review of Economics and Statistics, 53*, 225–236.
Mandelbrot, B. (1966). Forecasts of Future Prices, Unbiased Markets and Martingale Models. *Journal of Business, 39*, 242–255.
Mandelbrot, B. B. (1963). The Variation of Certain Speculative Prices. *Journal of Business, 36*, 394–419.
Mandelbrot, B. B., & Vanness, J. W. (1968). Fractional Brownian motions, fractional noises and applications. *SIAM Rev. 10*, 422–437.
Manganelli, S. (2005). Duration, volume and volatility impact of trades. *Journal of Financial Markets, 8*, 377–399.
Mankiw, N. G., Romer, D., & Shapiro, M. D. (1991). Stock market forecastability and volatility: A statistical appraisal. *Review of Economic Studies, 58*, 455–477.
Mann, C., Hamilton, D., Varma, P., & Cantor, R. (2003). What happens to fallen angels? *A statistical review* 1982–2003. Special comment. Moody's Investor Service.
Ammann, M., & Verhofen, M. (2008). The Impact of Prior Performance on the Risk-Taking of Mutual Fund Manager. *Annals of Finance, 5*, 69–90.
Marcus, A. J. (1984). Deregulation and bank financial policy. *Journal of Banking and Finance, 8*, 557–565.
Marcus, A. J., & Shaked, I. (1984). The valuation of FDIC deposit insurance using option-pricing estimates. *Journal of Money, Credit and Banking, 16*, 446–460.
Mardia, K. V., Kent, J. T., & Bibby, J. M. (1980). *Multivariate Analysis*. London: Academic Press.
Margrabe, W. (1978). The value of an option to exchange one asset for another. *Journal of Finance, 33*, 177–186.
Mark, N. (1995). Exchange rates and fundamentals: Evidence on long-horizon predictability. *American Economic Review, 85*, 201–218.
Mark, N., & Sul, D. (2001). Nominal exchange rates and monetary fundamentals: Evidence from a small post-Bretton Woods panel. *Journal of International Economics, 53*, 29–52.
Mark, N., & Yangru, W. (1998). Rethinking deviations from uncovered interest parity: the role of covariance risk and noise. *Economic Journal, 108*, 1686–1706.
Markowitz, H. (1959a). *Portfolio Selection: Efficient Diversification of Investments*. New York: Wiley.

Markowitz, H. M. (1959b). *Portfolio Selection*. New York: Wiley.
Markowitz, H. M. (1952). Portfolio selection. *The Journal of Finance, 7*, 77–91.
Markowitz, H. M. (1987). *Mean-Variance Analysis in Portfolio Choice and Capital Markets*. Oxford: Blackwell.
Marquardt, C., & Wiedman, C. (2004a). How are earnings managed? An examination of specific accruals. *Contemporary Accounting Research, 21*, 461–491.
Marquardt, C., & Wiedman, C. (2004b). The effect of earnings management on the value relevance of accounting information. *Journal of Business, Finance and Accounting, 31*, 297–332.
Marsh, T., & Rosenfeld, E. (1983). Stochastic Processes for Interest Rates and Equilibrium Bond Prices. *Journal of Finance, 38*, 635–646.
Marshall, A. W., & Olkin, I. (1979). *Inequalities: Theory of Majorization and Its Applications*. San Diego: Academic Press.
Marshall, C., & Siegel, M. (1997). Value At Risk: Implementing a Risk Measurement Standard. *Journal of Derivatives, 4*, 91–110.
Martens, M., & van Dijk, D. (2007). Measuring volatility with the realized range. *Journal of Econometrics, 138*, 181–207.
Martin, E. (2006). Performance Measurement of Hedge Funds using Data Envelopment Analysis. *Financial Markets and Port folio Management, 20*, 442–471.
Martin, J. (1996). Presidential Address: Another Puzzle: The Growth in Actively Managed Mutual Funds. *Journal of Finance, 51*, 783–810.
Martin, C. (1990). Corporate borrowing and credit constrains: Structural disequilibrium estimates for the U.K. *Review of Economics and Statistics, 72*(1), 78–86.
Martín, R., & Simin, T. (2002). Robust beta mining. In U. Fayyad, G. Grinstein, & A. Wierse (Eds.), *Information visualization in data mining and knowledge discovery* (pp. 309–312). San Francisco: Morgan Kaufmann Publishers Inc.
Mashal, R., & Naldi, M. (2003). Pricing Multi-name Default Swaps with Counterparty Risk. *Quantitative Credit Research*, November, 1–16.
Mason, R. (1987). The 10-year bond markets. Credit Suisse First Boston, CSFB Research.
Massimo Massa, & Xhang Lee. (2008). The Effects of Organizational Structure on Asset Management. http://ssrn.com/abstract=1328189.
Matsumoto, M., & Nishimura, T. (1998). Mersenne twister: a 623-dimensionally equidistributed uniform pseudorandom number generator. *ACM Transactions on Modeling and Computer Simulation, 8*, 3–30.
Matsumura, E. M., Tsui, K. W., & Wong, W. K. (1990). An extended multinomial-Dirichlet model for error bounds for dollar-unit sampling. *Contemporary Accounting Research, 6*, 485–500.
Maug, E., & Naik, N. (1996). Herding and delegated portfolio management: The Impact of Relative Performance Evaluation on Asset Allocation. IFA Working Paper 223. Available at: http://ssrn.com [Accessed 27 December 2011].
Mayer, W. J. (1989). Estimating disequilibrium models with limited a priori price-adjustment information. *Journal of Econometrics, 41*(3), 303–320.
Mayew, W. (2008). Evidence of management discrimination among analysts during earnings conference calls. *Journal of Accounting Research, 46*, 627–659.
Maynard, A., & Phillips, P. C. B. (2001). Rethinking and old empirical puzzle: Econometric evidence on the forward discount anomaly. *Journal of Applied Econometrics, 16*, 671–708.
Mazumder, M. I., & Varela, O. (2010). Market timing, the trading of international mutual funds: weekend, weekday and serial correlation strategies. *Journal of Business, Finance & Accounting, 37*(7–8), 979–1007.
Mazumder, R., Friedman, J. H., & Hastie, T. (2011). SparseNet: Coordinate Descent With Nonconvex Penalties. *Journal of the American Statistical Association, 106*, 1125–1138.
McAleer, M., & Medeiros, M. (2008). Realized volatility: A review. *Econometric Reviews, 27*, 10–45.
McBeth, J., & Merville, L. (1979). An Empirical Examination of the Black-Scholes Call Option Pricing Model. *Journal of Finance, 34*, 1173–1186.

McCabe, G. M. (1979). The empirical relationship between investment and financing: A new look. *Journal of Financial and Quantitative Analysis, 14*, 119–135.

McConnell, J., & Servaes, H. (1990). Additional evidence on equity ownership and corporate value. *Journal of Financial Economics, 27*, 595–612.

McCullagh, P. (1980). Regression models for ordinal data. *Journal of the Royal Statistical Society, 42*, 109–142.

McCulloch, C. E. (1997). Maximum likelihood algorithms for generalized linear mixed models. *Journal of the American Statistical Association, 92*, 162–170.

McCulloch, J. H. (1985). Interest-Risk Sensitive Deposit Insurance Premia: Stable ACH Estimates. *Journal of Banking and Finance, 9*, 137–156.

McCulloch, R. E., & Tsay, R. S. (1993). Bayesian inference and prediction for mean and variance shifts in autoregressive time series. *Journal of the American Statistical Association, 88*, 968–978.

McCurdy, T., & Morgan, I. G. (1987). Tests of the Martingale Hypothesis for Foreign Currency Futures with Time Varying Volatility. *International Journal of Forecasting, 3*, 131–148.

McDonald, J. B. (1984). Some Generalized Functions for the Size Distribution of Income. *Econometrica, 52*, 647–663.

McDonald, J. B. (1991). Parametric Models for Partially Adaptive Estimation with Skewed and Leptokurtic Residuals. *Economics Letters, 37*, 273–278.

McDonald, J. B., & Newey, W. K. (1988). Partially Adaptive Estimation of Regression Models Via the Generalized t Distribution. *Econometric Theory, 4*, 428–457.

McDonald, J. B., & Xu, Y. J. (1995). A Generalization of the Beta Distribution with Applications. *Journal of Econometrics, 66*, 133–152.

McDonald, J. G., Jacquillat, B., & Nussenbaum, M. (1979). Dividend, investment and financing decisions: Empirical evidence on French firms. *Journal of Financial and Quantitative Analysis, 10*, 741–755.

McDonald, R. L. (2006). The Role of Real Options in Capital Budgeting: Theory and Practice. *Journal of Applied Corporate Finance, 18*, 28–39.

McLachlan, G. J., & Krishnan, T. (2008). The EM Algorithm and Extensions (Wiley Series in Probability and Statistics), 2nd Ed., Wiley.

McNeil, A. (1998). Calculating quantile risk measures for financial time series using extreme value theory. Working paper, University of Zurich.

McNichols, M. (2000). Research design issues in earnings management studies. *Journal of Accounting and Public Policy, 19*, 313–345.

McQueen, G., & Roley, V. V. (1993). Stock prices, news, and business conditions. *Review of Financial Studies, 6*, 683–707.

Meddahi, N., & Renault, E. (2000). Temporal aggregation of volatility models. Document de Travail CIRANO, 2000-22.

Meddahi, N. (2001). An Eigenfunction Approach for Volatility Modeling. Working Paper, University of Montreal.

Meese, R., & Rogoff, K. (1983). Empirical exchange rate models of the seventies. Do they fit out of sample? *Journal of International Economics, 14*, 3–24.

Mehra, R., & Prescott, E. C. (1985). The Equity Premium: A Puzzle. *Journal of Monetary Economics, 15*, 145–161.

Mehran, H. (1995). Executive compensation structure, ownership, and firm performance. *Journal of Financial Economics, 38*(2), 163–184.

Meijun, Qian. (2009). Stale Prices and the Performance Evaluation of Mutual Funds. *Journal of Financial and Quantitative Analysis.*

Meinl, T., & Sun, E. W. (2012). A nonlinear filtering algorithm based on wavelet transforms for high-frequency financial data analysis. *Studies in Nonlinear Dynamics and Econometrics, 16*(3).

Melino, A., & Turnbull, S. (1990). Pricing Foreign Currency Options With Stochastic Volatility, *Journal of Econometrics, 45*, 239–265.

Melino, A., & Turnbull, S. (1995). Misspeciffcation and the Pricing and Hedging of Long-Term Foreign Currency Options. *Journal of International Money and Finance, 45*, 239–265.
Mell, P., & Grance, T. (2011). The NIST Definition of Cloud Computing. Special Publication 800–145.
Memmel, C. (2003). Performance hypothesis testing with the Sharpe ratio. *Finance Letters, 1*, 21–23.
Menkhoff, L., Sarno, L., Schmeling, M., & Schrimpf, A. (2012). Carry trades and global foreign exchange volatility. *Journal of Finance, 67*(2), 681–718.
Mensah, Y. M., & Miranti, P. J., Jr. (1989). Capital expenditure analysis and automated manufacturing systems: a review and synthesis. *Journal of Accounting Literature, 8*, 181–207.
Merton, C. R. (1973a). Theory of Rational Option Pricing. *The Bell Journal of Economics and Management Science, 4*, 141–183.
Merton, R. (1974). On the Pricing of Corporate Debt: The Risk Structure of Interest Rates. *Journal of Finance, 29*, 449–470.
Merton, R. C. (1973b). An Intertemporal Capital Assets Pricing Model. *Econometrica, 41*, 867–887.
Merton, R. C. (1976). Option pricing when underlying stock returns are discontinuous. *Journal of Financial Economics, 3*(1&2), 125–144.
Merton, R. C. (1977). An analytic derivation of the cost of deposit insurance and loan guarantees. *Journal of Banking and Finance, 1*, 3–11.
Merton, R. C. (1980). On estimating the expected return on the market: An exploratory investigation. *Journal of Financial Economics, 8*, 323–361.
Merton, R. C. (1981). On market timing and investment performance I: An equilibrium theory of value for market forecasts. *Journal of Business, 54*, 363–406.
Merville, L. J., & Petty, W. J. (1978). Transfer pricing for the multinational corporation. *The Accounting Review*, 935–951.
Mest, D., & Plummer, E. (2003). Analysts' rationality and forecast bias: Evidence from analysts' sales forecasts. *Review of Quantitative Finance and Accounting, 21*, 103–122.
Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., & Teller, E. (1953). Equations of state calculations by fast computing machines. *Journal of Chemical Physics, 21*, 1087–1092.
Meulbroek, L. K. (2001). The efficiency of equity-linked compensation: understanding the full cost of awarding executive stock options. *Financial Management, 30*, 5–44.
Meulbroek, L. K., Mitchell, M. L., Mulherin, J. H., Netter, J. M., & Poulsen, A. B. (1990). Shark repellants and managerial myopia: An empirical test. *Journal of Political Economy, 98*, 1108–1117.
Jensen, M. C. (1967). 'The Performance of Mutual Funds in the period 1945–1964. *Journal of Finance, 23*, 389–416.
Michaud, R. O. (1989). The Markowitz optimization enigma: Is 'optimized' optimal? *Financial Analysts Journal, 45*, 31–42.
Mikhail, M., Walther, D., & Willis, R. (1997). Do security analysts improve their performance with experience? *Journal of Accounting Research, 35*, 131–157.
Mikkelson, W. H., & Partch, M. M. (1986). Valuation effects of security offerings and the issuance process. *Journal of Financial Economics, 15*, 31–60.
Miller, A. J. (2002). *Subset Selection in Regression* (2nd Ed.). New York: Chapman & HALL/CRC.
Miller, M. H. (1977). Debt and Taxes. *Journal of Finance, 32*(2), 261–275.
Miller, M. H., & Modigliani, F. (1961). Dividend policy, growth, and the valuation of shares. *Journal of Business, 34*, 411–433.
Miller, M. H., & Modigliani, F. (1966). Some estimates of the cost of capital to the electric utility industry. *American Economic Review, 56*(3), 333–391.
Miller, M. H., & Rock, K. (1985). Dividend policy under asymmetric information. *Journal of Finance, 40*, 1031–1051.

Miller, M. H., Muthuswamy, J., & Whaley, R. E. (1994). Mean Reversion of Standard and Poor's 500 Index Basis Changes: Arbitrage-Induced or Statistical Illusion ? *Journal of Finance, 49*, 479–513.

Milojicic, S. D., et al. (2002). Peer-to-Peer Computing. HL Laboratories Research Report.

Miltersen, M., Sandmann, K., & Sondermann, D. (1997). Closed-form Solutions for Term Structure Derivatives with Lognormal Interest Rates. *Journal of Finance, 52*, 409–430.

Mincer, J., & Zarnowitz, V. (1969). The Evaluation of Economic Forecasts. In J. Mincer (Ed.), *Economic Forecasts and Expectations*. New York: National Bureau of Economic Research.

Mishkin F.S., (2001). The Transmission Mechanism and the Role of Asset Prices in Monetary Policy. NBER Working Paper 8617, December.

Mitchell, M., & Stafford, E. (2000). Managerial decisions and long-term stock price performance. *Journal of Business, 73*, 287–329.

Mittnik, S., Rachev, S., Doganoglu, T., & Chenyao, D. (1999). Maximum Likelihood Estimation of Stable Paretian Models. *Mathematical and Computer Modelling, 29*, 275–293.

Mitton, T. (2006). Stock market liberalization and operating performance at the firm level. *Journal of Financial Economics, 81*, 625–647.

Modigliani, F., & Perotti, E. (2000). Security markets versus bank finance: legal enforcement and investor protection. *International Review of Finance, 1*, 81–96.

Modigliani, F., & Miller, M. H. (1963). Corporate income taxes and the cost of capital: A correction. *American Economic Review, 63*, 433–443.

Modigliani, F., & Brumberg, R. (1955). Utility analysis and the consumption function. In K. Kurihara (Ed.), *Post Keynesian Economics*. London: G. Allen.

Modigliani, F., & Miller, M. (1958). The cost of capital, corporation finance and the theory of investment. *American Economic Review, 48*, 261–297.

Modigliani, Franco, & Modigliani, Leah. (1997). Risk Adjusted Performance. *Journal of Portfolio Management*, 45–54.

Modigliani, L. (1997). Don't pick your managers by their alphas. U.S. Investment Research, U.S. Strategy, Morgan Stanley Dean Witter, June.

Moerman, G. A., & van Dijk, M. A. (2010). Inflation risk and international asset returns. *Journal of Banking and Finance, 34*, 840–855.

Mohit, G., & Aggarwal, N. (2009). Mutual Fund Portfolio Creation using Industry Concentration. *The ICFAI Journal of Management Research, Viii*(3), 7–20.

Mohr, M. F. (2005). A trend-cycle (-season) filter. Working Paper Series European Central Bank, (499).

Molina, C. A. (2005). Are firms underleveraged? an examination of the effect of leverage on default probabilities. *Journal of Finance, 60*, 1427–1459.

Molina, G., Han, C.-H., & Fouque, J. P. (2010). MCMC Estimation of Multiscale Stochastic Volatility Models. *Handbook of Quantitative Finance and Risk Management*, Springer.

Molnár, P. (2010). Properties of range-based volatility estimators. *International Review of Financial Analysis, 23*, 20–29.

Monoyios, M. (2004). Performance of utility-based strategies for hedging basis risk. *Quantitative Finance, 4*, 245–255.

Montero, J. M., Garcia, M. C., & Fernandez-Aviles, G. (2011). Modelling the Volatility of the Spanish Wholesale Electricity Spot Market. Asymmetric GARCH Models vs. Threshold ARSV model. *Revista de Estudios de Economía Aplicada, 29*(2), 1–20.

Moody's. (2001). Moody's Country Credit Statistical Handbook, First Edition. New York.

Moody's. (2006). Bank Financial Strength Ratings: Revised Methodology. Moody's Global Credit Research Report.

Mora, N. (2006). Sovereign Credit Ratings: Guilty beyond Reasonable Doubt? *Journal of Banking and Finance, 30*(7), 2041–2062.

Morck, R., Yeung, B., & Yu, W. (2000). The information content of stock markets: why do emerging markets have synchronous stock price movements? *Journal of Financial Economics, 58*, 215–260.

Morck, R., Shleifer, A., & Vishny, R. W. (1988). Management ownership and market valuation: an empirical analysis. *Journal of Financial Economics, 20*, 293–315.

Moreira, M. J. (2002). *Tests with correct size in the simultaneous equations model*. Berkeley: University of California.

Morel, M. (2003). Endogenous Parameter Time Series Estimation of the Ohlson Model: Linear and Nonlinear Analyses. *Journal of Business Finance and Accounting, 30*, 1341–1362.

Morgan, D. P. (2002). Rating Banks: Risk and Uncertainty in an Opaque Industry. *The American Economic Review, 92*(4), 874–888.

Morgan, I. (1978). Market Proxies and the Conditional Prediction of Returns. *Journal of Financial Economics, 6*, 385–398.

Morgan, I., & Saint-Pierre, J. (1978). Dividend and investment decisions of Canadian firms. *Canadian Journal of Economics, 11*, 20–37.

Morgan, J. P. (1996). *RiskMetricsTM— Technical Document*. New York: Morgan Guaranty Trust Company.

Morlet, J. (1983). Sampling theory and wave propagation. *Issues in acoustic signal/image processing and recognition, 1*, 233–261.

Morrison, D. F. (1990). *Multivariate Statistical Methods* (3rd Ed.). New York: McGraw-Hill, Inc.

Morton, R., Kang, E. L., & Henderson, B. L. (2009). Smoothing splines for trend estimation and prediction in time series. *Environmetrics, 20*(3), 249–259.

Mosler, K., & Scarsini, M. (Eds.). (1991). *Stochastic Orders and Decision Under Risk*. Hayward, California: Institute of Mathematical Statistics.

Mossin, J. (1966). Equilibrium in a Capital Asset Market. *Econometrica, 35*, 768–783.

Muliere, P., & Scarsini, M. (1989). A note on stochastic dominance and inequality measures. *Journal of Economic Theory, 49*, 314–323.

Müller, A., & Stoyan, D. (2002). *Comparison Methods for Stochastic Models and Risks*. Chichester: Wiley.

Muller, D. C. (1985). Mergers and Market Share. *The Review of Economics and Statistics, 67*, 66–74.

Müller, G., Durand, R. B., & Maller, R. A. (2011). The risk–return tradeoff: A COGARCH analysis of Merton's hypothesis. *Journal of Empirical Finance, 18*, 306–320.

Müller, U., Dacorogna, M., Davé, R., Pictet, O., Olsen, R., & Ward, J. (1995). Fractals and Intrinsic Time - a Challenge to Econometricians. Working Papers, Olsen and Associates.

Munch, P., & Smallwood, D. (1978). *Solvency Regulation in the Property-casualty insurance industry*. Santa Monica, CA: Rand Corporation.

Muralidhar, K., Santhanam, R., & Wilson, R. L. (1990). Using the analytic hierarchy process for information system project selection. *Information and Management, 18*, 87–95.

Murphy, K. (1999). Executive compensation. In Orley Ashenfleter & David Card (Eds). *Handbook of Labor Economics*. North Holland: Amsterdam.

Murphy, K. (2003). Stock-based pay in new economy firms. *Journal of Accounting and Economics, 34*(1–3), 129–147.

Murphy, A. H. (1973). A new vector partition of the probability score. *Journal of Applied Meteorology, 12*, 595–600.

Murphy, A. H., & Winkler, R. (1977). Reliability of subjective probability forecasts of precipitation and temperature. *Applied Statistics, 26*, 41–47.

Musiela, M., & Rutkowski, M. (1997). Martingale Methods in Financial Modeling. Springer.

Muth, J. (1961). Rational expectations and the theory of price movements. *Econometrica, 29*, 315–335.

Myers, J. N. (1999). Implementing residual income valuation with linear information dynamics. *Accounting Reviews, 74*, 1–28.

Myers, R. J. (1991). Estimating time-varying optimal hedge ratios on futures markets. *Journal of Futures Markets, 11*(1), 39–53.

Myers, S. C. (1977a). Determinants of corporate borrowing. *Journal of Financial Economics, 5*, 146–175.

Myers, S. C., & Majluf, N. (1984). Corporate financing and investment decision when firms have information investors do not have. *Journal of Financial Economics, 13*, 187–221.

Myers, S. (1977b). Determinants of corporate borrowing. *Journal of Financial Economics, 5*, 147–175.
Myers, Stewart C. (1984). Finance Theory and Financial Strategy. *Interface 14*, 126–137, reprinted in 1987, *Midland Corporate Finance Journal 5*, 6–13.
Myers, Stewart C., & Borucki, Lynda S. (1994). Discounted Cash Flow Estimates of the Cost of Equity Capital—A Case Study. Financial Markets, Institutions & Instruments. *Estimating Cost of Capital: Methods and Practice, 3*(3), 9–45 (August 1994).
Myers, S. C., & Cohn, R. A. (1987). A Discounted Cash Flow Approach to Property-Liability Insurance Rate Regulation. In J. D. Cummins & S. E. Harrington (Eds.), *Fair Rate of Return in Property-Liability Insurance*. Norwell, MA: Kluwer Academic Publishers.
Nadenichek, J. (2006). The J-curve effect: an examination using a structural vector error correction model. *International Journal of Applied Economics, 3*(2), 34–47.
Naert, P. A. (1973). Measuring performance in a decentralized corporation with interrelated divisions: profit center versus cost center. *The Engineering Economist*, 99–114.
Nagpal, K., & Bahar, R. (2001). Measuring default correlation. *Risk, 14*(3), 129–132.
Nakayama, H. (1994). Some remarks on trade-off analysis in multi-objective programming. The XIth International Conference on Multiple Criteria Decision Making, Coimbra, Portugal, Aug. 1–6.
Nam, K., Pyun, C., & Avard, S. (2001). Asymmetric reverting behavior of short-horizon stock returns: An evidence of stock market overreaction. *Journal of Banking and Finance, 25*, 805–822.
Namboodiri, K., & Suchindran, C. M. (1987). *Life Tables and Their Applications*. Orlando: Academic Press.
Nandi, S. (1996). Pricing and Hedging Index Options under Stochastic Volatility. Working Paper, Federal Reserve Bank of Atlanta.
Nanni, A. J., Dixon, J. R., & Vollmann, T. E. (1990a). Strategic control and performance measurement. *Journal of Cost Management*, 33–42.
Nash, R. C., Netter, J. M., & Poulsen, A. B. (2001). Determinants of contractual relations between shareholders and bondholders: Investment opportunities and restrictive covenants. *Journal of Corporate Finance, 9*, 201–232.
Natarajan, T., Balasubramanian, S. A., & Manickavasagam, S. (2010). Customer's choice amongst self-service technology (SST) channels in retail banking: a study using analytical hierarchy process (AHP). *Journal of Internet Banking and Commerce, 15*, 1–15.
Nayak, S., & Prabhala, N. R. (2001). Disentangling the dividend information in splits: a decomposition using conditional event-study methods. *Review of Financial Studies, 14*, 1083–1116.
Neftci, S. (2000). Value at risk calculations, extreme events, and tail estimation. *Journal of Derivatives, 7*, 23–37.
Neil, R. (2001). A portfolio manager's view of downside risk. In Frank A. Sortino and Stephen E. Satchell (eds.), Managing downside risk in financial markets, Butterworth-Heinemann, UK.
Nelsen, R. B. (2006). *An introduction to copulas* (2nd Ed.). New York: Springer.
Nelson, C. (1976). Inflation and rate of returns on common stocks. *Journal of Finance, 31*, 471–483.
Nelson, C. (1979). Recursive structure in U.S. income, prices and output. *Journal of Political Economy, 87*, 1307–1327.
Nelson, C. R. (1993). *Applied Time Series Analysis for Managerial Forecasting*. New York: Holden-Day.
Nelson, C. R., & Siegel, A. F. (1987). Parsimonious modeling of yield curve. *Journal of Business, 60*, 473–489.
Nelson, D. (1991a). Conditional heteroskedasticity in asset returns: a new approach. *Econometrica, 59*, 347–370.
Nelson, D. B. (1990a). ARCH Models as Diffusion Approximations. *Journal of Econometrics, 45*, 7–38.
Nelson, D. B. (1990b). Stationarity and Persistence in the GARCH(1,1) Model. *Econometric Reviews, 6*, 318–334.

Nelson, D. B. (1991b). Conditional Heteroskedasticity in Asset Returns: A New Approach. *Journal of Econometrics, 45*, 347–370.
Nelson, D. B., & Foster, D. P. (1991). Estimating Conditional Variances with Misspecified ARCH Models: Asymptotic Theory. Graduate School of Business, University of Chicago, mimeo.
Nelson, D. B., & Foster, D. P. (1994). Asymptotic Filtering Theory for Univariate ARCH Models. *Econometrica, 62*, 1–41.
Nelson, D. B. (1990c). ARCH as Diffusion Approximations. *Journal of Econometrics, 45*, 7–38.
Nelson, R. B. (1998). *An Introduction to Copula*. New York: Springer.
Nerlove, M. (1971). A note on error components model. *Econometrica, 39*, 383–396.
Newbold, P., Carlson, W. L., & Thorne, B. M. (2009). *Statistics for Business and Economics* (7th Ed.). New Jersey: Pearson Education Canada.
Newey, W., & West, K. (1987). A simple positive semi-definite heteroscedasticity and autocorrelation consistent covariance matrix. *Econometrica, 55*, 703–708.
Neyman, J. (1952). *Lectures and conferences on mathematical statistics and probability* (2nd Ed., pp. 143–154). Washington, DC: U.S. Department of Agriculture.
Neyman, J., & Scott, E. (1948). Consistent estimates based on partially consistent observations. *Econometrica, 16*, 1–32.
Ng, W.-Y., Choi, K.-W., & Shum, K.-H. (1996). Arbitrated Matching: Formulation and Protocol. *European Journal of Operational Research, 88*(2), 348–357.
Ng, A. (2000). Volatility spillover effects from Japan and the US to the Pacific-Basin. *Journal of International Money and Finance, 19*(2), 207–233.
Nickell, P., Perraudin, W., & Varotto, S. (2000). Stability of ratings transitions. *Journal of Banking and Finance, 24*, 203–222.
Niederreiter, H. (1992). Random Number Generation and Quasi-Monte Carlo Methods. In CBMS-NSF Regional Conference Series in Appl. Math.
Nielsen, M., & Frederiksen, P. (2005). Finite sample comparison of parametric, semiparametric, and wavelet estimators of fractional integration. *Econometric Reviews, 24*, 405–443.
Nikos, K. (2006). Commodity prices and the influence of the US dollar. World Gold Council.
Nissim, D., & Ziv, A. (2001). Dividend changes and future profitability. *Journal of Finance, 56*, 2111–2133.
Nofsinger, J. R. W. S. (1999). Herding and feedback trading by institutional and individual investors. *Journal of Finance, 54*, 2263–2295.
Nohel, T., & Todd, S. (2005). Compensation for managers with career concerns: the role of stock options in optimal contracts. *Journal of Corporate Finance, 11*, 229–251.
Nordman, E., & Wagner, J. (2012). Public purchases and private preferences: Challenges for analyzing public open space acquisitions. *Urban Forestry & Urban Greening, 11*(2), 179–186.
Noreen, E. W. (1989). Computer intensive methods for testing hypotheses: an introduction, Wiley.
North, D. C. (1990). *Institutions, institutional change, and economic performance*. Cambridge: Cambridge University Press.
Noussair, C. N., & Tucker, S. (2006). Futures Markets and Bubble Formation in Experimental asset markets. *Pacific Economic Review, 11*, 167–184.
Numerical Algorithm Group. (1990). *Numerical Algorithm Group Library. Mark 13*. Oxford: Oxford University.
Nutt, S., Easterwood, J., & Easterwood, C. (1999). New evidence on serial correlation in analyst forecast errors. *Financial Management, 28*, 106–117.
Nyberg, H. (2012). Risk-return tradeoff in U.S. stock returns over the business cycle. *Journal of Financial and Quantitative Analysis, 47*, 137–158.
O'Brien, T. J., & Dolde, W. (2000). A currency index global capital asset pricing model. *European Financial Management, 6*(1), 7–18.
O'Hanlon, J., & Pope, P. (1999). The value-relevance of U.K. dirty surplus accounting flows. *British Accounting Review, 31*, 459–482.
Odders-White, E. R. (2000). On the Occurrence and Consequences of Inaccurate Trade Classification. *Journal of Financial Markets, 3*, 259–286.

OECD. (2010). Pension markets in focus, 7. Available at: http://www.oecd.org [Accessed 22 July 2011].

Oehler, A., & Goeth-Chi, G. (2006). "Institutional Herding in Bond Markets." Working paper, Digital archive for economics and business studies. Available at: http://ssrn.com [Accessed 27 December 2011].

Ofek, E., & Richardson, M. (2003). A survey of market efficiency in the internet sector. *Journal of Finance, 58*, 1113–1138.

Officer, R. (1972). The Distribution of Stock Returns. *Journal of the American Statistical Association, 67*, 807–812.

Ogryczak, W., & Ruszczynski, A. (2002). Dual stochastic dominance and related mean-risk models. *SIAM Journal on Optimization, 13*, 60–78.

Ogryczak, W., & Ruszczynski, A. (1999). From stochastic dominance to mean-risk models: semideviations as risk measures. *European Journal of Operational Research, 116*, 33–50.

Ogryczak, W., & Ruszczynski, A. (2001). On consistency of stochastic dominance and mean-semideviation models. *Mathematical Programming, 89*, 217–232.

Ohanissian, A., Russell, J. R., & Tsay, R. S. (2004). True or spurious long memory in volatility: does it matter for pricing options? Working paper, University of Chicago.

Ohanissian, A., Russell, J. R., & Tsay, R. S. (2008). True or spurious long memory? a new test. *Journal of Business and Economic Statistics, 26*, 161–175.

Ohlson, J. A. (1995a). Earnings, book values, and dividends in equity valuation. *Contemporary Accounting Research, 11*(2), 661–687.

Ohlson, J. A. (2001a). Earnings, book values, and dividends in equity valuation: An empirical perspective. *Contemporary Accounting Research, 18*(1), 107–120.

Ohlson, J. A. (1980a). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research, 18*, 109–131.

Ohlson, J. A. (1995b). Earnings, book values, and dividends in equity valuation. *Contemporary Accounting Research, 11*(2), 661–687.

Ohlson, J. A. (2001b). Earnings, book values, and dividends in equity valuation: An empirical perspective. *Contemporary Accounting Research, 18*(1), 107–120.

Ohlson, J. A., & Zhang, X. J. (1998). Accrual accounting and equity valuation. *Journal of Accounting Research, 36*(3), 85–111.

Ohlson, J. A., & Zhang, X. J. (1999). On the theory of forecast horizon in equity valuation. *Journal of Accounting Research, 37*(2), 437–449.

Ohlson, J. S. (1980b). Financial ratios and the probabilistic prediction of bankruptcy. *Journal of Accounting Research, 19*, 109–131.

Okimoto, T. (2008). New evidence of asymmetric dependence structures in international equity markets. *Journal of Financial and Quantitative Analysis, 43*, 787–815.

Oksendal, B. (1998). Stochastic Differential Equations. 5th Edition. Springer.

Onsi, M. (1970). A transfer pricing system based on opportunity cost. *The Accounting Review*, 535–543.

Onur, A., Edwards, E., & Samant, A. (2008). Risk-adjusted Performance of International Mutual Funds. *Managerial Finance, 34*(1), 5–22.

Opler, T., & Titman, S. (1993). The determinants of leveraged buyout activity: Free cash flow vs. financial distress costs. *Journal of Finance, 48*, 1985–1999.

Ortiz-Molina, H. (2006). Top management incentives and the pricing of corporate public debt. *Journal of Financial and Quantitative Analysis, 41*, 317–340.

Osterberg, W. P. (2000). New results on the rationality of survey measures of exchange-rate expectations. *Federal Reserve Bank of Cleveland Economic Reveiw, 36*, 14–21.

Oswiecimka, P., Kwapien, J., & Drozdz, S. (2006). Wavelet versus detrended fluctuation analysis of multifractal structures. *Physical Review, E 74*, 016103.

Ota, K. (2011). Analysts' awareness of systematic bias in management earnings forecasts. *Applied Financial Economics, 21*, 1317–1330.

Otto, C. A. (2011). CEO optimism and incentive compensation. Working paper, London Business School.

Ou, J., & Penman, S. (1989). Financial Statement Analysis and the Prediction of Stock Returns. *Journal of Accounting and Economics, 11*, 295–329.

Oyer, P., & Schaefer, S. (2005). Why do some firms give stock options to all employees?: an empirical analysis of alternative theories. *Journal of Financial Economics, 76*, 99–133.

Pagan, A. R., & Hong, Y. S. (1989). Nonparametric estimation and the risk premium. In W. A. Barnett, J. Powell, & G. E. Tauchen (Eds.), *Semiparametric and Nonparametric Methods in Econometrics and Statistics* (pp. 51–75). Cambridge: Cambridge University Press.

Palia, D. (2001). The endogeneity of managerial compensation in firm valuation: a solution. *Review of Financial Studies, 14*, 735–764.

Palliam, R. (2005). Application of a multi-criteria model for determining risk premium. *The Journal of Risk Finance, 6*, 341–348.

Palmon, O., & Venezia, I. (2012). A Rationale for Hiring Irrationally Overconfident Managers, forthcoming, Encyclopedia of Finance 2nd edition.

Palmon, O., Chen, R. R., Bar Yosef, S., & Venezia, I. (2008). Optimal Strike Prices of Stock Options for Effort Averse Executives. *Journal of Banking and Finance, 32*, 229–239.

Pan, J. (2002). The jump-risk premia implicit in options: Evidence from an integrated time-series study. *Journal of Financial Economics, 63*, 3–50.

Pan, J., & Poteshman, A. (2006). The information in option volume for stock prices. *Review of Financial Studies, 19*, 871–908.

Panzone, L. (2011). The lost scent of Eastern European wines in Western Europe: A hedonic model applied to the UK market. *British Food Journal, 113*(8), 1060–1078.

Papahristoulou, C., & Dotzauer, E. (2004). Optimal portfolios using linear programming models. *Journal of the Operational Research Society, 55*, 1169–1177.

Paparoditis, E., & Politis, D. (2003). Residual-based block bootstrap for unit root testing. *Econometrica, 71*, 813–855.

Pardo, A., Meneu, V., & Valor, E. (2002). Temperature and seasonality influences on Spanish Electricity Load. *Energy Economics, 24*, 55–70.

Pardoux, E., & Talay, D. (1985). Discretization and Simulation of Stochastic Differential Equations. *Acta Applicandae Mathematicae, 3*, 23–47.

Park, H. Y., & Sears, R. S. (1985). Changing Volatility and the Pricing of Options on Stock Index Futures. *Journal of Financial Research, 8*, 265–296.

Park, S. Y., & Jei, S. Y. (2010). Estimation and hedging effectiveness of time-varying hedge ratio – flexible GARCH approaches. *Journal of Futures Markets, 30*, 71–99.

Parker, C., & Weber, Bruce. (2012). How IT Can Disrupt Markets: A Simulation Analysis, London Business School, Working Paper.

Parkinson, M. (1980). The extreme value method for estimating the variance of the rate of return. *Journal of Business, 53*, 61–65.

Parks, R. W. (1967). Efficient estimation of a system of regression equations when disturbances are both serially and contemporaneously correlated. *Journal of the American Statistical Association, 62*, 500–509.

Parzen, M. I., Wei, L. J., & Ying, Z. (1994). A resampling method based on pivotal estimating functions. *Biometrika, 81*, 341–350.

Pas, P. F., & Lin, C. S. (2005). A hybrid ARIMA and support vector machines model on stock price forecasting. *OMEGA, 33*, 497–505.

Pasaribu, R. B. (2009). Relevance of accrual anomaly information in stock portfolio formation – case study: Indonesia stock exchange. *Journal of Accounting and Business 10*(2), August.

Paskov, S. H., & Traub, J. F. (1995). Faster Valution of Financial Derivatives. *Journal of Portfolio Management, 22*(1), 113–120.

Pasquariello, P., & Vega, C. (2007). Informed and strategic order flow in the bond markets. *Review of Financial Studies, 20*, 1975–2019.

Pasquariello, P., & Vega, C. (2009). The on-the-run liquidity phenomenon. *Journal of Financial Economics, 92*, 1–24.

Pástor, L., & Stambaugh, R. (2003). Liquidity risk and expected stock retums. *Journal of Political Economy*, 642–685.

Pastor, L., Sinha, M., & Swaminthan, B. (2008). Estimating the intertemporal risk-return tradeoff using the implied cost of capital. *Journal of Finance, 63*, 2859–2897.

Patrick, B., Ferreira, F., & McMillan, R. (2007). A Unified Framework for Measuring Preferences for Schools and Neighborhoods. *Journal of Political Economy, 115*(4), 588–638.

Patterson, D. M., & Sharma, V. (2006). Do Traders Follow Each Other at the NYSE? Working Paper, University of Michigan-Dearborn.

Patton, A. (2006a). Modelling asymmetric exchange rate dependence. *International Economic Review, 47*, 527–556.

Patton, A. (2006b). Estimation of multivariate models for time series of possibly different lengths. *Journal of Applied Econometrics, 21*, 147–173.

Patton, A. (2009). Copula-based models for financial time series. In T. G. Andersen, R. A. Davis, J. P. Kreiss, & T. Mikosch (Eds.), *Handbook of Financial Time Series*. New York: Springer.

Patton, A. J. (2011). Volatility forecast comparison using imperfect volatility proxies, *Journal of Econometrics, 160*, 246–256.

Patton, A. J. (2004). On the Out-of-Sample Importance of Skewness and Asymmetric Dependence for Asset allocation. *Journal of Financial Econometrics, 2*, 130–168.

Patton, A. J. (2006c). Modelling asymmetric exchange rate dependence. *International Economic Review, 47*, 527–555.

Patton, A. J. (2006). Estimation of multivariate models for time series of possibly different lengths. *Journal of Applied Econometrics, 21*(2), 147–173.

Patton, A. J. (2006). Modelling Asymmetric Exchange Rate Dependence. *International Economic Review, 47*(2), 527–556.

Patton, A. J. (2004). On the Out-of-Sample Importance of Skewness and Asymmetric Dependence for Asset Allocation. *Journal of Financial Econometrics, 2*(1), 130–168.

Pearson, E. S., & Please, N. W. (1975). Relation between the shape of population distribution and the robustness of four simple test statistics. *Biometrika, 62*, 223–241.

Pearson, K. (1897). Mathematical contributions to the theory of evolution- on a form of spurious correlation which may arise when indices are used in the measurement of organs. *Proceedings of the Royal Society of London, 60*, 489–497.

Peasnell, K., Pope, P., & Young, S. (2005). Board Monitoring and Earnings Management: Do Outside Directors Influence Abnormal Accruals?

Peavy, J. W., III. (1990). Returns on Initial Public Offerings of Closed-End Funds. *Review of Financial Studies, 3*, 695–708.

Peltonen, S., Gabbouj, M., & Astola, J. (2001). Nonlinear filter design: methodologies and challenges. In Proceedings of the 2nd International Symposium on Image and Signal Processing and Analysis, 2001. ISPA 2001, pp. 102–107.

Pena, D., Tiao, G. C., & Tsay, R. S. (2001). *A Course in Time Series Analysis*. New York: Wiley.

Peña, J., Vera, J. C., & Zuluaga, L. F. (2010). Static-arbitrage lower bounds on the prices of basket options via linear programming. *Quantitative Finance, 10*, 819–827.

Penman, S. H., & Sougiannis, T. (1998). A comparison of dividend, cash flow, and earnings approaches to equity valuation. *Contemporary Accounting Research, 15*(3), 343–383.

Penman, S. H. (2000). *Financial Statement Analysis and Security Valuation*. New York: McGraw-Hill.

Percival, D., & Walden, A. (2000). Wavelet Methods for Time Series Analysis. Cambridge University Press.

Pérez-Rodríguez, J. (2006). The Euro and other major currencies floating against the U.S. dollar. *Atlantic Economic Journal, 34*, 367–384.

Perkins, A. B., & Perkins, M. C. (1999). *The internet bubble: Inside the overvalued world of high-tech stocks*. New York: Harper Business.

Perrakis, S. (1986). Option pricing bounds in discrete time: extensions and the pricing of the American put. *Journal of Business, 59*, 119–141.

Perrakis, S., & Ryan, J. (1984). Option pricing in discrete time. *Journal of Finance, 39*, 519–525.

Perron, P. (1989). The great crash, the oil price shock and the unit root hypothesis. *Econometrica, 57*, 1361–1401.

Perron, P., & Ng, S. (1996). Useful modifications to some unit root tests with dependent errors and their local asymptotic properties. *Review of Economic Studies, 63*, 435–463.

Perry, T. (2000). Incentive compensation for outside directors and CEO turnover. Working Paper, Arizona State University.

Pesando, J. (1974). The supply of money and common stock prices: Further observations on econometric evidence. *Journal of Finance, 29*, 909–921.

Pesaran, H. M., & Weale, M. (2006). Survey expectations. In G. E. C. W. J. Granger, & A. Timmermann (Eds.), *Handbook of Economic Forecasting*. North-Holland.

Pesaran, H. M. (2004). General diagnostic tests for cross section dependence in panels. CESifoWorking Paper Series 1229, CESifo Group Munich.

Petersen, M. (2009). Estimating standard errors in finance panel data sets: comparing approaches. *Review of Financial Studies, 22*, 435–480.

Petersen, M., & Rajan, R. (1994). The benefits of lending relationships: Evidence from small business data. *Journal of Finance, 49*, 3–37.

Peterson, P. P., & Benesh, G. A. (1983). A reexamination of the empirical relationship between investment and financing decisions. *Journal of Financial and Quantitative Analysis, 18*, 439–453.

Petkova, R., & Zhang, L. (2005). Is value riskier than growth. *Journal of Financial Economics, 78*, 187–202.

Pettengill, G. N., Sundaram, S., & Mathur, I. (1995). The conditional relation between beta and returns. *Journal of Financial and Quantitative Analysis*, March.

Peyer, U., & Theo, V. (2009). The nature and persistence of buyback anomalies. *Review of Financial Studies, 22*, 1693–1745.

Pfaff, B. (2008). VAR, SVAR and SVEC Models: Implementation Within R Package vars. *Journal of Statistical Software*, 27(4).

Philbrick, D., & Ricks, W. (1991). Using value line and IBES analyst forecasts in accounting research. *Journal of Accounting Research, 29*(2), 397–417.

Phillips P. C. B., Shi, S., & Yu, J. (2011). Testing for multiple bubbles. Singapore Management University working paper.

Phillips, A., et al. (1984). Effects of Information Technology on Financial Services Systems. Washington, D. C.: U.S. Congress, Office of Technology Assessment, OTA-CIT-202.

Phillips, P., & Hansen, B. E. (1990). Statistical inference in instrumental variables regression with I(1) processes. *Review of Economic Studies, 57*, 99–125.

Phillips, P. C. B., & Maynard, A. (2001). Rethinking an old empirical puzzle: Econometric evidence on the forward discount anomaly. *Journal of Applied Econometrics, 16*, 671–708.

Phillips, P. C. B., & Perron, P. (1988). Testing for a unit root in time series regression. *Biometrika, 75*, 335–346.

Phillips, P. C. B., Wu, Y., & Yu, J. (2009). Explosive behavior in the 1990s Nasdaq: when did exuberance escalate asset values? Cowles Foundation Discussion Paper No.1699.

Piazzesi, M. (2001). Macroeconomic Jump Effects and the Yield Curve. Working Paper, UCLA.

Piazzesi, M. (2003). Affine Term Structure Models, in Handbook of Financial Econometrics, North-Holland, 2010.

Piazzesi, M. (2004). Affine Term Structure Models, Manuscript prepared for the Handbook of Financial Econometrics, University of California at Los Angeles.

Piazzesi, M. (2005). Bond Yields and the Federal Reserve. *Journal of Political Economy, 113*, 311–344.

Pike, R. (1983). The capital budgeting behavior and corporate characteristics of capital-constrained firms. *Journal of Business Finance and Accounting, 10*, 663–671.

Pike, R. (1988). An empirical study of the adoption of sophisticated capital budgeting practices and decision-making effectiveness. *Accounting and Business Research, 18*, 341–351.
Pike, R., & Sharp, J. (1989). Trends in the use of management science techniques in capital budgeting. *Managerial and Decision Economics, 10*, 135–140.
Pilbeam, K. (1995). Exchange rate models and exchange rate expectations: an empirical investigation. *Applied Economics, 27*, 1009–1015.
Pinches, G. E., & Mingo, K. A. (1973). A multivariate analysis of industrial bond ratings. *Journal of Finance, 28*(1), 1–18.
Pindyck, R. (1984). Risk, inflation, and the stock market. *American Economic Review, 74*, 335–51.
Pindyck, R. S. (1986). Risk aversion and determinants of stock market behavior, NBER Working Paper, No. 1921.
Plott, C., & Sunder, S. (1982). Efficiency of experimental markets with insider information: An application of rational expectations models. *Journal of Political Economy, 90*, 663–698.
Politis, D. N., Romano, J. P., & Wolf, M. (1999). *Subsampling*. New York: Springer and Verlag.
Pollock, D. S. G. (2001). Methodology for trend estimation. *Economic Modelling, 18*(1), 75–96.
Pompe, P. M., & Bilderbeek, J. (2005). The prediction of bankruptcy of small- and medium-sized industrial firms. *Journal of Business Venturing, 20*, 847–868.
Pong, S., Shackleton, M. B., & Taylor, S. J. (2008). Distinguishing short and long memory volatility specifications. *Econometrics Journal, 11*, 617–637.
Pontiff, J. (1997). Excess Volatility and Closed-End Funds. *American Economic Review, 87*(1), 155–169.
Poon, W. P. H., Firth, M., & Fung, H.-G. (1999). A Multivariate Analysis of the Determinants of Moody's Bank Financial Strength Ratings. *Journal of International Financial Markets, Institutions & Money, 9*(3), 267–283.
Poon, S. H., & Granger, C. W. J. (2003). Forecasting volatility in financial markets: A review. *Journal of Economic Literature, 41*, 478–639.
Poon, S. H., Rockinger, M., & Tawn, J. (2004). Extreme value dependence in financial markets: Diagnostics, models and financial implications. *Review of Financial Studies, 17*, 581–610.
Popescu, M., & Kumar, R. (2008). An Ex-Ante Measure of the Probability of Informed Trading. http://ssrn.com/abstract=891717.
Porter, D. P., & Smith, V. L. (2003). Stock market bubbles in the laboratory. *Journal of Behavioral Finance, 4*(1), 7–20.
Portnoy, S. L., & Koenker, R. (1997). The gaussian hare and the Laplacian tortoise: computability of squared-error vs. absolute-error estimator. *Statistical Science, 12*, 279–300.
Poskitt, D. S., Athanasopoulos, G., & Vahid, F. (2007). Two canonical VARMA forms: Scalar component models vis-vis the Echelon form.
Poterba, J., & Summers, L. (1986). The persistence of volatility and stock market fluctuations. *American Economic Review, 76*, 1142–1151.
Poterba, J. (1988). Financing Constraints and Corporate Investment: Comment. *Brookings Papers on Economic Activity, 1*, 200–204.
Poterba, J. M., & Summers, L. H. (1986). The persistence of volatility and stock market fluctuations. *American Economic Review, 76*, 1142–1151.
Poudyal, N., Hodges, D., & Merrett, C. (2009). A hedonic analysis of the demand for and benefits of urban recreation parks. *Land Use Policy, 26*(4), 975–983.
Powell. (1986). Censored regression quantiles. Journal of Econometrics, 32, 143–55.
Prasch, R. E. (2001). Work, Leisure, and the Labor Supply Curve: A Reply to David Spencer. *Journal of Economic Issues, 35*, 1001–1007.
Pratt, J. W. (1964). Risk aversion in the small and in the large. *Econometrica, 32*, 122–136.
Prekopa, A. (2003). Probabilistic Programming. In A. Ruszczynski & A. Shapiro (Eds.), *Stochastic Programming* (pp. 267–351). Amsterdam: Elsevier.
Prescott, E. S. (1997). The pre-commitment approach in a model of regulatory banking capital. *FRB of Richmond Economic Quarterly, 83*, 23–50.

Previts, G., & Bricker, R. (1994). A content analysis of sell-side financial analyst company reports. *Accounting Horizons, 8*, 55–70.
Price, K., Price, B., & Nantel, T. J. (1982). Variance and lower partial moment measures of systematic risk: some analytical and empirical results. *Journal of Finance, 37*, 843–855.
Pritsker, M. (1998). Nonparametric Density Estimators and Tests of Continuous Time Interest Rate Models. *Review of Financial Studies, 11*, 449–487.
Protter P. (1990). Stochastic Integration and Differential Equations: A new approach. Springer-Verlag.
Pruitt, S. W., & Gitman, L. J. (1991). The interactions between the investment, financing, and dividend decisions of major US firms. *Financial Review, 26*, 409–430.
Puckett, A., & Yan, X. (2007). The Determinants and Impact of Short-Term Institutional Herding. Working Paper. Available at: http://ssrn.com [Accessed 27 December 2011].
Puri, M. (1996). Commercial banks in investment banking: conflict of interest or certification role? *Journal of Financial Economics, 40*, 373–401.
Qiang, B., & Lacey, N. (2008). On Understanding Mutual fund terminations. *Journal of Economics and Finance, 33*, 80–99.
Qiu, P. (2003). A jump-preserving curve fitting procedure based on local piecewise-linear kernel estimation. *Journal of Nonparametric Statistics, 15*(4), 437–453.
Qiu, P., & Yandell, B. (1998). A local polynomial jump-detection algorithm in nonparametric regression. *Technometrics, 40*(2), 141–152.
Qu, Z., & Perron, P. (2007). Estimating and testing structural changes in multivariate regressions. *Econometrica, 75*, 459–502.
Quandt, R. E. (1958). The estimation of the parameters of a linear regression system obeying two separate regimes. *Journal of the American Statistical Association, 53*, 873–880.
Quandt, R. E. (1960). Tests of the hypothesis that a linear regression system obeys two separate regimes. *Journal of the American Statistical Association, 55*, 324–330.
Quandt, R. E. (1988). *The Econometrics of Disequilibrium*. New York: Basil Blackwell Inc.
Quenouille, M. (1949). Approximate Tests of Correlation in Time Series. *Journal of the Royal Statistical Society, Series B, 11*, 18–84.
Quiggin, J. (1982). A theory of anticipated utility. *Journal of Economic Behavior and Organization, 3*, 225–243.
Quiggin, J. (1993). *Generalized Expected Utility Theory - The Rank-Dependent Expected Utility Model*. Dordrecht: Kluwer.
Quintos, C. E., & Phillips, P. C. B. (1993). Parameter constancy in cointegrating regressions. *Empirical Economics, 18*, 675–706.
Quirin, G. D., & Waters, W. R. (1975). Market Efficiency and the Cost of Capital: The Strange Case of Fire and Casualty Insurance Companies. *Journal of Finance, 30*(2), 427–224.
Quirk, J. P., & Saposnik, R. (1962). Admissibility and measurable utility functions. *Review of Economic Studies, 29*, 140–146.
Rachev, S. T. (2003). Handbook of heavy tailed distributions in finance. North-Holland.
Rachev, S. T., Mittnik, S., Fabozzi, F. J., Focardi, S. M., & Jašić, T. (2008). *Financial econometrics: from basics to advanced modelling techniques*. Hoboken, New Jersey: John Wiley & Sons.
Ragunathan, V., & Peker, A. (1997). Price variability, trading volume and market depth: evidence from the Australian futures market. *Applied Financial Economics, 7*, 447–454.
Rahul, B., Gallo, J., & Swason, P. T. (2001). The Performance, Asset Allocation and Investment Style of International Equity Manager. *Review of Quantitative Finance and Planning, 17*, 377–395.
Raimondo, M. (1998). Minimax estimation of sharp change points. *Annals of Statistics, 26*(4), 1379–1397.
Raimondo, M., & Tajvidi, N. (2004). A peaks over threshold model for change-point detection by wavelets. *Statistica Sinica, 14*(2), 395–412.
Rajan, R. G. (1992). Insiders and outsiders: the choice between informed and arm's length debt. *Journal of Finance, 47*, 1367–1400.

Rajan, R. G. (1994). Why credit policies fluctuate: A theory and some evidence. *Quarterly Journal of Economics, 436*, 399–442.

Rajan, R. G., & Zingales, L. (1998). Which capitalism? lessons from the East Asian crisis. *Journal of Applied Corporate Finance, 11*, 40–48.

Rajgopal, S., & Shevlin, T. (2002). Empirical evidence on the relation between stock option compensation and risk taking. *Journal of Accounting and Economics, 33*(2), 145–171.

Ramakrishnan, L., Canon, R. S., Muriki, K., Sakrejda, I., & Wright, N. J. (2011). Evaluating interconnect and virtualization performance for high performance computing. Proceedings of 2nd International Workshop on Performance Modeling, Benchmarking and Simulation of High Performance Computing Systems, 1-2.

Ramaswamy, K., & Sundaresan, S. M. (1985). The valuation of options on futures contracts. *Journal of Finance, 40*(5), 1319–1340.

Ramaswamy, R., & Sloek, T. (1998). The real effects of monetary policy in the European Union: What are the differences? *IMF Staff Papers, 45*, 347–395.

Ramesh, C. (2006a). Investment Manager's market Timing abilities: Empirical evidences. *The ICFAI Journal of Applied Finance, 12*(8), 15–31.

Ramnath, S. (2002). Investor and analyst reactions to earnings announcements of related firms: An empirical analysis. *Journal of Accounting Research, 40*, 1351–1376.

Ramsey, J. B. (1999). The contribution of wavelets to the anlaysis of economic and financial data. *Philosophical Transactions of the Royal Society of London A, 357*, 2593–2606.

Ramsey, J. B. (2002). Wavelets in economics and finance: Past and future. *Studies in Nonlinear Dynamics & Econometrics, 3*, 1–29.

Ramsey, J. B., & Zhang, Z. (1997). The analysis of foreign exchange data using waveform dictionaries. *Journal of Empirical Finance, 4*, 341–372.

Ramsey, J. B., Zaslavsky, G., & Usikov, D. (1995). An analysis of U. S. stock price behavior using wavelets. *Fractals, 3*, 377–389.

Ranaldo, A., & Soderlind, P. (2009). Safe haven currencies. *Review of Finance, 14*(3), 385–407.

Ranaldo, A., & Favre, L. (2003). How to Price Hedge Funds: From Two- to Four-Moment CAPM. EDHEC Publications.

Rappoport, P., & Reichlin, L. (1989). Segmented trends and non-stationary time series. *The Economic Journal, 99*, 168–177.

Rauh, J. (2006). Investment and Financing Constraints: Evidence from the Funding of Corporate Pension Plans. *Journal of Finance, 61*, 33–71.

Ravid, S. A., & Sarig, O. H. (1991). Financial signaling by committing to cash outflows. *Journal of Financial and Quantitative Analysis, 26*, 165–80.

Ravn, M. O., & Uhlig, H. (2002). On adjusting the Hodrick-Prescott filter for the frequency of observations. *Review of Economics and Statistics, 84*(2), 371–376.

Redmond, C. H., & Cubbage, F. W. (1988). Portfolio risk and returns from timber asset investments. *Land Economics, 64*, 325–337.

Reeves, G. A., & Hedin, S. R. (1993). A generalized interactive goal programming procedure. *Computers and Operations Research, 20*, 747–753.

Reeves, G. F., Lawrence, S. M., & Gonzalez, J. J. (1988). A multiple criteria approach to aggregate industrial capacity expansion. *Computers and Operations Research, 15*, 333–339.

Reeves, G. R., & Franz, L. S. (1985). A simplified interactive multiple objective linear programming procedure. *Computers and Operations Research, 12*, 589–601.

Reiner, E., & Rubinstein, M. (1991). Breaking Down the Barriers. *Risk, 4*, 28–35.

Reinhart, C. M. (2002). Default, Currency Crises, and Sovereign Credit Ratings. *World Bank Economic Review, 16*, 151–170.

Rendleman, R., & Bartter, B. (1979). Two-state asset pricing. *Journal of Finance, 34*, 1093–1110.

Reno, R. (2008). Nonparametric estimation of the diffusion coefficient of stochastic volatility models. *Econometric Theory, 24*(5), 1174–1206.

Rhodes. (1993). Power, Empire Building, and Mergers. Lexington Books.

Riahi-Belkaoui, A., & Pavlik, E. (1993). Effects of ownership structure, firm performance, size and diversification strategy on CEO compensation: A path analysis. *Managerial Finance, 19*(2), 33–54.

Richardson, S., Teoh, S., & Wysocki, P. (2004). The walk-down to beatable analyst forecasts: The roles of equity issuance and insider trading incentives. *Contemporary Accounting Research, 21*, 885–924.

Ricke, M. (2004). What is the link between margin loans and stock market bubbles? University of Muenster, Department of Banking No. 03-01.

Riddick, L., & Whited, T. (2009). The Corporate Propensity to Save. *Journal of Finance, 64*, 1729–1766.

Rigobon, R., & Sack, B. (2005). The effect of war on US financial markets. *Journal of Banking and Finance, 29*(7), 1769–1789.

Ripple, R. D., & Moosa, I. A. (2009). The effect of maturity, trading volume, and open interest on crude oil futures price range-based volatility. *Global Finance Journal, 20*, 209–219.

Ritchken, P., & Sankarasubramanian, L. (1992). Pricing the quality option in Treasury bond futures. *Mathematical Finance, 2*(3), 197–214.

Ritchken, P. (1985). On option pricing bounds. *Journal of Finance, 40*, 1219–1233.

Ritchken, P., & Kuo, S. (1988). Option bounds with finite revision opportunities. *Journal of Finance, 43*, 301–308.

Ritchken, P., & Kuo, S. (1989). On stochastic dominance and decreasing absolute risk averse option pricing bounds. *Management Science, 35*, 51–59.

Ritchken, P., & Trevor, R. (1999). Pricing options under generalized GARCH and stochastic volatility processes. *Journal of Finance, 54*, 377–402.

Ritter, J. (1991). The long-run performance of initial public offerings. *Journal of Finance, 46*, 1–27.

Robbins, E. H., & Schatzberg, J. D. (1986). Callable bonds: A risk-reducing signaling mechanism. *Journal of Finance, 41*, 935–50.

Robbins, H., & Siegmund, D. (1970). Boundary Crossing Probabilities for the Wiener Process and Sample Sums. *The Annals of Mathematical Statistics, 41*, 1410–1429.

Robin, S., & Ruffieux, B. (2001). Price Bubbles in Laboratory Asset markets with Constant Fundamental Values. *Experimental Economics, 4*, 87–105.

Robinson, L. (1991). The time has come to report comprehensive income. *Accounting Horizons, 5* (June), 107–112.

Robinson, P. M., & Zaffaroni, P. (1998). Nonlinear time series with long memory: a model for stochastic volatility. *Journal of Statistical Planning and Inference, 68*, 359–371.

Rockafellar, R. T., & Uryasev, S. (2000). Optimization of conditional value-at-risk. *Journal of Risk, 2*, 21–41.

Rockafellar, R. T., & Uryasev, S. (2002). Conditional value-at-risk for general loss distribution. *Journal of Banking and Finance, 26*, 1443–1471.

Rodriguez, J. C. (2007). Measuring financial contagion: a copula approach. *Journal of Empirical Finance, 14*, 401–423.

Rodriguez, R. (2003). Option pricing bounds: synthesis and extension. *Journal of Financial Research, 26*, 149–164.

Rogers, D. S., & Van Dyke, C. J. V. (2006). Measuring the volatility of hedge fund returns. Journal of Wealth Management, 45–53.

Rogers, L., & Satchell, S. (1991). Estimating variance from high, low and closing prices. *Annals of Applied Probability, 1*, 504–512.

Rogers, L., & Zhou, F. (2008). Estimating correlation from high, low, opening and closing prices. *The Annals of Applied Probability, 18*, 813–823.

Rogers, L., Satchell, S., & Yoon, Y. (1994). Estimating the volatility of stock prices: a comparison of methods that use high and low prices. *Applied Financial Economics, 4*, 241–247.

Rogers, R., & Grant, J. (1997). Content analysis of information cited in reports of sell-side financial analysts. *Journal of Financial Statement Analysis, 3*, 17–30.

Rogoff, K., & Reinhart, C. M. (2010). Growth in a Time of Debt. NBER Working Paper 15795.

Roll, R. (1977a). A critique of the asset pricing theory's tests. *Journal of Financial Economics, 4*, 129–176.

Roll, R. (1984). A simple implicit measure of the effective bid-ask spread in an efficient market. *Journal of Finance, 39*, 1127–1139.

Roll, R. (1992). A mean/variance analysis of tracking error. *Journal of Portfolio Management*, summer, 13–22.

Roll, R. (1986). The hubris hypothesis of corporate takeovers. *Journal of Business, 59*, 197–216.

Roll, R. (1989). Price volatility, international market links, and their implications for regulatory policy. *Journal of Financial Services Research, 3*, 211–246.

Roll, R., & Ross, S. (1980). An empirical investigation of the arbitrage pricing theory. *Journal of Finance, 35*, 1073–1103.

Roll, R., & Ross, S. A. (1984). A Critical Reexamination of the Empirical Evidence on the Arbitrage Pricing Theory: A Reply. *Journal of Finance, 39*(2), 347–350.

Roll, R. R. (1982). A mean variance analysis of tracking error. *Journal of Portfolio Management, 18*, 13–22.

Roll, R. R. (1977b). A critique of the asset pricing theory's tests - part 1: On past and potential testability of the theory. *Journal of Financial Economics, 4*, 129–176.

Roll, R. R. (1980). Orthogonal portfolios. *Journal of Financial and Quantitative Analysis, 15*, 1005–1023.

Roll, R. R. (1985). A note on the geometry of Shanken's CSR T2 test for mean/variance efficiency. *Journal of Financial Economics, 14*, 349–357.

Roll, R. (1977c). A Critique of the Asset Pricing Theory's Tests. *Journal of Financial Economics, 4*, 129–176.

Ronen, J., & McKinney, G. (1970). Transfer pricing for divisional autonomy. *Journal of Accounting Research*, 99–112.

Ronen, J., & Yaari, V. (2008). Earnings management: Emerging insights in theory, practice, and research. Springer Series in Accounting Scholarship, N.Y.

Ronen, T., & Zhou, X. (2012). Trade and Information in the Corporate Bond Market, forthcoming. *Journal of Financial Markets*.

Rosen, S. (1982). Authority, control, and distribution of earnings. *Bell Journal of Economics, 13*, 311–323.

Rosen, S. (1974). Hedonic Prices and Implicit Markets: Product Differentiation in Pure Competition. *Journal of Political Economy, 82*(1), 34–55.

Rosenberg, B., & Marathe, V. (1975). The Prediction of Investment Risk: Systematic and Residual Risk. In Proceedings of the Seminar on the Analysis of Security Prices, Center for Research in Security Prices. University of Chicago.IL, pp. 185–226.

Rosenberg, J., & Engle, R. (2002). Empirical pricing kernels. *Journal of Financial Economics, 64*(3), 341–372.

Rosenblatt, M. (1952). Remarks on a Multivariate Transformation. *Annals of Mathematical Statistics, 23*, 470–472.

Rosenstein, S., & Wyatt, J. (1990). Outside directors, board independence, and shareholder wealth. *Journal of Financial Economics, 26*, 175–192.

Rosett, R. N. (1959). A statistical model of friction in economics. *Econometrica: Journal of the Econometric Society*, 263–267.

Ross, S. A. (2004). Compensation, incentives, and the duality of risk aversion and riskiness. *Journal of Finance, 59*(1), 207–225.

Ross, S. (1976a). The arbitrage theory of capital asset pricing. *Journal of Economic Theory, 13*, 341–360.

Ross, S. (1977). Return, risk, and arbitrage. In Friend and Bicksler (Eds.), *Risk and Return in Finance*. Vol. I, MA: Cambridge.

Ross, S. (2001). *Simulation* (3rd ed.). Boston, Massachusetts: Academic Press.

Ross, S. A. (1989). Information and volatility: the no-arbitrage Martingale approach to timing and resolution irrelevancy. *Journal of Finance, 47*, 1–17.

Ross, S. (1995). Hedging Long-Run Commitments: Exercises in Incomplete Market Pricing. Working Paper, Yale School of Management.

Ross, S. (1977b). The Determination of financial structure: The incentive signaling approach. *Bell Journal of Economics, 8*, 23–40.

Ross, S. A. (1976b). Arbitrage Theory of Capital Asset Pricing. *Journal of Economic Theory, 4*, 341–360.

Ross, S. A. (1978). Mutual Fund Separation in Financial Theory-the Separating Distribution. *Journal of Economic Theory, 17*, 254–286.

Ross, S. A. (1977). Return, Risk, and Arbitrage. In Risk and Return in Finance, Vol. 1 (Cambridge, MA: Ballinger), pp. 187–208.

Rotch, W. (1990). Activity-based costing in service industries. *Journal of Cost Management*, 4–14.

Roth, G., & McDonald, C. G. (1999). Shareholder-management conflict and event risk covenants. *Journal of Financial Research, 22*, 207–225.

Rothenberg, T. J., Griliches, Z., & Intriligator, M. D. (1984). Approximating the distributions of econometric estimators and test statistics. In: Handbook of Econometrics, Volume II, Handbooks in Economics Series, Elsevier/North Holland.

Rothman, P. (1999). *Nonlinear Time Series Analysis of Economic and Financial Data*. Boston, NJ: Kluwer Academic Publishers.

Rothschild, M., & Stiglitz, J. E. (1969). Increasing risk: I. A definition. *Journal of Economic Theory, 2*, 225–243.

Roy, A. D. (1952). Safety first and the holding of assets. *Econometrica, 20*, 431–449.

Roy, A. D. (1951). Some thoughts on the distribution of earnings. *Oxford Economic Papers, 3*, 135–146.

Rozeff, M. S. (1982). Growth, beta and agency costs as determinants of dividend payout ratios. *Journal of Financial Research, 5*, 249–259.

Rua, A., & Nunes, L. (2009). International comovement of stock market returns: A wavelet analysis. *Journal of Empirical Finance, 16*, 632–639.

Rubinstein, M. (1976). The valuation of uncertain income stream and the pricing of options. *The Bell Journal of Economics, 7*, 407–425.

Rubinstein, M. (1994). Implied binomial trees. *Journal of Finance, 49*, 771–818.

Rubinstein, M. (1985). Nonparametric Tests of Alternative Option Pricing Models Using All Reported Trades and Quotes on the 30 Most Active CBOE Options Classes From August 23, 1976 Through August 31, 1978, *Journal of Finance*, 455–480.

Ruefli, T. W. (1971). A generalized goal decomposition model. Management Science, 505–518.

Ruey, T. S. (2010). *Analysis of Financial Time Series*. New York: John Wiley & Sons.

Ruppert, D., & Wand, M. P. (1995). Multivariate locally weighted least squares regression. *The Annal of Statistics, 23*, 1346–1370.

Ruppert, D., Sheather, S. J., & Wand, M. P. (1995). An effective bandwidth selector for local least squares regression. *Journal of the American Statistical Association, 90*, 1257–1270.

Ruszczynski, A., & Vanderbei, R. J. (2003). Frontiers of stochastically nondominated portfolios. *Econometrica, 71*, 1287–1297.

Ruszczynski, A. (1986). A regularized decomposition method for minimizing a sum of polyhedral functions. *Mathematical Programming, 35*, 309–333.

Ryan, H. E., & Wiggins, R. A. (2001). The influence of firm- and manager-specific characteristics on the structure of executive compensation. *Journal of Corporate Finance, 7*, 101–123.

Ryan, H., & Wiggins, R. (2004). Who is whose pocket? Director compensation, board independence, barriers to effective monitoring. *Journal of Financial Economics, 73*, 497–524.

Rydberg, T. H. (1997). The normal inverse Gaussian Levy process: simulation and approximation. *Communications in Statistics: Stochastic models, 13*, 887–910.

Rydberg, T. H. (1999). Generalized hyperbolic diffusion processes with applications in finance. *Mathematical Finance, 9*, 183–201.

Saaty, T. L. (1980). *The analytic hierarchy process*. New York: McGraw-Hill.
Sadeghi, M., & Shavvalpour, S. (2006). Energy risk management and value at risk modeling. *Energy Policy, 34*, 3367–3373.
Sadhak, H. (2003). *Mutual Funds in India: Marketing Strategies and Investment Practices*. New Delhi: Response Books.
Sadorsky, P. (1999). Oil price shocks and stock market activity. *Energy Economics, 21*, 449–69.
Sahut, J. M., Gharbi, S., & Gharbi, H. O. (2011). Stock Volatility, Institutional Ownership and Analyst Coverage. Bankers Markets & Investors 110 (Janvier-Febrier).
Saikkonen, P. (1991). Asymptotically efficient estimation of cointegration regressions. *Econometric Theory, 7*, 1–21.
Samarakoon, L. P., & Hasan, T. (2005). Portfolio performance evaluation. In C.F. Lee (ed.) *The Encyclopedia of Finance*. Springer, 617–622.
Samuelson, P. (1965). Proof that Properly Anticipated Prices Fluctuate Randomly. *Industrial Management Review, 31*, 41–49.
Samuelson, W., & Rosenthal, L. (1986). Price movements as indicators of tender offer success. *Journal of Finance, 41*, 481–499.
Sander, F. (1963). On subjective probability forecasting. *Journal of Meteorology, 2*, 191–201.
Sander, H., & Polasky, S. (2009). The value of views and open space: Estimates from a hedonic pricing model for Ramsey County, Minnesota, USA. *Land Use Policy, 26*(3), 837–845.
Santa-Clara, P., & Sornette, D. (2001). The Dynamics of the Forward Interest Rate Curve with Stochastic String Shocks. *Review of Financial Studies, 14*, 2001.
Santhanam, R., & Hartono, E. (2003). Issues in linking information technology capability to firm performance. *MIS Quarterly, 27*, 125–153.
Santhanam, R., Muralidhar, K., & Schniederjans, M. (1989). Zero-one goal programming approach for information system project selection. *Omega: International Journal of Management Science, 17*, 583–598.
Santos, J. A. (2001). Bank capital regulation in contemporary banking theory: A review of the literature. *Financial Markets, Institutions and Instruments, 10*, 41–84.
Sanz, J. (1999). Los planes de pensiones y jubilación como sistema complementario de las pensiones. *Acciones e investigaciones sociales, 9*, 157–166.
Sarbanes-Oxley Act. (2002). Public Law 107-204, 107th Cong., 2nd Session. Washington, DC: GPO.
Sargan, J. D. (1958). The estimation of economic relationships using instrumental variables. *Econometrica, 26*, 393–415.
Sarkar, S., & Sriram, R. S. (2001). Bayesian Models for Early Warning of Bank Failures. *Management Science, 47*(11), 1457–1475.
Sarma, M., Thomas, S., & Shah, A. (2000). Performance evaluation of alternative VaR models. Working paper, Indira Gandhi Institute of Development Research.
Sarno, L., & Valente, G. (2000). The cost of carry model and regime shifts in stock index futures markets: An empirical investigation. *Journal of Futures Markets, 20*, 603–624.
Sarno, L., Valente, G., & Leon, H. (2006). Nonlinearity in deviations from uncovered interest parity: An explanation of the forward bias puzzle. *Review of Finance, 10*, 443–482.
Saunders, A. (1999). Financial Institutions Management: A Modern Perspective, Irwin Series in Finance.
Saunders, A., & Yourougou, P. (1990). Are Banks Special? The Separation of Banking from Commerce and Interest Rate Risk. *Journal of Economics and Business, 42*, 171–182.
Savu, C., & Trede, M. (2010). Hierarchies of Archimedean Copulas. *Quantitative Finance, 10*, 295–304.
Schachter, B. (1998). An Irreverent Guide to Value at Risk. Risks and Rewards (March), 17–18.
Scharfstein, D. S., & Stein, J. C. (1990). Herd behavior and investment. *American Economic Review, 80*, 465–479.
Schaumburg, J. (2011). Predicting extreme VaR: Nonparametric quantile regression with refinements from extreme value theory, Discussion Paper 2010-009, CRC 649, Humboldt Universität zu Berlin.

Scheetz, T. E., Kim, K.-Y. A., Swiderski, R. E., Philp1, A. R., Braun, T. A., Knudtson, K. L., Dorrance, A. M., DiBona, G. F., Huang, J., Casavant, T. L., Shefield, V. C. & Stone, E. M. (2006). Regulation of gene expression in the mammalian eye and its relevance to eye disease. *Proceeding of the National Academy of Sciences, 103*, 14429–14434.

Scheinkman, J. A., & Xiong, W. (2003). Overconfidence and speculative bubbles. *Journal of Political Economy, 111*, 1183–1219.

Schere, F. M. (1980). Industrial Market Structure and Economic Performance. Houghton Mofflin.

Schink, G. R., & Bower, R. S. (1994). Application of the Fama-French Model to Utility Stocks. *Financial Markets, Institutions & Instruments. Estimating Cost of Capital: Methods and Practice, 3*(3), 74–96.

Schipper, K. (1991). Commentary on analysts' forecasts. *Accounting Horizons, 5*, 105–121.

Schipper, K., & Thompson, R. (1981). Common stocks as hedges against shifts in the consumption on investment opportunity set. *Journal of Business, 54*, 305–328.

Schleicher, C. (2002). An Introduction to Wavelets for Economists. Working Paper, Bank of Canada.

Schlicht, E. (2005). Estimating the smoothing parameter in the socalled Hodrick-Prescott filter. *Journal of the Japan Statistical Society, 35*(1), 99–119.

Schmidt, A. (2005). Quantitative Finance for Physicists: An Introduction. Academic Press Advanced Finance Series.

Schniederjans, M. J., & Wilson, R. L. (1991). Using the analytic hierarchy process and goal programming for information system project selection. *Information and Management, 20*, 333–342.

Schniederjans, M. J., & Garvin, T. (1997). Using the analytic hierarchy process and multi-objective programming for the selection of cost drivers in activity-based costing. *European Journal of Operational Research, 100*, 72–80.

Scholes, M., & Williams, J. (1977). Estimating Beta from Nonsynchronous Trading. *Journal of Financial Economics, 5*, 309–327.

Schönbucher, P. & Schubert, D. (2001). Copula dependent default risk in intensity models. Working paper, Department of Statistics, Bonn University.

Schönbucher, P. J. (2003). Credit derivatives pricing models. Wiley Finance.

Schroder, M. (1989). Computing the constant elasticity of variance option pricing formula. *Journal of Finance, 44*(1), 211–219.

Schultz, P. (2003). Pseudo market timing and the long-run underperformance of IPOs. *Journal of Finance, 58*, 483–517.

Schultz, P., & Zaman, M. (2001). Do the individuals closest to internet firms believe they are overvalued? *Journal of Financial Economics, 59*, 347–381.

Schwartz, E. S., & Trigeogis, L. (2004). Real Options and Investment under Uncertainty: Classical Readings and Recent Contributions. MIT Press.

Schwartz, G. (1978). Estimating the dimension of a model. *Annals of Statistics, 6*, 461–464.

Schwartz, R., & Weber, B. (1997). Next-Generation Securities Market Systems: An Experimental Investigation of Quote-Driven and Order-Driven Trading. *Journal of Management Information Systems*, 57–79.

Schwert, G. W. (1989). Why does stock market volatility change over time. *Journal of Finance, 44*, 1115–1153.

Scott, J. (1976). A theory of optimal capital structure. *Bell Journal of Economics, 7*, 33–54.

Scott, L. (1997). Pricing stock options in a jump diffusion model with stochastic volatility and interest rates: applications of Fourier inversion methods. *Mathematical Finance, 7*, 413–424.

Scott, L. (1987). Option Pricing When the Variance Changes Randomly: Theory, Estimators, and Applications. *Journal of Financial and Quantitative Analysis, 22*, 419–438.

Scott, W. (1979). On optimal and data-based histograms. *Biometrika, 33*, 605–610.

Scruggs, J. T. (1998). Resolving the puzzling intertemporal relation between the market risk premium and conditional market variance: a two-factor approach. *Journal of Finance, 52*, 575–603.

Sealey, C. W., Jr. (1979). Credit rationing in the commercial loan market: Estimates of a structural model under conditions of disequilibrium. *Journal of Finance, 34*, 689–702.

Sears, R. S., & John Wei, K. C. (1988). The structure of skewness preferences in asset pricing model with higher moments: an empirical test. *Financial Review, 23*(1), 25–38.

Securities Regulation Committee. (1997). Department of Treasury, series 00588, volume 6, 1997.

Segal, M. R., Dahlquist, K. D., & Conklin, B. R. (2003). Regression approach for microarray data analysis. *Journal of Computational Biology, 10*, 961–980.

Segen, J., & Sanderson, A. (1980). Detection of Changes in a Time Series. *IEEE Transaction in Information Theory, 26*, 249–257.

Seifert, J., & Uhrig-Homburg, M. (2007). Modelling jumps in electricity prices: Theory and empirical evidence. *Review of Derivatives Research, 10*(1), 59–85.

Seiford, L., & Yu, P. L. (1979). Potential solutions of linear systems: the multi-criteria multiple constraint level program. *Journal of Mathematical Analysis and Applications*, 283–303.

Sentana, E., & Wadhwani, S. (1992). Feedback Traders and Stock Return Autocorrelations: Evidence from A Century of Daily Data. *Economic Journal, 102*, 415–425.

Sentana, E. (1995). Quadratic ARCH models. *Review of Economic Studies, 62*, 639–661.

Serfling, R. J. (1980). *Approximation Theorems of Mathematical Statistics*. New York: John Wiley & Sons Inc.

Seth, A. (1990). Value Creation in Acquisitions: A Re-examination of Performance Issues. *Strategic Management Journal, 11*, 90–115.

Shanken, J. (1982). The arbitrage pricing theory: Is it testable? *Journal of Finance, 37*, 1129–1140.

Shanken, J., & Weinstein, M. I. (2006). Economic forces and the stock market revisited. *Journal of Empirical Finance, 13*, 129–144.

Shanken, J. (1987). Multivariate proxies and asset pricing relations: Living with the Roll critique. *Journal of Financial Economics, 18*, 91–110.

Shao, J. (1997). An Asymptotic Theory for Linear Model Selection. *Statistica Sinica, 7*, 221–264.

Shao, X., Lian, Y., & Yin, L. (2009). Forecasting value-at-risk using high frequency data: The realized range model. *Global Finance Journal, 20*, 128–136.

Shapiro, A. F. (2005). Fuzzy regression models. Working Paper, Penn State University.

Shapour, M., Raeei, R., Ghalibaf, H., & Golarzi, H. (2010). Analysis of herd behavior of investors in Tehran stock exchange using with state-space model. *Journal of financial accounting research summer, 2*(4), 49–60.

Sharifi, S., Crane, M., Shamaie, A., & Ruskin, H. (2004). Random matrix theory for portfolio optimization: a stability approach. *Physica A, 335*, 629–643.

Sharma, V. (2011). Stock return and product market competition: beyond industry concentration. *Review of Quantitative Finance and Accounting, 37*(3), 283–299.

Sharpe, W. F. (1964). Capital asset prices: A theory of market equilibrium under conditions of risk. *Journal of Finance, 19*, 425–442.

Sharpe, W. F. (1966). Mutual funds performance. *Journal of Business, 39*, 119–138.

Sharpe, W. F. (1971). A linear programming approximation for the general portfolio analysis problem. *Journal of Financial and Quantitative Analysis, 6*, 1263–1275.

Shaw, W. (2006). Stochastic volatility: models of Heston type, lecture notes. http://www.mth.kcl.ac.uk/~shaww/web_page/papers/StoVolLecture.pdf http://www.mth.kcl.ac.uk/~shaww/web_page/papers/StoVolLecture.nb

Shea, J. (1997). Instrument relevance in multivariate linear models: a simple measure. *Review of Economics and Statistics, 79*, 348–352.

Sheather, S. J., & Maritz, J. S. (1983). An estimate of the asymptotic standard error of the sample median. *Australian Journal of Statistics, 25*, 109–122.

Shefrin, H. (2000). *Beyond greed and fear; Understanding behavioral finance and the psychology of investing*. Cambridge: HBS Press.

Shehata, M. (1991). Self selection bias and the economic consequences of accounting regulations: An application of two-stage switching regression to SFAS No. 2. *The Accounting Review, 66*, 768–787.

Shen, W. T., & Huang, S. L. (2001). An empirical analysis of China's mutual funds: performance and evaluation. *Economic Research, 9*, 22–30 (in Chinese).

Shephard, N. (1996). Statistical aspects of ARCH and stochastic volatility. In D. R. Cox, D. -V. Hinkley, & O. E. Barndorff-Nielsen (Eds.), *Likelihood, Time Series with Econometric and Other Applications* (pp. 1–67). London: Chapman and Hall.
Shepherd, P. (1997). Likelihood analysis of Non-Gaussian measurement time series. *Biometrika, 84*(3), 653–667.
Sheppard, K. (2009). MFE MATLAB function reference. Mimeo.
Sheppard, S. (1999). Hedonic Analysis of Housing Markets. In Paul C. Chesire and Edwin S. Mills (eds.), *Handbook of Regional and Urban Economics*. vol. 3. Elsevier.
Shi, Y. (1991). Optimal linear production systems: models, algorithms, and computer support systems. Ph. D. dissertation, School of Business, University of Kansas.
Shi, Y., & Lee, H. (1992). A binary integer linear programming with multi-criteria and multi-constraint levels. Working paper #92-21. School of Business, University of Nebraska at Omaha.
Shi, Y., & Yu, P. L. (1992). Selecting optimal linear production systems in multiple criteria environments. *Computer and Operations Research, 19*, 585–608.
Shiller, R. J. (2000). *Irrational Exuberance*. Princeton, NJ: Princeton University Press.
Shiller, R. J. (1979). The volatility of long term interest rates and expectations models of the term structure. *Journal of Political Economy, 87*, 1190–1219.
Shiller, R. J. (1988). Fashions, fads, and bubbles in financial markets. In J. Coffee Jr., L. Lowenstein, & S. Ross-Ackerman (Eds.), *Knights, raiders, and targets*. New York: Oxford Press.
Shiller, R. (1984). Stock prices and social dynamics. *Brookings Papers on Economic Activity, 15*, 457–498.
Shiller, R. J. (1999). Human behavior and the efficiency of the financial system. In: J. B. Taylor & M. Woodford (ed.), *Handbook of Macroeconomics*, edition 1, volume 1, chapter 20, pages 1305-1340 Elsevier.
Shin, H., & Stulz, R. (1998). Are Internal Capital Markets Efficient? *Quarterly Journal of Economics, 113*, 531–552.
Shirai, S. (2004). Assessing the impact of financial and capital market reforms on firms' corporate financing patterns in India. *South Asia Economic Journal, 5*, 189–208.
Shivdasani, A. (1993). Board composition, ownership structure, and hostile takeovers. *Journal of Accounting & Economics, 16*, 167–188.
Shleifer, A., & Vishny, R. (1997). A survey of corporate governance. *Journal of Finance, 52*, 737–783.
Shleifer, A., & Vishny, R. W. (1990). Equilibrium short horizons of investors and firms. *American Economic Review, 80*, 148–153.
Shorack, G., & Wellner, J. (1986). *Empirical Processes with Applications to Statistics*. New York, Massachusetts: John Wiley & Sons.
Shreve, E. (2007). Stochastic Calculus for Finance II: Continuous-Time Models. Springer Finance.
Shreve, S. E. (2000). Stochastic Calculus for Finance II: Continuous-Time Models. Springer.
Shreve, S. E. (2008). *Stochastic Calculus for Finance II: Continuous-Time Models*. New York: Springer.
Shrieves, E., & Dahl, D. (1992). The relationship between risk and capital in commercial banks. *Journal of Banking and Finance, 16*, 439–457.
Shu, J. H., & Zhang, J. E. (2006). Testing range estimators of historical volatility. *Journal of Futures Markets, 26*, 297–313.
Shumway, T. (2001). Forecasting bankruptcy more accurately: A simple hazard model. *The Journal of Business, 74*, 101–124.
Shyam-Sunder, L. (1991). The stock price effect of risky versus safe debt. *Journal of Financial and Quantitative Analysis, 26*, 549–58.
Sias, R. W., & Starks, L. T. (1997). Return autocorrelation and institutional investors. *Journal of Financial Economics, 46*, 103–131.
Sibbertsen, P., & Kruse, R. (2009). Testing for a break in persistence under longrange dependencies. *Journal of Time Series Analysis, 30*, 268–285.

Siddiqui, M. (1960). Distribution of quantiles in samples from a bivariate population. *Journal of Research of the National Bureau of Standards, 64B*, 145–150.

Siegel, A. F., & Woodgate, A. (2007). Performance of portfolios optimized with estimation error. *Management Science, 53*, 1005–1015.

Siegel, J. J. (2003). What Is an Asset Price Bubble? An Operational Definition. *European Financial Management, 9*, 11–24.

Silvennoinen, A., & Teräsvirta, T. (2008). Multivariate autoregressive conditional heteroskedasticity with smooth transition in conditional correlations. SSE/EFI Working paper series in Economics and Finance.

Silvennoinen, A., & Teräsvirta, T. (2009). Modeling multivariate autoregressive conditional heteroskedasticity with the double smooth transition conditional correlations GARCH Model. *Journal of Financial Econometrics, 7*, 373–411.

Sim, N. (2012). Modeling the correlations of asset returns by the dependence between their quantiles. Working paper.

Simaan, Y. (1997). Estimation risk in portfolio selection: The mean variance model versus the mean absolute deviation model. *Management Science, 43*, 1437–1446.

Sims, C. A. (1972). Money, Income, and Causality. *American Economic Review, 62*, 540–552.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica, 48*, 1–48.

Singh, P., & Montegomery, T. (1987). Using Simulated Mergers to Evaluate Corporate Diversification Strategies. *Strategic Management Journal, 8*, 377–386.

Singhal, J., Marsten, R. E., & Morin, T. L. (1989). Fixed order brand-and-bound methods for mixed-integer programming: the ZOOM system. *ORSA Journal on Computing, 1*, 44–51.

Singleton, K. J. (2006). *Empirical Dynamic Asset Pricing - Model Specification and Econometric Assessment*. Princeton and Oxford: Princeton University Press.

Skinner, D. J. (2008). The evolving relation between earnings, dividends, and stock repurchases. *Journal of Financial Economics, 87*, 582–609.

Sklar, A. (1959). Fonctions de repartition a n dimensions et leurs marges. *Pub. Inst. Statist. Univ. Paris, 8*, 229–231.

Sloan, R. (1996). Do stock prices fully reflect information in accruals and cash flows about future earnings? *Accounting Review, 71*, 289–315.

Smarr, L., & Catlett, C. (1992). Metacomputing. *Communications of the ACM archive, 35*(6), 44–52.

Smirlock, M., & Starks, L. (1985). A further examination of stock price changes and transactions volume. *Journal of Financial Research, 8*, 217–225.

Smirlock, M., & Marshall, W. (1983). An Examination of the empirical relationship between the dividend and investment decisions: a note. *Journal of Finance, 38*, 1659–1667.

Smirnov N. (1939). On the Estimation of the Discrepancy Between Empirical Curves of Distribution for Two Independent Samples. Bulletin Mathematique de l.Universite.de Moscou, 2, fasc. 2.

Smith, C., & Stulz, R. (1985). The determinants of firms' hedging policies. *Journal of Financial and Quantitative Analysis, 20*, 391–405.

Smith, C., & Watts, R. (1992a). The investment opportunity set and corporate financing, dividend, and compensation policies. *Journal of Financial Economics, 32*(3), 117–61.

Smith, C. W., Jr., & Watts, R. L. (1992b). The investment opportunity set and corporate financing, dividend, and compensation policies. *Journal of Financial Economics, 32*, 263–292.

Smith, C., & Watts, R. (1992c). The investment opportunity set and corporate financing, dividend and compensation policies. *Journal of Financial Economics, 32*, 263–292.

Smith, D. M. (2009). The Economics of Mutual Funds, Chapter-3. In J. A Haslem (Ed.) *A Companion to Mutual Funds*, John Wiley Sons, USA

Smith, M. (2000). Modelling and Short-Term Forecasting of New South Wales Electricity System Load. *Journal of Business & Economic Statistics*, Oct 2000, 465–478. 4.

Smith, P., & Reither, C. (1996). Comprehensive income and the effect of reporting it. Financial Analysts Journal (November/December) 14–19.

Smith, V., & Williams, A. (1992). Experimental Market Economics. Scientific American, December 1992, pp. 116–121.

Smith, V., Suchanek, G., & Williams, A. (1988). Bubbles, crashes, and endogenous expectations in experimental spot asset markets. *Econometrica, 56*, 1119–1151.

Smithson, C., & Minton, L. (1996a). Value-at-Risk(2):The debate on the use of VaR. Risk, 9(February).

Smithson, C., & Minton, L. (1996b). Value-at-Risk(1): The debate on the use of VaR. *Risk, 9*, 25–27.

Snedecor, G. W., & Cochran, W. G. (1989). *Statistical Methods*. Ames, Iowa: Iowa State Press.

Snir, M, et al. (1996). MPI: The Complete Reference. Massachusetts Institute of Technology.

So, M. K. P., & Yu, P. L. H. (2006). Empirical analysis of GARCH models in value at risk estimation. *International Financial Markets, Institutions & Money, 16*, 180–197.

So, M. K. P., Lam, K., & Li, W. K. (1998). A stochastic volatility model with Markov switching. *Journal of Business and Economic Statistics, 16*, 244–253.

Sobehart, J. R., Keenan, S. C., & Stein, R. M. (2000). *Benchmarking Quantitative Default Risk Models: A Validation Methodology*. New York, NY: Moody's Investors Service.

Sobol, I. M. (1967). On the Distribution of Points in a Cube and the Approximate Evaluation of Integers. *U.S.S.R. Computational Mathematics and Mathematical Physics, 7*(4), 86–112.

Sodal, S., Koekebakker, S., & Aadland, R. (2008). Market Switching in Shipping: A Real Option Model Applied to the Valuation of Combination Carriers. *Review of Financial Economics, 17*, 183–203.

Solnik, B. (1983). The relation between stock prices and inflationary expectations: The international evidence. *Journal of Finance, 38*, 35–48.

Solnik, B. (1996). *International Investments* (3rd ed.). New York: Addison-Wesley.

Solomon, D. (2005). Accounting rule exposes problems but draws complaints about cost. Wall Street Journal, May 2, p. A.1.

Solomon, D., & Frank, R. (2003). Stock analysis: You don't like our stock? You are off the list – SEC sets new front on conflicts by taking aim at companies that retaliate against analysts. Wall Street Journal, Jun 19, p. C.1.

Sondhi, H. J. (2007). *Financial Performance of Equity Mutual Funds in India*. New Delhi: Deep & Deep.

Song, F. (1994). A Two Factor ARCH Model for Deposit-institution Stock Returns. *Journal of Money Credit and Banking, 26*, 323–340.

Song, H., Liu, X., & Romilly, P. (1996). A time varying parameter approach to the Chinese aggregate consumption function. *Economics of Planning, 29*, 185–203.

Sonlinik, B., Boucrelle, C., & Le Fur, Y. (1996). International market correlation and volatility. *Financial Analysts Journal, 52*(5), 17–34.

Sorensen, M. (2007). How smart is smart money? a two-sided matching model of venture capital. *The Journal of Finance, 62*, 2725–2762.

Sortino, F. A., & van der Meer, R. (1991). Downside risk. *Journal of Portfolio Management, 17*, 27–31.

Speranza, M. G. (1993). Linear programming models for portfolio optimization. *Finance, 14*, 107–123.

Spiegel, Y., & Wilkie, S. (1996). Investment in a new technology as a signal of firm value under regulatory opportunism. *Journal of Economics and Management Strategy, 5*, 251–76.

Spitz, A. E. (1970). Mutual fund performance and cash inflows. *Applied Economics, 2*(2), 141–145.

Sriboonchitta, S., Wong, W. K., Dhompongsa, D., & Nguyen, H. T. (2009). *Stochastic Dominance and Applications to Finance, Risk and Economics*. Boca Raton, Florida: Chapman and Hall/CRC.

Srivastava, S. C., Hamid, S., & Choudhury, A. H. (1999). Stock and Bond Market Linkage in the Empirical Study of Interest Rate Sensitivity of Bank Returns. *Journal of Applied Business Research, 15*, 47–58.

Staiger, D., & Stock, J. H. (1997). Instrumental variables regression with weak instruments. *Econometrica, 65*, 557–586.

Stambaugh, R. F. (1999). Predictive regressions. *Journal of Financial Economics, 54*, 375–421.

Stambough, R. (1982). On the exclusion of assets from tests of the two-parameter model: A sensitivity analysis. *Journal of Financial Economics, 10*, 237–268.

Standard & Poor's. (1999). Sovereign Ratings Display Stability over Two Decades. Credit Week.

Standard & Poor's. (2001). Sovereign Ratings History since 1975. New York, NY.

Standard & Poor's. December 12 (2001). Playing out the credit cliff dynamics. Standard & Poor's. New York [Re-printed Oct. 7, 2004]

Standard & Poor's Corporation. (1996). Standard & Poor's Corporation Ratings Criteria. McGraw Hill.

Stangeland, D., & Zheng, S. (2007). IPO underpricing, firm quality, and analyst forecasts. *Financial Management, 36*, 1–20.

Stata. (2008). Stata version 10. College Station, TX.

Statman, M. (2000). Socially responsible Mutual Funds. *Financial Analysts Journal, 56*, 30–38.

Statman, M. (2002). Lottery players/stock traders. Journal of Financial Planning January/February, 14–21.

Stavroyiannis, S., Makris, I., Nikolaidis, V., & Zarangas, L. (2012). Econometric modeling and value-at-risk using the Pearson type-IV distribution. *International Review of Financial Analysis, 22*, 10–17.

Stein, R. M. (2002). *Benchmarking Default Prediction Models: Pitfalls and Remedies in Model Validation*. New York, NY: Moody's KMV.

Stein, C. (1972). A Bound for the Error in the Normal Approximation to the Distribution of a Sum of Dependent Random Variables. Proceedings of the Sixth Berkeley Symposium on Mathematical Statistics and Probability, University of California Press, Berkeley, 583–602.

Stein, E., & Stein, J. (1991). Stock Price Distributions With Stochastic Volatility. *Review of Financial Studies, 4*, 727–752.

Stein, J. (2003). Agency Information and Corporate Investment. In G. Constantinides, M. Harris, & R. Stulz (Eds.), *Handbook of the Economics of Finance*. Amsterdam: Elsevier, North-Holland.

Stein, J. L. (1973). The simultaneous determination of spot and futures prices. *American Economic Review, 51*, 1012–1025.

Stein, J. C. (1988). Takeover threats and managerial myopia. *Journal of Political Economy, 96*, 61–80.

Steinitz, J. (2006). Gold: volatility, correlation and teacups. World Gold Council.

Stephan, J., & Whaley, R. (1990). Intraday price changes and trading volume relations in the stock and stock option markets. *Journal of Finance, 45*, 191–220.

Stephens, C., & Weisbach, M. (1998). Actual share reacquisition in open-market repurchase programs. *Journal of Finance, 53*, 313–333.

Stephens, M. A. (1974). EDF statistics for goodness of fit and some Comparisons. *Journal of the American Statistical Association, 69*, 730–737.

Steuer, R. (1986). *Multiple criteria optimization: theory, computation and application*. New York: John Wiley & Sons.

Stevenson, B. G., & Fadil, M. W. (1995). Modern Portfolio Theory: Can It Work for Commercial Loans? *Commercial Lending Review, 10*(2), 4–12.

Stevenson, S. (2001). Emerging markets, downside risk and the asset allocation decision. *Emerging Markets Review, 2*, 50–66.

Stickel, S. (1991). Common stock returns surrounding earnings forecast revisions: More puzzling evidence. *The Accounting Review, 66*, 402–416.

Stiglitz, J. (1990). Symposium on bubbles. *Journal of Economic Perspective, 4*, 13–18.

Stober, T. (1999). Empirical Applications of the Ohlson [1995] and Feltham-Ohlson [1995,1996] Valuation Models. *Managerial Finance, 25*(2), 3–16.

Stock, J. H. (1994). Unit roots, structural breaks and trends. In Handbook of conometrics Chapter 46, 2740–2841.

Stock, J. H., & Yogo, M. (2002). Testing for weak instruments in linear IV regression. National Bureau of Economic Research, NBER Technical Working Paper: 284.

Stock, J. H., & Watson, M. W. (1998). A Dynamic Factor Model Framework for Forecast Combination. Manuscript, Kennedy School, Harvard University and Woodrow Wilson School, Princeton University.

Stock, J. H., Wright, J. H., & Yogo, M. (2002). A survey of weak instruments and weak identification in generalized method of moments. *Journal of Business and Economic Statistics, 20*, 518–529.
Stock, J. H., & Watson, M. W. (2002). Forecasting using principal components from a large number of predictors. *Journal of the American Statistical Association, 97*, 1167–1179.
Stock, J., & Watson, M. (2001). Vector autoregressions. *Journal of Economic Perspectives, 15*, 101–115.
Stockman, A. C. (1987). Economic theory and exchange rate forecasts. *International Journal of Forecasting, 3*, 3–15.
Stokes-Rees, I., et al. (2007). Multi-cluster parallel job submission: Experiences with Monte Carlo simulations for computational finance on Grid5000. IRIS PicsouGrid Project Report.
Stoll, H. R. (1976). Dealer Inventory Behavior: An Empirical Investigation of NASDAQ Stocks. *Journal of Financial and Quantitative Analysis, 11*, 356–380.
Stone, B. (1974). Systematic Interest Rate Risk in a Two-Index Model of Returns. *Journal of Financial and Quantitative Analysis, 9*, 709–721.
Stowe, J. D. (1978). Life Insurance Company Portfolio Behavior. *Journal of Risk and Insurance, 45*, 431–447.
Strandell, E., Tilak, S., Chou, H. M., Wang, Y. T., Lin, F. P., Arzberger, P., & Fountain, T. (2007). Data Management at Kenting's Underwater Ecological Observatory. In Proceedings of the third International Conference on Intelligent Sensors, Sensor Networks and Information Processing.
Strebulaev, I. (2002). Liquidity and asset pricing: Evidence from the US Treasury securities market. Working paper, London Business School.
Strong, N., & Walker, M. (1993). The Explanatory Power of Earnings for Stock Returns. *The Accounting Review, 68*, 385–399.
Struzik, Z. (2001). Wavelet methods in financial time-series processing. *Physica A, 296*, 307–319.
Stulz, R. (1982). Options on the minimum or the maximum of two risky assets: Analysis and applications. *Journal of Financial Economics, 10*(2), 161–185.
Stulz, R., & Wasserfallen, W. (1995). Foreign Equity Investment Restrictions, Capital Flight, and Shareholder Wealth Maximization: Theory and Evidence. *Review of Financial Studies, 8*, 1019–1057.
Stulz, R. (1995). International Portfolio Choice and Asset Pricing: An Integrative Survey. In R. Jarrow, V. Maksimovic, & W. T. Ziemba (Eds.), *Handbooks in Operation Research and Management Science* (Vol. 9). Amsterdam: North-Holland.
Su, J. B., & Hung, J. C. (2011). Empirical analysis of jump dynamics, heavy-tails and skewness on value-at-risk estimation. *Economic Modelling, 28*, 1117–1130.
Su, Y. (1996). Comparison of relative accuracy of management and analysts earnings forecast. Unpublished thesis, Accounting Department, Dong Wu University.
Subbotin, A. (2008). A multi-horizon scale for volatility. Working Paper, University of Paris-1.
Subha, M. V., & Bharati, J. S. (2007). An Empirical study on the performance of select mutual fund schemes in India, *Journal of Contemporary Research in Management, 1*(1), 2 Jan-June.
Sudhakar, A., & Sasi, K. K. (2005). Performance Evaluation of Mutual Funds: A Case Study. Southern Economist, November, pp. 19–23.
Sufi, A. (2007). Information asymmetry and financing arrangements: evidence from syndicated loans. *Journal of Finance, 62*, 629–668.
Suh, C. K., Suh, E. H., & Choi, J. (1993). A two-phased DSS for R&D portfolio selection. Proceedings of International DSI, 351–354.
Sullivan, R., Timmermann, A., & White, H. (1999). Data-Snooping, Technical Trading Rule Performance, and the Bootstrap. *The Journal of Finance, 54*(5), 1647–1691.
Sullivan, R., Timmermann, A., & White, H. (2001). Dangers of data-driven inference: The case of calendar effects in stock returns. *Journal of Econometrics, 249*, 286.
Sun, C. (2007). Variation of federal cost-share programs in the United States and the inducement effects on tree planting. *Journal of Forest Economics, 12*, 279–296.

Sun, C., & Zhang, D. (2001). Assessing the financial performance of forestry-related investment vehicles: Capital Asset Pricing Model vs. Arbitrage Pricing Theory. *American Journal of Agricultural Economics, 83*, 617–628.

Sun, E. W., & Meinl, T. (2012). A new wavelet-based denoising algorithm for high-frequency financial data mining. *European Journal of Operational Research, 217*(3), 589–599.

Sun, J., & Qiu, P. (2007). Jump detection in regression surfaces using both first-order and second-order derivatives. *Journal of Computational and Graphical Statistics, 16*(2), 289–311.

Sun, W., Rachev, S. T., Fabozzi, F., & Stoyan, S. (2008). Multi-variate skewed student's t copula in analysis of nonlinear and asymmetric dependence in German equity market. Studies in Nonlinear Dynamics and Econometrics, 12(2).

Sun, W., Rachev, S. T., Chen, Y., & Fabozzi, F. J. (2009). Measuring intra-daily market risk: a neural network approach. Technical Report, Chair of Econometrics, Statistics and Mathematical Finance School of Economics and Business Engineering, University of Karlsruhe.

Sunder, S. (1973). Relationship Between Accounting Changes and Stock Prices: Problems of Measurement and Some Empirical Evidence. In Empirical Research in Accounting: Selected Studies, Journal of Accounting Research, 1–45.

Sunder, S. (1992). Experimental asset markets: A survey. In J. H. Kagel & A. E. Roth (Eds.), *Handbook of Experimental Economics*. Princeton, NY: Princeton University Press.

Suntheim, F. (2012). CEO overconfidence in Banking. Working paper, Bocconi University.

Sutton, C. D. (1993). Computer-intensive methods for tests about the mean of an asymmetrical distribution. *Journal of the American Statistical Association, 88*, 802–808.

Sweeney, R. J., & Warga, A. D. (1986). The Pricing of Interest-Rate Risk: Evidence from the Stock Market. *Journal of Finance, 41*, 393–410.

Switzer, L. (1984). The determinants of industrial R&D: A funds flow simultaneous equation approach. *The Review of Economics and Statistics, 66*, 163–168.

Sy, A. N. R. (2004). Rating the Rating Agencies: Anticipating Currency Crises or Debt Crises? *Journal of Banking and Finance, 28*(11), 2845–2867.

Talmor, S. (1996). History repeat itself. THE BANKER (March), 75–76.

Tan, L., Chiang, T. C., Mason, J. R., & Nelling, E. (2008). Herding Behavior in Chinese Stock Markets: An Examination of A and B Shares. *Pacific-Basin Finance Journal, 16*, 61–77.

Tanaka, S. (2003). Joint Development Project Applies Grid Computing Technology to Financial Risk Management. Nikkei Computer, special section on grid computing.

Tang, G. Y. N. (1998). Monthly Pattern and Portfolio Effect on Higher Moments of Stock Returns: Empirical Evidence from Hong Kong. *Asia-Pacific Financial Markets, 5*, 275–307.

Tang, L. T., & Shieh, J. S. (2006). Long memory in stock index futures markets: a value-at-risk approach. *Physica A: Statistical Mechanics and its Applications, 366*, 437–448.

Tang, R. Y. W. (1992). Transfer pricing in the 1990s. Management Accounting, 22–26.

Tapinos, E., Dyson, R. G., & Meadows, M. (2011). Does the balanced scorecard make a difference to the strategy development process? *The Journal of the Operational Research Society, 62*, 888–899.

Tarjan, R. E. (1972). Depth-First Search and Linear Graph Algorithms. *SIAM Journal on Computing, 1*, 146–160.

Tastan, H. (2006). Estimating time-varying conditional correlations between stock and foreign exchange markets. *Physica A, 360*, 445–458.

Tatsuoka, M. M. (1988). *Multivariate Analysis: Techniques for Educational and Psychological Research* (2nd ed.). New York: John Wiley & Sons, Inc.

Tauchen, G., & Pitts, M. (1983). The price variability-volume relationship on speculative markets. *Econometrica, 51*, 485–505.

Taylor, J. W. (2008a). Using exponentially weighted quantile regression to estimate value at risk and expected shortfall. *Journal of Financial Econometrics, 6*, 382–406.

Taylor, S. J. (2008b). *Modelling Financial Time Series* (2nd ed.). Singapore: World Scientific Publishing.

Taylor, J. W., & Snyder, R. D. (2012). Forecasting intraday time series with multiple seasonal cycles using parsimonious seasonal exponential smoothing. *OMEGA, 40*, 748–757.
Taylor, S. (1986). *Modelling Financial Time Series*. New York: John Wiley and Sons.
Taylor, S. J. (1994). Modeling stochastic volatility: a review and comparative study. *Mathematical Finance, 4*, 183–204.
Taylor, S. J. (2005). *Asset Price Dynamics, Volatility, and Prediction*. Princeton: Princeton University Press.
Teverovsky, V., & Taqqu, M. (1997). Testing for long -range dependence in the presence of shifting means or a slowly declining trend, using a variance-type estimation. *Journal of Time Series Analysis, 18*, 279–304.
Tewari, M., & Ramanlal, P. (2010). Is the put option in U.S. structured bonds good news for both bondholders and stockholders? *International Research Journal of Finance and Economics, 52*, 50–59.
Thaler, R. (1991). *Quasi-rational economics*. New York: Russel Sage Foundation.
Thanassoulis, E. (1985). Selecting a suitable solution method for a multi-objective programming capital budgeting problem. *Journal of Business Finance and Accounting, 12*, 453–471.
Thanou, E. (2008). *Mutual Fund Evaluation During Up and Down Market Conditions: The Case of Greek Equity Mutual Funds"*. Issue: International Research Journal of Finance and Economics. 13.
The International Monetary Fund. (2001). World Economic Outlook. Washington, DC.
Theil, H. (1966). *Applied Economic Forecasting*. Amsterdam: North-Holland.
Theil, H. (1958). *Economic Forecasts and Policy*. Amsterdam: North-Holland.
Theil, H. (1971). *Principal of Econometrics*. New York: John Wiley.
Theodossiou, P. (1998). Financial Data and the Skewed Generalized t Distribution. *Management Science, 44*, 1650–1661.
Thomas, A. L. (1980). A behavioral analysis of joint-cost allocation and transfer pricing, Stipes Publishing.
Thomas, S. (2002). Firm diversification and asymmetric information: Evidence from analysts' forecasts and earnings announcements. *Journal of Financial Economics, 64*, 373–396.
Thomas, S., Ramiah, V., Heaney, R., & Mitchell, H. (2011). Seasonal factors and outlier effects in rate of return in Australia's National Electricity Market. *Applied Economics, 43*(3), 355–369.
Thompson, H. E. (1979). Estimating the Cost of Equity Capital for Electric Utilities: 1958–1976. *Bell Journal of Economics, 10*(2), 619–635.
Thompson, S. B. (2008). Identifying Term Structure Volatility from the LIBOR-swap Curve. *Review of Financial Studies, 21*, 819–854.
Thompson, S. B. (2011). Simple formulas for standard errors that cluster by both firm and time. *Journal of Financial Economics, 99*(1), 1–10.
Thompson, S. B. (2010). Simple formulas for standard errors that cluster by both firm and time, Working paper.
Thompson, S. Identifying Term Structure Volatility from the LIBOR-Swap Curve. Review of Financial Studies, 21, 819–854.
Thrane, C. (2009). Explaining Variation In Wine Prices: The Battle Between Objective And Sensory Attributes Revisited. *Applied Economics Letters, 16*(13), 1383–1386.
Thursby, J. G. (1985). The relationship among the specification error tests of Hausman, Ramsey and Chow. *Journal of the American Statistical Association, 80*(392), 926–928.
Tibshirani, R. (1996). Regression shrinkage and selection via LASSO. *Journal of the Royal Statistical Society, Series B, 58*, 267–288.
Tilak, S., Hubbard, P., Miller, M., & Fountain, T. (2007). The Ring Buffer Network Bus (RBNB) DataTurbine Streaming Data Middleware for Environmental Observing Systems. In Proc. Third IEEE International Conference on e-Science and Grid Computing, 125–133.
Tinbergen, J. (1939). The dynamics of share-price formation. *Review of Economics and Statistics, 21*, 153–160.

Tirole, J. (1982). On the possibility of speculation under rational expectations. *Econometrica, 50*, 1163–1187.

Titman, S. (1984). The effect of capital structure on a firm's liquidation decision. *Journal of Financial Economics, 13*, 137–151.

Titman, S., & Wessels, R. (1988). The determinants of capital structure choice. *Journal of Finance, 43*(1), 1–19.

Tobin, J. (1958). Estimation of relationships for limited dependent variables. Econometrica: Journal of the Econometric Society, 24–36.

Tobin, J. (1958). Liquidity preference as behavior towards risk. *Review of Economic Studies, 25*, 65–86.

Todorova, N. (2012). Volatility estimators based on daily price ranges versus the realized range. *Applied Financial Economics, 22*, 215–229.

Topaloglou, N., Vladimirou, H., & Zenios, S. A. (2002). CVaR models with selective hedging for international asset allocation. *Journal of Banking and Finance, 26*, 1535–1561.

Treacy, W. F., & Carey, M. S. (2000). Credit Risk Rating Systems at Large US Banks. *Journal of Banking & Finance, 24*(1–2), 167–201.

Treynor, J. L., & Mazuy, K. K. (1966). Can mutual funds outguess the market? *Harvard Business Review, 44*, 131–136.

Treynor, J. L. (1965). How to rate management of investment funds. Harvard Business Review, January-February, 63–75.

Treynor, J. L., & Black, F. (1973). How to use security analysis to improve portfolio selection. *Journal of Business, 46*, 66–86.

Treynor, J. L., & Mazuy, K. K. Can Mutual Funds Outguess the Markets. Harward Business Review, 1066, Vol. 44, 131–136.

Triantis, A., & Borison, A. (2001). Real Options: State of the Practice. *Journal of Applied Corporate Finance, 14*, 8–24.

Trigeorgis, L. (1993). The Nature of Option Interactions and the Valuation of Investments with Multiple Real Options. *Journal of Financial and Quantitative Analysis, 28*, 1–20.

Trivedi, P. K., & Zimmer, D. A. (2007). *Copula modeling: An introduction for practitioners*. Hanover, Massachusetts: Now Publishers Inc.

Trivedi, P. K., & Zimmer, D. M. (2006). Using Trivariate Copulas to Model Sample Selection and Treatment Effects: Application to Family Health Care Demand. *Journal of Business and Economic Statistics, 24*(1), 63–76.

Trolle, A. B., & Schwartz, E. S. (2009). A General Stochastic Volatility Model for the Pricing of Interest Rate Derivatives. *Review of Financial Studies, 22*(5), 2007–2057.

Trueman, B. (1986). Why Do Managers Voluntarily Release Earnings Forecasts. *Journal of Accounting and Economics, 8*, 53–71.

Trueman, B. (1990). Theories of earnings-announcement timing. *Journal of Accounting and Economics, 13*, 285–301.

Trueman, B. (1994). Analyst Forecasts and herding behavior. *Review of Financial Studies, 7*, 97–124.

Tsafack, G., & Garcia, R. (2011). Dependence Structure and Extreme Comovements in International Equity and Bond Markets. *Journal of Banking and Finance, 35*, 1954–1970.

Tsai, G. M. (2005). Alternative dynamic capital asset pricing models: Theories and empirical results. Ph.D. Dissertation, Rutgers University.

Tsai, R. S. (2005). Analysis of Financial Time Series. 2nd Edition. Wiley.

Tsaig, Y., Levy, A., & Wang, Y. (2011). Analysing the impact of credit migration in a portfolio setting. *Journal of Banking and Finance, 35*, 3145–3157.

Tsay, R. C., Rena, D., & Pankratz, A. (2000). Outliers in Multivariate Time Series. *Biometrica, 87*, 789–804.

Tsay, R. S., Lin, Y. M., & Wang, H. W. (2008). Residual income, value-relevant information and equity valuation: a simultaneous equations approach. *Review of Quantitative Finance and Accounting, 31*, 331–358.

Tsay, R. S. (2005). Analysis of financial time series. Wiley-Interscience.

Tsay, R. S. (1998). Testing and modeling multivariate threshold models. *Journal of American Statistical Association, 93*(443), 1188–1202.

Turnbull, S., & Milne, F. (1991). A simple approach to interest-rate option pricing. *The Review of Financial Studies, 4*(1), 87–120.

Turner, A. L. (1987). Insurance in an Equilibrium Asset-Pricing Model. In J. D. Cummins & S. E. Harrington (Eds.), *Fair Rate of Return in Property-Liability Insurance*. Norwell, MA: Kluwer Academic Publishers.

Turner, C. M., Startz, R., & Nelson, C. R. (1989). A Markov model of heteroskedasticity, risk, and learning in the stock market. *Journal of Financial Economics, 25*, 3–22.

Turner, P. (2003). Bond market development: what are the policy issues? In R. E. Litan, M. Pomerleano, & V. Sundararajan (Eds.), *The future of domestic capital markets in developing countries*. Washington: Brookings Institute Press.

Turney, S. T. (1990). Deterministic and stochastic dynamic adjustment of capital investment budgets. *Mathematical Computation and Modeling, 13*, 1–9.

Turnovsky, S. (1970). Empirical evidence on the formation of price expectations. *Journal of the American Statistical Association, 65*, 1441–1454.

Tversky, A., & Kahnemann, D. (1986). Rational choice and the framing of decisions. *Journal of Business, 59*, 251–278.

Uchida, H., & Nakagawa, R. (2007). Herd Behavior in the Japanese Loan Market: Evidence from Bank Panel Data. *Journal of Financial Intermediation, 16*, 555–583.

Unal, H., & Kane, E. J., (1987). Two Approaches to Assessing the Interest-Rate Sensitivity of Deposit-Institutions' Equity Returns. In Research in Finance, 7, Jai Press Inc., Greenwich Conn., 113–132.

Upper, C. (2000). How safe was the "safe-haven"? Financial market liquidity during the 1998 turbulences. Discussion paper 1/00, Deutsche Bundesbank.

Van Der Klaauw, W. (2002). Estimating the effect of financial aid offers on college enrollment: A regression-discontinuity approach. *International Economic Review, 43*(4), 1249–1287.

Vanderbei, R. J. (2001). Linear Programming: Foundations and Extensions. Kluwer Academic Publishers, 2nd edition.

Vasicek, O. A. (1977). An Equilibrium Characterization of the Term Structure. *Journal of Financial Economics, 5*, 177–188.

Vasicek, O. A. (1997). The loan loss distribution, Technical Report, KMV Corporation.

Vayanos, D., & Weill, P. (2008). A search-based theory of the on-the-run phenomenon. *The Journal of Finance, 63*, 1361–1398.

Vazza, D., Aurora, D., & Schneck, R. (2005a). *Crossover credit: A 24-year study of fallen angel rating behavior. Global Fixed Income Research*. New York: Standard & Poor's.

Vazza, D., Leung, E., Alsati, M., & Katz, M. (2005). *CreditWatch and ratings outlooks: Valuable predictors of rating behaviour. Global Fixed Income Research (May)*. New York: Standard & Poor's.

Vellekoop, M., & Nieuwenhuis, J. (2006). Efficient pricing of derivatives on assets with discrete dividends. *Applied Mathematical Finance, 13*, 265–284.

Venezian, E. C. (1984). Are Insurers Under-Earning? *Journal of Risk and Insurance, 51*, 150–156.

Verbeek, M. (2004). *A guide to modern econometrics*. Hoboken, New Jersey: John Wiley & Sons.

Vercher, E., Bermudez, J. D., & Segura, J. V. (2007). Fuzzy portfolio optimization under downside risk measures. *Fuzzy Sets and Systems, 158*, 769–782.

Viceira, L. (2001). Optimal portfolio choice for long-horizon investors with nontradable labor income. *Journal of Finance, 56*, 433–470.

Vijh, A. M. (1988). Potential biases from using only trade prices of related securities on different exchanges. *Journal of Finance, 43*, 1049–1055.

Villani, G. (2008). An R&D Investment Game under Uncertainty in Real Option Analysis. *Computational Economics, 32*, 199–219.

Vinod, H. D. (2004). Ranking Mutual Funds Using Unconventional Utility Theory and Stochastic Dominance. *Journal of Empirical Finance, 11*, 353–377.

Viswanathan, S., & Wei, B. (2008). Endogenous events and long-run returns. *Review of Financial Studies, 21*, 855–888.

Vlaar, P. J. G. (2000). Value at risk models for Dutch bond portfolios. *Journal of Banking & Finance, 24*, 1131–1154.

Volkman, D. A., & Wohar, M. E. (1995). Determinants of Persistence in Relative Performance of Mutual Funds". *Journal of Financial Research, 18*, 415–430.

Von Eije, H., & Megginson, W. (2008). Dividends and share repurchases in the European Union. *Journal of Financial Economics, 89*, 347–374.

von Neumann, J., & Morgenstern, O. (1944). *Theory of Games and Economic Behavior*. Princeton: Princeton University Press.

Voronkova, S., & Bohl, M. (2005). Institutional Traders' Behavior in an Emerging Stock Market: Empirical Evidence on Polish Pension Fund Investors. *Journal of Business Finance and Accounting, 32*(7–8), 1537–1560.

Voth, H. J., & Temin, P. (2003). Riding the South Sea Bubble. MIT Economics Working Paper No. 04-02.

Vuong, Q. (1989). Likelihood Ratio Tests for Model Selection and Non-Nested Hypotheses. *Econometrica, 57*, 307–333.

Vuorenmaa, T. (2005). A wavelet analysis of scaling laws and long-memory in stock market volatility. Bank of Finland Research Discussion Paper.

Wade, M., & Hulland, J. Review: The resource-based view and information systems research: review, extension, and suggestions for future research. MIS Quarterly, 28, 107–142.

Wahal, S., & McConnell, J. J. (2000). Do institutional investors exacerbate managerial myopia. *Journal of Corporate Finance: Contracting, Governance and Organization, 6*, 307–29.

Wahal, S., & Wang, A. Y. (2010). Competition among Mutual Funds. Journal of Financial Economics, March, source: http://ssrn.com/abstract=1130822.

Walkenbach, J. (2010). *Excel 2010 Power Programming with VBA*. Indianapolis: Wiley Publishing, Inc.

Walker, M. B. (2006). Credit Default Swaps with Counterparty Risk: A Calibrated Markov Model. *Journal of Credit Risk, 2*(1), 31–49.

Wallace, H. A. (1926). Comparative Farmland Values in Iowa. *Journal of Land and Public Utility Economics, 2*, 385–392.

Wallingford, B. A., II. (1972a). An inter-temporal approach to the optimization of dividend policy with predetermined investments. *Journal of Finance, 27*, 627–635.

Wallingford, B. A., II. (1972b). A correction to "An inter-temporal approach to the optimization of dividend policy with predetermined investments". *Journal of Finance, 27*, 944–945.

Wallingford, B. A., II. (1974). An inter-temporal approach to the optimization of dividend policy with predetermined investments: Reply. *Journal of Finance, 29*, 264–266.

Walter, A., & Weber, F. (2006). Herding in the German Mutual Fund Industry. *European Financial Management, 12*(3), 375–406.

Wang, D. (2008). Herd behavior towards the market index: evidence from 21 financial markets. IESE Business School Working paper. Available at: http://ssrn.com [Accessed 27 December 2011]

Wang, G., & Yau, J. (2000). Trading volume, bid-ask spread, and price volatility in futures markets. *Journal of Futures Markets, 20*, 943–970.

Wang, H. (2009). Forward regression for ultra-high dimensional variable screening. *Journal of the American Statistical Association, 104*, 1512–1524.

Wang, H., Li, R., & Tsai, C. L. (2007). Tuning parameter selectors for the smoothly clipped absolute deviation method. *Biometrika, 94*, 553–568.

Wang, J., & Zivot, E. (2000). A Bayesian time series model of multiple structural changes in level, trend, and variance. *Journal of Business and Economic Statistics, 18*, 374–386.

Wang, J. J., Wang, J., Zhang, Z., & Guo, S. (2012a). Stock index forecasting based on a hybrid mode. *OMEGA, 40*, 758–766.

Wang, K. L., Fawson, C., Barrett, C. B., & McDonald, J. B. (2001). A Flexible Parametric GARCH Model with an Application to Exchange Rates. *Journal of Applied Econometrics, 16*, 521–536.
Wang, L.-R., & Shen, C. H. (1999). Do foreign investments affect foreign exchange and stock markets? - The case of Taiwan. *Applied Economics, 31*(11), 1303–1314.
Wang, S. S., & Yong, V. R. (1998). Ordering risks: expected utility versus Yaari's dual theory of risk. *Insurance Mathematics and Economics, 22*, 145–161.
Wang, S. S., Yong, V. R., & Panjer, H. H. (1997). Axiomatic characterization of insurance prices. *Insurance Mathematics and Economics, 21*, 173–183.
Wang, S. Y., & Lee, C. F. (2011). Fuzzy Multi-Criteria Decision Making for Evaluating Mutual Fund Strategies. *Applied Economics, 43*(24), 3405–3414.
Wang, Y. (1995). Jump and sharp cusp detection by wavelets. *Biometrika, 82*(2), 385–397.
Wang, Y. (1998). Change curve estimation via wavelets. Journal of the American Statistical Association, 93(441).
Wang, Y. J., Chung, H., & Guo, J. H. (2012). A value-at-risk analysis of carry trades using skew-GARCH models. Studies in Nonlinear Dynamics & Econometrics, (in press).
Wansbeek, T. J. (2001). GMM Estimation in Panel Data Models with Measurement Error. *Journal of Econometrics, 104*, 259–268.
Warfield, T., Wild, J. J., & Wild, K. (1995). Managerial ownership, accounting choices, and informativeness of earnings. *Journal of Accounting and Economics, 20*, 61–91.
Warga, A. (1992). Bond returns, liquidity, and missing data. *Journal of Financial and Quantitative Analysis, 27*, 605–617.
Warren, J. M., & Shelton, J. P. (1971). A Simultaneous Equation Approach to Financial Planning. *Journal of Finance, 26*(5), 1123–1142.
Washburn, C. L., & Binkley, C. S. (1990). On the use of period-average stumpage prices to estimate forest asset pricing models. *Land Economics, 66*, 379–393.
Washburn, C. L., & Binkley, C. S. (1993). Do forest assets hedge inflation? *Land Economics, 69*, 215–224.
Washburn, C. (2008). *Forestry investments in a slowing global economy, International Forest Investment (London, UK)*. Athens, GA: Center for Forest Business.
Wasserman, L., & Roeder, K. (2009). High-dimensional variable selection. *Annals of Statistics, 37*, 2178–2201.
Watson, D. J. H., & Baumler, J. V. (1975). Transfer pricing: a behavioral context. The Accounting Review, 466–474.
Watson, M. (1994). Vector autoregressions and cointegration. In R. F. Engle & D. McFadden (Eds.), *Handbook of Econometrics* (Vol. IV). Amsterdam: Elsevier Science Ltd.
Watts, R. L., & Zimmerman, J. L. (1979). The demand for and supply of accounting theories: The Market for Excuses. *Accounting Review, 54*, 273–306.
Waymire, G. (1986). Additional evidence on the accuracy of analyst forecasts before and after voluntary managed earnings forecasts. *Accounting Review, 61*, 129–141.
Wei, G. C. G., & Tanner, M. A. (1990). A Monte Carlo implementation of the EM algorithm and the Poor Man's data augmentation algorithms. *Journal of the American Statistical Association, 85*, 699–704.
Wei, K. C. (1984). The arbitrage pricing theory versus the generalized intertemporal capital asset pricing model: theory and empirical evidence. Ph.D. Dissertation, University of Illinois at Urbana-Champaign.
Wei, K. C. J. (1988). An Asset-Pricing Theory Unifying the CAPM and APT. *Journal of Finance, 43*(4), 881–892.
Wei, L. J., Lin, D. Y., & Weissfeld, L. (1989). Regression analysis of multivariate incomplete failure time data by modeling marginal distributions. *Journal of The American Statistical Association, 84*, 1065–1073.
Wei, W.W. (2006). Time Series Analysis: Univariate and Multivariate Methods, Pearson Education, Inc.

Wei, Z., Xie, F., & Zhang, S. (2005). Ownership structure and firm value in China's privatized firms: 1991-2001. *Journal of Financial and Quantitative Analysis, 40*, 87–108.

Weiner, R. J. (2006). Do Birds of a Feather Flock Together? Speculator Herding in the World Oil Market. Discussion paper. Resources for the Futures, RFF DP 06-31. Available at: http://ideas.repec.org/p/rff/dpaper/dp-06-31.html [Accessed 27 December 2011].

Weingartner, H. M. (1963). *Mathematical programming and the analysis of capital budgeting problems*. Englewood Cliffs, NJ: Prentice-Hall.

Weisbach, M. (1988). Outside directors and CEO turnover. *Journal of Financial Economics, 20*, 431–460.

Weiss, K. (1989). The Post-Offering Price Performance of Closed-end Funds. *Financial Management, 18*, 57–67.

Welch, G., & Bishop, G. (1995). *An introduction to the Kalman filter*. Chapel Hill, NC: University of North Carolina at Chapel Hill.

Welch, I., & Goyal, A. (2008). A comprehensive look at the empirical performance of equity premium prediction. *Review of Financial Studies, 21*, 1455–1508.

Welch, P. (1967). The use of fast Fourier transforms for the estimation of power spectra: A method based on time averaging over short modified periodograms. *IEEE Transactions on Audio and Electroacoustics, 15*, 70–73.

Wells, E., & Harshbarger, S. (1997). *Microsoft Excel 97 Developer's Handbook*. Redmond: Microsoft Press.

Wermers, R. (1999). Mutual fund herding and the impact on stock prices. *Journal of Finance, 54*, 581–622.

Weron, R. (2004). *Handbook of Computational Statistics*. Berlin: Springer.

West, K. (1987). A specification test for speculative bubbles. *Quarterly Journal of Economics, 102*, 553–580.

West, K. (1988). Bubbles, fads, and stock price volatility tests: A partial evaluation. *Journal of Finance, 43*, 639–655.

Weston, J. F., & Chung, K. S. (1983). Do Merger Make Money? A Reach Summary. *Mergers& Acquisitions, 18*, 40–48.

Weston, J. F., & Mansinghka, S. K. (1971). Test of the Efficiency: Performance of Conglorate Firms. *The Journal of Finance, 26*, 919–939.

Whaley, R. (1982). Valuation of American Call Options on Dividend Paying Stocks. *Journal of Financial Economics, 10*, 29–58.

White, E. (1990a). The stock market boom and crash of 1929 revisited. *Journal of Economic Perspectives, 4*, 67–84.

White, J. A. (1990b). More institutional investors selling short: but tactic is part of a wider strategy. *Wall Street Journal, 7*, 493–501.

White, H. (1980a). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica, 48*(4), 817–838.

White, H. (1980b). A heteroskedasticity-consistent covariance matrix estimator and a direct test for heteroskedasticity. *Econometrica, 48*, 817–838.

White, H. (1982). Maximum Likelihood Estimation of Misspecified Models. *Econometrica, 50*, 1–25.

White, H. (2000). A Reality Check for Data Snooping. *Econometrica, 68*, 1097–1126.

Whited, T. (2001). Is It Inefficient Investment that Causes the Diversification Discount? *Journal of Finance, 56*, 1667–1691.

Whited, T. (2006). External Finance Constraints and the Intertemporal Pattern of Intermittent Investment. *Journal of Financial Economics, 81*, 467–502.

Whitelaw, R. F. (1994). Time variations and covariations in the expectation and volatility of stock market returns. *Journal of Finance, 49*, 515–541.

Whitelaw, R. F. (2000). Stock market risk and return: An equilibrium approach. *Review of Financial Studies, 13*, 521–547.

Whitmore, G. A., & Findlay, M. C. (Eds.). (1978). *Stochastic Dominance: An Approach to Decision-Making Under Risk*. Lexington, MA: D.C.Heath.

Wiggins, J. (1991). Empirical tests of the bias and efficiency of the extreme-value variance estimator for common stocks. *Journal of Business, 64*, 417–432.
Wiggins, J. (1987). Option Values Under Stochastic Volatilities. *Journal of Financial Economics, 19*, 351–372.
Wild, C. J., & Seber, G. A. F. (1999). Chance Encounters: A First Course in Data Analysis and Inference. John Wiley and Sons Ltd.
Wilde, T. (1997). CreditRisk+: a credit risk management framework. Credit Suisse FirstBoston.
Willamson, O. E. (1981). The Modern Corporation: Origins, Evolution, Attributes. *Journal of Economic Literature, 14*, 1537–1568.
Williamson, O. E. (1988). Corporate finance and corporate governance. *Journal of Finance, 43*, 567–592.
Windmeijer, F. (2005). A finite sample correction for the variance of linear efficient two-step GMM estimators. *Journal of Econometrics, 126*, 25–51.
Winker, R. L. (1989). Combining forecasts: a philosophical basis and some current issues. *International Journal of Forecasts, 5*, 605–608.
Winkler, R. L. (1996). Scoring rules and the evaluation of probabilities. *Test, 5*(1), 1–60.
Winkler, R. L., & Makridakis, S. (1983). The combination of forecasts. *Journal of the Royal Statistical Society, A, 146*, 150–157.
Winkler, R. L. (2003). Introduction to Bayesian Inference and Decision, Probabilistic Publishing.
Wintoki, M. B., Linck, J. S., & Netter, J. M. Endogeneity and the dynamics of internal corporate governance. Journal of Financial Economics. Forthcoming.
Wirl, F. (2008). Optimal Maintenance and Scrapping versus the Value of Back ups. *Computational Management Science, 5*, 379–392.
Wolf, A. (1982). Fundamentals of commodity options on futures. *Journal of Futures Markets, 2*(4), 391–408.
Womack, K. (1996). Do brokerage analysts' recommendations have investment value? *Journal of Finance, 54*, 137–157.
Wong, W. K., Manzur, M., & Chew, B. K. (2003). How rewarding is technical analysis? Evidence from Singapore stock market. *Applied Financial Economics, 13*, 543–551.
Wong, E. (1960). The Construction of a Class of Stationary Markov Processes. In R. Bellman (Ed.), *Sixtenn Symposia in Applied Mathematics: Stochastic Processes in Mathematical Physics and Engineering*. Providence, RI: American Mathematical Society.
Wong, W. K. (2006). Stochastic dominance theory for location-scale family. Journal of Applied Mathematics and Decision Sciences 2006, 1–10.
Wong, W. K. (2007). Stochastic dominance and mean-variance measures of profit and loss for business planning and investment. *European Journal of Operational Research, 182*, 829–843.
Wong, W. K., & Li, C. K. (1999). A note on convex stochastic dominance theory. *Economics Letters, 62*, 293–300.
Wong, W. K., & Ma, C. (2008). Preferences over location-scale family. *Economic Theory, 37*, 119–146.
Wong, W. K., & Bian, G. (2000). Robust Bayesian inference in asset pricing estimation. *Journal of Applied Mathematics & Decision Sciences, 4*, 65–82.
Wong, W. K., & Miller, R. B. (1990). Analysis of ARIMA-noise models with repeated time series. *Journal of Business and Economic Statistics, 8*, 243–250.
Wong, W. K., & Chan, R. (2004). The estimation of the cost of capital and its reliability. *Quantitative Finance, 4*, 365–372.
Wong, W. K., & Chan, R. (2008). Markowitz and prospect stochastic dominances. *Annals of Finance, 4*, 105–129.
Wong, W. K., Chew, B. K., & Sikorski, D. (2001). Can P/E ratio and bond yield be used to beat stock markets? *Multinational Finance Journal, 5*, 59–86.
Wong, W. K., Thompson, H. E., Wei, S., & Chow, Y. F. (2006). Do Winners perform better than losers? A stochastic dominance approach. *Advances in Quantitative Analysis of Finance and Accounting, 4*, 219–254.

Wong, W. K., Wright, J. A., Yam, S. C. P., & Yung, S. P. (2012). A mixed Sharpe ratio. *Risk and Decision Analysis, 3*(1–2), 37–65.

Wong, W. K., Phoon, K. F., & Lean, H. H. (2008). Stochastic dominance analysis of Asian hedge funds. *Pacific-Basin Finance Journal, 16*, 204–223.

Wongswan, J. (2006). Transmission of information across international equity markets. *Review of Financial Studies, 19*, 1157–1189.

Wooldridge, J. M. (2002). *Econometric Analysis of Cross Section and Panel Data*. Cambridge, Massachusetts: MIT Press.

Wooldridge, J. M. (2010). Econometric Analysis of Cross Section and Panel Data, 2nd Edition. MIT Press.

Woolridge, J. (2009). Introductory Econometrics: A Modern Approach, 4th Edition. Cengage Learning.

World Bank. (2000a). Global Development Finance 2000. Washington, D.C.

World Bank. (2000b). World Development Indicators 2000. Washington, D.C.

Wu, C. (1990). On The Asymptotic Properties of the Jackknife Histogram. *Annals of Statistics, 18*, 1438–1452.

Wu, C., & Liang, S. (2011). The economic value of range-based covariance between stock and bond returns with dynamic copulas. *Journal of Empirical Finance, 18*, 711–727.

Wu, C. F. G. (1986). Jackknife, bootstrap and other resampling methods in regression analysis. *Annals of Statistics, 14*, 1261–1295.

Wu, C.-R., Lin, C.-T., & Tsai, P.-H. (2011a). Financial service sector performance measurement model: AHP sensitivity analysis and balanced scorecard approach. *The Service Industries Journal, 31*, 695.

Wu, D. (1973). Alternative tests of independence between stochastic regressors and disturbances. *Econometrica, 41*, 733–750.

Wu, G., & Xiao, Z. (2002). Are there speculative bubbles in stock markets? Evidence from an alternative approach. Preprint.

Wu, G., & Xiao, Z. (2002). An Analysis of Risk Measures. *Journal of Risk, 4*(4), 53–75.

Wu, J. W., Hung, W. L., & Lee, H. M. (2000). Some Moments and Limit Behaviors of the Generalized Logistic Distribution with Applications. Proceedings of the National Science Council ROC, 24, 7–14.

Wu, P. C. (2010). Applying a Factor Copula to Value Basket Credit Linked Notes with Issuer Default Risk. *Finance Research Letters, 7*(3), 178–183.

Wu, P. C., Kao, L. J., & Lee, C. W. (2011b). How Issuer Default Risk Affects Basket Credit Linked Note Coupon Rate. *International Journal of Information and Management Sciences, 22*(1), 59–71.

Wylie, S. (2005). Fund Manager Herding: A test of the accuracy of Empirical Results using U.K. Data. *The Journal of Business, 78*(1), 381–403.

Xiao, Z., & Phillips, P. C. B. (2002). A CUSUM test for cointegration using regression residuals. *Journal of Econometrics, 108*, 43–61.

Xiao, Z., & Koenker, R. (2009). Conditional Quantile Estimation and Inference for GARCH Models. *Journal of the American Statistical Association, 104*(488), 1696–1712.

Xiao, Z., Shao, J., & Palta, M. (2008). A Unified Theory for GMM Estimation in Panel Data Models with Measurement Error. Working paper, University of Wisconsin Madison.

Xie, B., Davidson, W., & DaDalt, P. (2003). Earnings management and corporate governance: The role of the board and the audit committee. *Journal of Corporate Finance, 9*, 295–316.

Xu, J. (1992). The effect of management earnings forecast on portfolio management. Unpublished thesis, Accounting Department, College of Politics, Taiwan.

Xu, X. (1990). A examination of analysts earnings forecasts accuracy. Unpublished thesis, Accounting Department, Politics University.

Xu, Y. (2004). Incorporating Estimation Risk in Copula-based Portfolio Optimization. Department of Economics, University of Miami.

Yaari, M. E. (1987). The dual theory of choice under risk. *Econometrica, 55*, 95–115.
Yajima, Y. (1988). On estimation of a regression model with long-momery stationary errors. *Ann Statist, 16*, 791–807.
Yamaguchi, K. (1991). *Event History Analysis*. Beverly Hills, California: Sage Publications.
Yan, A. (2006). Leasing and debt financing: substitutes or complements? *Journal of Financial and Quantitative Analysis, 41*, 709–731.
Yang, D., & Zhang, Q. (2000). Drift-independent volatility estimation based on high, low, open, and closing Prices. *Journal of Business, 73*, 477–491.
Yang, C. C. (1989). The impact of new equity financing on firms' investment, dividend and bebt-financing decisions. Ph.D. Dissertation, The University of Illinois at Urbana-Champaign.
Yang, C. C., Lee, C. F., Gu, Y. X., & Lee, Y. W. (2010). Co-determination of capital structure and stock returns—A LISREL approach: An empirical test of Taiwan stock markets. *Quarterly Review of Economic and Finance, 50*(2), 222–233.
Yang, T. C. (2005). The pricing of credit risk of portfolio base on listed corporation in Taiwan market. Master Thesis at National Tsing Hua University in Taiwan.
Yao, Y. C. (1984). Estimation of a noisy discrete-time step function: Bayes and empirical Bayes approaches. *The Annals of Statistics, 12*, 1434–1447.
Yao, Y., Partington, G., & Stevenson, M. (2005). Run length and the predictability of stock price reversals. *Journal of Accounting and Finance, 45*, 653–671.
Hoshino, Y. (1982). The Performance of Corporate Mergers in Japan. *Journal of Business Financial and Accounting, 9*, 145–163.
Yates, J. (1982). External correspondence: Decompositions of the mean probability score. *Organizational Behaviour and Human Performance, 30*, 132–156.
Yee, K. K. (2008). A Bayesian framework for combining valuation estimates. *Review of Quantitative Finance and Accounting, 30*, 339–354.
Yeh, S. K., & Lin, B. H. (2003). Term structure fitting and information content: An empirical study on the Taiwanese government bond market. *Review of Pacific Basin Financial Markets and Policies, 6*(3), 1–44.
Yeh, J., Wang, J., & Kuan, C. (2009). A noise-robust estimator of volatility based on interquantile ranges. Working paper.
Yelick, K., Coghlan, S., Draney, B. & Canon, R. S. (2011). The Magellan Report on Cloud Computing for Science. ASCR, DOE.
Yermack, D. (1997). Good timing: CEO stock option awards and company news announcements. *Journal of Finance, 52*(2), 449–476.
Yie, M. (1996). Examination on factors in management and financial analyst earnings forecasting accuracy. Unpublished thesis, Accounting Department, Zhong Zheng University.
Yoon, K. C. (2006). Relative Portfolio performance evaluation and incentive strcture. *Journal of Business, 79*(2), 903–921.
Yoon, P. S., & Starks, L. T. (1995). Signaling, investment opportunities, and dividend announcements. *Review of Financial Studies, 8*, 995–1018.
Yourougou, P. (1990). Interest Rate and the Pricing of Depository Financial Intermediary Common Stock: Empirical Evidence. *Journal of Banking and Finance, 14*, 803–320.
Yu, F. (2005). Correlated default in reduced-form models. *Journal of Investment Management, 3*(4), 16–27.
Yu, J., & Yuan, Y. (2011). Investor sentiment and the mean–variance relation. *Journal of Financial Economics, 100*, 367–381.
Yu, J. R., & Lee, W. Y. (2011). Portfolio rebalancing model using multiple criteria. *European Journal of Operational Research, 209*, 166–175.
Yu, J. (2010). Simulation-based Estimation Methods for Financial Time Series Models. In J.-C. Duan, J. E. Gentle and W. Hardle (Eds.), *Handbook of Computational Finance*. Springer.
Yu, K., & Jones, M. C. (1998). Local linear quantile regression. *Journal of the American Statistical Association, 98*, 228–237.

Yu, M., Takahashi, S., Inoue, H., & Wang, S. (2010). Dynamic portfolio optimization with risk control for absolute deviation model. *Journal of Operational Research, 201*, 349–364.

Yu, P. L. (1985). *Multiple criteria decision making: concepts, techniques and extensions*. New York: Plenum.

Yu, Y. L., & Zeleny, M. (1975). The set of all nondominated solutions in the linear cases and a multicriteria simplex method. *Journal of Mathematical Analysis and Applications, 49*, 430–458.

Yuan, L., & Zhongjie, X. (2000). The wavelet detection of the jump and cusp points of a regression function. *Acta Mathematicae Applicatae Sinica (English Series), 16*(3), 283–291.

Yunker, P. J. (1983). Survey study of subsidiary, autonomy, performance evaluation and transfer pricing in multinational corporations. Columbia Journal of World Business, 51–63.

Yan, Y. (1999). Measuring the Timing Ability of Mutual Fund Managers. *Annals of Operations Research*, 233–243.

Zacharatos, N., & Sutcliffe, C. (2002). Is the forward rate for the greek drachma unibased? a VECM analysis with both overlap43 Rationality and Heterogeneity of Yen-Dollar Survey Forecasts 29 ping and non-overlapping data. *Journal of Financial Management and Analysis, 15*, 27–37.

Zaffaroni, P. (2000). Contemporaneous aggregation of GARCH processes. Econometrics Discussion Paper 378, LSE.

Zaffaroni, P. (2006). Aggregation and memory of models of changing volatility. Journal of Econometrics 136, 237–249.

Zarnowitz, V. (1985). Rational expectations and macroeconomic forecasts. *Journal of Business & Economic Statistics, 3*, 293–311.

Zavoina, R., & McKelvey, W. (1975). A statistical model for the analysis of ordinal level dependent variables. *Journal of Mathematical Sociology, 4*, 103–120.

Zeileis, A., Leisc, F., Hornik, K., & Kleiber, C. (2002). Strucchange: An R package for testing for structural change in linear regression models. *Journal of Statistical Software, 7*, 1–38.

Zeleny, M. (1974). *Linear multi-objective programming*. Berlin: Springer-Verlag.

Zellner, A. (1970). Estimation of regression relationships counting un-observable variables. *International Economic Review, 11*, 441–454.

Zellner, A. (1986). Biased predictors, rationality and the evaluation of forecasts. *Economics Letters, 21*, 45–48.

Zhang, P. G. (1994a). Bounds for option prices and expected payoffs. *Review of Quantitative Finance and Accounting, 4*(2), 179–197.

Zhang, C. (2010). Nearly unbiased variable selection under minimax concave penalty. *Annals of Statistics, 38*, 894–942.

Zhang, F. (2006a). Information uncertainty and analyst forecast behavior. *Contemporary Accounting Research, 23*, 1–28.

Zhang, F. (2006b). Information uncertainty and stock returns. *Journal of Finance, 61*, 105–137.

Zhang, J. E., & Zhu, Y. (2006). VIX Futures. *The Journal of Futures Markets, 26*(6), 521–531.

Zhang, L., & Mykland, P. (2005). A tale of two time scales: Determining integrated volatility with noise high frequency data. *Journal of American Statistics, 100*, 1394–1411.

Zhang, P. (1994b). Bounds for option prices and expected payoffs. *Review of Quantitative Finance and Accounting, 4*, 179–197.

Zhang, S. S. (2011). Andreas Storkenmaier, Martin Wagener, and Christof Weinhardt. “The Quality of Electronic Markets”, 44th Hawaii International Conference on System Sciences, pp. 1–10.

Zhang, X. F. (2010). High Frequency Trading, Stock Volatility, and Price Discovery. Yale University, Working Paper.

Zhou, C. (2001). An analysis of default correlations and multiple defaults. *Review of Financial Studies, 14*(2), 555–576.

Zhou, J., & He, X. (2008). Dimension reduction based on constrained canonical correlation and variable filtering. *Annals of Statistics, 36*, 1649–1668.

Zhou, T., & Lai, R. (2009). Herding and information based trading. *Journal of Empirical Finance, 16*, 388–393.

Zhou, X. (2009). Information-based trading in the junk bond market, Working paper, Cornell University.

Zhu, L., Li, L., Li, R., & Zhu, L. (2011). Model-free feature screening for ultrahigh dimensional data. Journal of the American Statistical Association, to appear.

Zimmermann, H. J. (1978). Fuzzy programming and linear programming with several objective functions. *Fuzzy Sets and Systems, 1*, 44–45.

Zinkhan, F. C. (1988). Forestry projects, modern portfolio theory, and discount rate selection. *Southern Journal of Applied Forestry, 12*, 132–135.

Zinkhan, F. C. (2008). History of modern timberland investments. International Forest Investment: An Executive Education Conference (Munich, Germany). Center for Forest Business, Athens, GA.

Zinkhan, F. C., & Cubbage, F. W. (2003). Financial analysis of timber investments. In E. O. Sills & K. L. Abt (Eds.), *Forests in a Market Economy*. Dordrecht, The Netherlands: Kluwer Academic Publishers.

Zitzewitz, E. (2002). Regulation Fair Disclosure and the private information of analysts. Working Paper, Stanford University.

Zivot, E. (2000). Cointegration and forward and spot exchange rate regressions. *Journal of International Money and Finance, 19*, 785–812.

Zmijewski, M. E. (1984). Methodological issues related to the estimation of financial distress prediction models. *Journal of Accounting Research, 22*, 59–68.

Zou, H. (2006). The adaptive lasso and its oracle properties. *Journal of the American Statistical Association, 101*, 1418–1429.

Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society, Series B, 67*, 301–320.

Zou, H., & Li, R. (2008). One-step sparse estimates in nonconcave penalized likelihood models. *Annals of Statistics, 36*, 1509–1533.

Zou, H., & Zhang, H. H. (2009). On the adaptive elastic-net with a diverging number of parameters. *Annals of Statistics, 37*, 1733–1751.